

Transportation (2010) 37:1011–1030
DOI 10.1007/s11116-010-9281-1

State-of-the art of freight forecast modeling: lessons learned and the road ahead

Joseph Y. J. Chow · Choon Heon Yang · Amelia C. Regan

Published online: 17 June 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract In this paper we review freight forecasting models and current advances and needs with respect to data and model development. We then present a case study to suggest which models should be developed for the State of California in the US. We suggest several alternatives including an aggregate commodity flow model, a disaggregate regional logistics model and a hybrid regional logistics model with a truck touring model. We point out however, that the data requirements for the latter model would be extensive. In addition, the development of hybrid models, for example progress in the integration of regional logistics models with urban truck touring models, will introduce new problems such as reconciling the outputs of multiple models for consistency.

Keywords Freight demand models · Commodity flow · Truck tours · Supply chain · Logistics · Activity-based · Disaggregate

Motivation

Continuing population growth and vigorous economic activity has led to a steep increase in freight movements in transportation networks in the U.S. As freight transportation demand increases, freight and passenger mobility on transportation systems decrease in the absence of accompanying increases in capacity. The Federal Highway Administration (FHWA) estimates domestic freight volumes to grow by more than 65%, increasing from 13.5

J. Y. J. Chow

Department of Civil and Environmental Engineering, Institute of Transportation Studies,
University of California, Irvine, CA 92697, USA

C. H. Yang

Korea Institute of Construction Technology, Highway Research Division 2311,
Daehwa-Dong, Ilsan-Gu, Goyang-Si, Gyeonggi-Do 411-712, Republic of Korea

A. C. Regan (✉)

Department of Computer Science and Institute of Transportation Studies,
University of California, Irvine, CA 92697, USA
e-mail: aregan@uci.edu

billion tons in 1998 to 22.5 billion tons in 2020 (FHWA 2002). According to another report, volumes of goods shipped by trucks and railroads are projected to increase by 98% and 88%, respectively, by 2035 (US-GAO 2007). This growth in freight transportation is expected to significantly outpace growth in passenger transportation (RITA 2008). An increase in freight movements provides motivation for the need to accurately estimate the movements of goods as well as to forecast the expected future truck and underlying commodity flows.

Our research was originally motivated by the need to clearly identify, compile, and relate the many different data sources available to California for freight demand modeling—and then to generalize our experience to other large metropolitan areas—particularly those where international gateways make goods movement a growing concern. An understanding of freight data sources goes hand in hand with a strong understanding of the models that use them for the different objectives of regional agencies. As part of the effort to compile data sources, this study was conducted to identify the state of the art in freight demand models from the point of view of their objectives and data requirements.

As pointed out by Fischer et al. (2005), California is the number one freight destination by value in the U.S., with over \$800 billion in freight movement. California is also home to the Ports of Long Beach and Los Angeles in the San Pedro Bay, the largest container port in the U.S. and the fifth largest in the world (Giuliano et al. 2007).

Hensher and Puckett (2005) and Giuliano et al. (2007) discuss some of the attributes that make an ideal freight demand model: a strong behavioral foundation; a multimodal scope; incorporating freight and passenger interactions; and capable of handling policy changes. However, many modeling efforts fall far short of meeting these objectives. In the remaining sections, we start by identifying and discussing gaps in the state of practice. Building from the review developed by Regan and Garrido (2001), a literature survey is conducted for subsequent freight demand model development to relate the models to the objectives of regional agencies and their data needs and availability. We use California as a case study throughout these evaluations as a representative region.

State-of-the-practice

The statewide freight forecasting toolkit released by the National Cooperative Highway Research Program (NCHRP 606 2008) gives a comprehensive review of current models used in practice with a primary focus on the US. An exhaustive review of model developments in Europe (de Jong et al. 2004) examines different models and presents very similar findings. In the NCHRP report, five model classes are defined and categorized by components, as shown in Table 1. In that table we also provide specific examples where these models have been applied in the US.

The first class, which we label *Class A* is the direct facility flow factoring method. Link-by-link flows are obtained by applying growth rates to observed truck traffic volumes. The *Class A* model is very straightforward and mainly used for the short-term forecasts of freight volumes on transportation system links.

Class B, the O–D Factoring Method makes use of two additional components of mode split and traffic assignment with the Origin–Destination (O–D) factoring. Unlike *Class A*, *Class B* explicitly considers the O–D travel patterns of commercial vehicles. Ohio developed an interim freight model to provide a clear picture of current and future freight movements on important highway corridors (ODOT 2008).

Table 1 Freight model classes by component (re-created from NCHRP 606 2008)

Model class	Model component					
	Direct factoring	Trip generation	Trip distribution	Mode split	Traffic assignment	Economic/land use modeling
Direct facility flow factoring method	Of facility flows					
O–D factoring method	Of O–D tables			Included	Included	
Truck model		Based on exogenously supplied zonal activity	Included	Not applicable	Included	
Four-step commodity model		Based on exogenously supplied zonal activity	Included	Included	Included	
Economic activity model		Based on outputs of economic model	Included	Included	Included	Included

Class C Truck Models generate aggregate truck trips and assign them to the road network. The Portland Metropolitan Planning Organization developed a truck model called the Tactical Model System (FHWA 2007).

Class D Four-Step Commodity Models are commodity-based versions of the *Class C* vehicle-based models. Several states have well documented statewide freight models which were developed using such a structure. For example, Wisconsin's model predicts both passenger and truck traffic volumes for a network (Proussaloglou et al. 2007). Other states such as Texas, Pennsylvania, Iowa, and Florida have also developed their freight forecasting models based on four-step process commodity models (NCHRP 2008; FHWA 2007; Iowa DOT 2008; PennDOT 2007).

Class E Economic Activity Models are economic land use models that incorporate feedback mechanisms with freight transport costs. Oregon developed a statewide passenger and freight forecasting model based on an economic and land use behavioral model (Hunt et al. 2001).

More detailed descriptions of each model class with specific case studies can be obtained from the NCHRP 606 report. In their comprehensive review paper, de Jong et al. (2004) examine the national and international freight models that have been developed, mainly for application in Europe. Their primary classifications are: trend and time series models, system dynamics models, zonal trip rate models and I/O and related models correspond roughly to our classes A or B, E, C or D and E, respectively.

Gaps in the state-of-practice

Several gaps are mentioned in NCHRP 606 in terms of the analytical needs that are not met by any of the classes discussed earlier. These unmet policy and analytical needs, obtained from the report are shown in Table 2. Note that the acronym STIP stands for the Statewide

Table 2 State needs versus model classes (re-created from NCHRP 606 2008)

	Policy and analytical needs	Type of tool				
		Economic flow factor	O–D factor models	Truck models	Commodity models	Economic-based models
1	State transportation planning	–	P	P	P	P
2	Project prioritization, STIP development	P	S	P	P	P
3	Modal diversion analysis	–	S	–	P	P
4	Pavement, bridge, and safety management	P	S	P	P	P
5	Policy studies	–	–	–	–	–
6	Needs analysis	P	S	P	P	P
7	Commodity flow analysis	–	P	–	P	P
8	Rail planning	–	S	–	P	P
9	Trade corridor and border planning	–	–	–	–	–
10	Operations, safety, security, truck size and weight issues, etc.	–	–	–	–	–
11	Project development or design needs; e.g. forecasts and loadings	P	S	S	S	S
12	Terminal access planning	–	S	–	S	P
13	Truck flow analysis and forecasting	–	S	P	P	P
14	Performance measurement/program evaluation	–	–	–	–	–
15	Bottleneck analysis	–	–	S	S	S

P Primary, *S* secondary

Transportation Improvement Program which is required for every U.S. state seeking federal funding for transportation projects.

The needs listed under 5 (policy studies), 9 (trade corridor planning), 10 (safety, security, operations), and 14 (performance measurement) are not met by any of the five model classes. These insufficiencies in practice agree with the general conclusions in current models made by Liedtke and Schepperle (2004), Hesse and Rodrigue (2004), Friesz and Holguín-Veras (2005), Fischer et al. (2005), among others. Hensher and Figliozzi (2007) emphasize the inadequacies of the four-step modeling approach in dealing with a “21st century global customer-driven economy”, concluding that it is crucial for freight models to account for supply chain relationships and logistics constraints.

Essentially, freight demand models in practice rely on aggregate approaches that are insensitive to economic behavior at the level of the firms that act as the decision-makers. Furthermore, the two primary types of models, vehicle-based or commodity-based, both have flaws. Vehicle-based models such as those in class C fail to model the underlying economic behavior such as commodity flows from which the demand is actually derived. Commodity-based models such as those in class D fail to realistically account for vehicle activities, especially in urban settings, for which evaluation and impact assessment are most crucial (Holguín-Veras and Thorson 2003).

As a result, there has been a recent flurry of new developments toward more disaggregate types of models that incorporate supply chain behavioral mechanics or truck touring aspects. Regan and Garrido (2001) alluded to this trend relatively early on and

supported it with their comprehensive literature survey. The paper by de Jong et al. (2004) had similar conclusions.

Building on that survey, a review of research developments is conducted in two additional classes of freight demand models: Class F Logistics Models and Class G Vehicle Touring Models. As Fischer et al. (2005) concluded, these two categories result from the need to improve on the sensitivity of models to economics of commodities for policymaking (class F) and more realistically capture the movements of vehicles for impact assessment (class G).

What separates our survey from other recent efforts is the identification of regional agency objectives that the models serve and data gaps for implementation. The goal is to supplement the NCHRP 606 by providing both researchers and practitioners with enough general guidelines to construct their own class F, G or hybrid model to suit their needs and data availability. We end our discussion with several recommended alternate models using California as a case study. These reflect the spectrum of data needs and the state-of-the-art in model development.

Class F—logistics models

Logistics models share several common traits. As the name suggests, these models incorporate more than a single origin and destination, often having multiple intermediate stops to represent distribution channels. These models are equipped with behavioral distinctions which apply to the many decision-makers within the chain. Because supply chains involve the movements of raw goods through finished products, class F models generally focus on units of commodities rather than on vehicles.

Despite these commonalities, logistics models can differ significantly from one another. Tavasszy et al. (1998), Liedtke and Schepperle (2004), and de Jong and Ben-Akiva (2007) use aggregate data along with disaggregate logistics choices to generate commodity flows through regional supply chains. On the other hand, the work of Boerkamps et al. (2000), Wisetjindawat and Sano (2003), and Wisetjindawat et al. (2006) focus on urban logistics models derived from purely disaggregate firm choice data.

Regional logistics models

Strategic model for integrated logistic evaluations (SMILE)

In developing their decision support system, Tavasszy et al. (1998) were concerned with answering two primary questions: understanding how socioeconomic trends interact with the performance of logistics and transport systems, and finding ways to measure the performance of the systems for analyzing policy options. They argue that SMILE would enable decision-makers to answer questions such as how transport would be impacted by having a central European distribution.

The SMILE model consists of three levels: Production, Inventory, and Transport. At the Production level, the volumes of commodities produced and consumed are generated by a production function f_1 :

$$\text{Production Chains} = f_1(\text{Make/Use table, demand function, shipment valuedensity, socio - economic factors}) \quad (1)$$

The model relies heavily on the availability of Make/Use tables, which are difficult to obtain at a detailed level. In the U.S., Make/Use tables by industry are available from the

Bureau of Economic Analysis (BEA) at the national level for multiple years from 1997 to 2007. Regional BEA zone level I–O tables by industry can be obtained from the Regional Input–Output Modeling System (RIMS II 1997).

From the production chains and sales and sourcing processes at each location, a spatial distribution based on price differences can be obtained from trade theory, f_2 .

$$\text{Trade Flows} = f_2(\text{Production Chains, logistics costs}) \quad (2)$$

At the second Inventory level, the demand for transport is obtained using two steps. The first step involves obtaining optimal distribution locations, f_3 , given three alternative channels: direct, single distribution center (DC), and two DC's. The choice of which channel to distribute the commodities is based on a multinomial logit model, f_4 , using inventory and transport costs.

$$\text{Distribution Locations} = f_3(\text{Trade Flows, inventory, handling, and transport costs}) \quad (3)$$

$$\text{Distribution Choice} = f_4(\text{Distribution Locations, inventory, handling, transport costs}) \quad (4)$$

A survey is used to obtain product characteristics for determining the logistics costs by commodity for the Inventory level. This type of data can be extremely difficult to obtain as it involves detailed information on handling costs, shipment frequencies, etc.

At the last Transport level, six modes of transportation are considered in a mode choice model, f_5 , using the shortest route per mode for the choice disutility.

$$\text{Mode Choice} = f_5(\text{Distribution Choice, multimodal network, transport costs}) \quad (5)$$

While SMILE incorporates distribution and mode choice in the commodity-based freight demand, its structure suggests behavior attuned to the commodities as opposed to specific agents such as a shipper or carrier. A significant contribution is the inclusion of the second level described by Eqs. 3, 4, although in practicality especially in the U.S. it may be difficult to obtain survey data to estimate the behavior.

Activity-based freight transport model

Liedtke and Schepperle's (2004) freight model has its roots in activity-based passenger demand modeling. Similar to SMILE, their model was developed to better understand the effects of new information technologies on freight transport. The activity-based approach was taken to emphasize the behavioral aspects of freight.

Activity-based modeling is meant to overcome several deficiencies in practice including: crude conversions of commodity flows to vehicular flows, poor explanation for empty vehicles, and inability of aggregate models to forecast impacts due to changes in logistic structures. For example, models in practice are generally unable to address a logistics policy question such as "what if shipment sizes decrease by $x\%$ ".

To fit the approach to freight modeling, the activity can be defined as a freight order. However, Liedtke and Schepperle show that real data for a full freight activity chain is not possible because of the many dimensions of uses and actors present. Instead, they combine two classification methods—the Classification of Products by Activity (CPA) (RAMON 2008) and the Standard Goods Classification for Transport Statistics (NST/R) to obtain the necessary schema.

Employment information from the CPA is used to obtain annual production in tons per employee by commodity, f_6 .

$$\text{Production} = f_6(\text{CPA employment}) \quad (6)$$

Tour type distribution is defined by apply fuzzy clustering of 1.7 million trip samples from the NST/R, f_7 . Four tour types are defined: Collection/distribution, consolidation, trucking segment, and shuttle tours.

$$\text{Tour Type} = f_7(\text{NST/R sample trips}) \quad (7)$$

A gravity model is used to distribute the productions to their tour types, which effectively transforms the economic-based CPA into transport-based tours. The final schema allows decision-makers to perform microscopic simulation of commodity tours.

As an initial proposed framework, it makes sufficient use of available data in Europe to model commodity tours. Whereas SMILE requires abundant logistics data to model the distribution of commodities, Liedtke and Schepperle's framework makes use of the NST/R trip samples to estimate the tour choice of different commodities. While the NST/R trip sample data does not have an equivalent version in the U.S., similar surveys can be conducted by an agency.

Joint shipment size and transport chain choice model

Disaggregate joint shipment size and mode choice models have been around since McFadden et al. (1985), as pointed out in Regan and Garrido (2001). The model developed by de Jong and Ben-Akiva (2007) expands on the mode choice aspect into a set of commodity distribution chain choices. Their motivation is to account for logistics elements because of recent logistics changes such as just-in-time delivery. Compared to SMILE, this model is specified at the level of the decision-maker, from one sender to one receiver.

The model takes aggregate production–consumption matrices for zonal totals as inputs. Several choices are modeled, including frequency/shipment size, number of legs or stops, location and use of consolidation and distribution centers, and mode and vessel type for each leg. Inventory decisions are generally assumed to be made by the receiver.

The aggregate production flows are first disaggregated to annual firm-to-firm flows using the number of employees per firm by zone, f_8 .

$$\text{Firm Flows} = f_8(\text{Production} - \text{wholesale} - \text{consumption matrices}) \quad (8)$$

Once the annual flows are provided, the inventory portion of the model determines the shipment size at the destination of a chain such that total logistics costs are minimized using the economic order quantity (EOQ) model, f_9 . Although purchase costs of goods are not directly included because the senders and receivers are predetermined, the variable is indirectly incorporated in the capital costs. Note that the transport costs are not actually determined until the transport chains are determined. Hence, the transport costs are either assumed to be constant with respect to the shipment size or approximated iteratively.

$$\text{Destination Shipment Size} = f_9(\text{Firm Flows}, \text{Logistics Costs}) \quad (9)$$

After shipment size is determined at the destination, the same EOQ cost function is used for minimizing disutility in a random utility discrete choice model, f_{10} .

$$\begin{aligned} & \text{Transport Chain(No. of Legs, Mode/Vehicle Type/Loading Unit per Leg,} \\ & \text{Transshipment Location per Leg)} \\ & = f_{10}(\text{Destination Shipment Size, Available Modes, Firm Flows}) \end{aligned} \quad (10)$$

Empty truck movements are included in this model by adding an additional commodity called “empties” and having the flows mirror the vehicle flows, f_{11} . This method uses exogenously determined return loads. The method is based on the empty truck model by Holguín-Veras and Thorson (2003).

$$\text{Empty Truck Flows} = f_{11}(\text{Transport Chain (Truck Flows)}) \quad (11)$$

The model is more data costly than the other two models discussed in this subclass. Not only are logistics costs and initial I–O tables needed, but surveys would need to be conducted for individual sender–receiver pairs to obtain the estimates of the transport chain models. It is also possible to calibrate the model to aggregate data such as zonal mode shares by commodity type. Disaggregation to firm flows requires firm registration by zone information. In California most of this information should be available at the BEA zonal level (of which there are five coarse metropolitan-based zones) except for firm transport chain survey data.

Service delivery truck tours are not included in this model since it focuses on the movement of commodities. Like the other models in this subclass, integrated shipper-carrier operations are assumed.

Urban logistics models

One of the early supply chain models identified by Regan and Garrido (2001) is the GoodTrip urban logistics model. The following subclass of commodity-based models incorporates logistics behavior like the regional logistics models, but they tend to focus more on the urban setting with commodity-based truck tours.

Urban logistics models (and the class G models below) also fall into a category called city logistics, which is defined as “the process of totally optimizing urban logistics activities by considering the social, environmental, economic, financial, and energy impacts of urban freight movement” (Taniguchi et al. 2001). This category of models encompasses optimization methods more so than the behavioral methods in regional freight modeling.

Goodtrip model

Boerkamps et al. (2000) developed a four-step modeling framework for supply chain elements of urban freight movement and applied it to the city of Groningen in the Netherlands. Their conceptual framework considers the behavior of multiple actors including the sender, the transporter, and the receiver, as well as multiple distribution channels. The model can be used to analyze changes in consumption patterns, different supply chain organizations, different delivery requirements, different distribution patterns and mode choices, and impacts of environmental improvements.

The initiating process is the consumer demand for a commodity where the production is related to the land uses in a zone, f_{12}

$$\text{Volumes of Goods by Type by Zone} = f_{12}(\text{land use, consumer demand}) \quad (12)$$

Goods demand flows can be determined as a two step process of supplier choice (by the receiver) and spatial choice. After the goods flows have been estimated, they are combined using groupage probabilities, f_{13} , and assigned to vehicle tours by mode.

$$\text{Combined goods flow} = f_{13}(\text{Spatial Choice} * \text{Supplier Choice}(\text{Goods Volume}), \text{groupage probabilities}) \tag{13}$$

Depending on the origin’s activity type/land use, the transport mode, vehicle capacity, maximum load factor, and maximum number of stops per tour are determined, f_{14} . Examples of activity types are consumers, supermarkets, stores, offices, distribution centers of retailers, and producers. The mode choice refers specifically to the type of truck, such as traditional distribution trucks, urban DC trucks, or underground logistics trucks.

$$\begin{aligned} &(\text{Mode, Vehicle Capacity, Maximum Load, Maximum No. of Stops}) \\ &= f_{14}(\text{Origin activity type}) \end{aligned} \tag{14}$$

The destination’s activity type is assumed to determine the minimal delivery frequency, f_{15} .

$$\text{Delivery Frequency} = f_{15}(\text{Destination activity type}) \tag{15}$$

Groupage probabilities would need to be obtained from shipment surveys. The model was calibrated and validated for the food retail sector in Groningen using data from a traditional four-step traffic model and empirical data from a traditional method of distribution.

Urban freight micro-simulation

Wisetjindawat and Sano (2003) developed an initial urban freight commodity flow model that was later extended by Wisetjindawat et al. (2006) to incorporate the fractional split distribution method developed by Sivakumar and Bhat (2002). The model consists of two components: commodity generation and commodity distribution.

The production and consumption volumes by each firm is obtained from a regression model related to the firm size, number of employees, and floor area, f_{16} . This is different from f_{12} in GoodTrip, which obtains productions by zone.

$$\text{Commodity Productions/Consumptions by Firm} = f_{16}(\text{firm size, no. employees, floor area}) \tag{16}$$

The fraction of a commodity that is assigned to a customer j from shipper i is estimated from a spatial mixed logit model, f_{17} . It is equal to the product of the probability of a distribution channel chosen, the probability of a zone being chosen conditional on the distribution channel, and the probability of a shipper i chosen by customer j conditional on the distribution channel and zone.

$$\text{Commodity Fraction} = f_{17}(\text{Distribution Channel Choice, Zone Choice, Shipper Choice}) \tag{17}$$

The distribution choice is estimated from empirical data; the zone choice is estimated from a spatial mixed logit model, f_{18} , with a maximizing utility function based on zonal attractiveness. Shipper choice is obtained from an assumed logistic function of commodity productions by firm, f_{19} .

$$\text{Zone Choice} = f_{18}(\text{Number of firms in Zone, Number of Commodities in Zone}) \quad (18)$$

$$\text{Shipper Choice} = f_{19}(\text{Commodity Productions by Firm}) \quad (19)$$

The commodity flows from shipper i to receiver j can then be determined, f_{20} . These firm-to-firm commodity flows can be aggregated up to zonal levels for validation with commodity flow survey data.

$$\text{Commodity Flow} = f_{20}(\text{Commodity Fraction, Commodity Consumed by Firm}) \quad (20)$$

The model is calibrated with data from the Tokyo Metropolitan Goods Movement Survey, which is collected from 46,000 firms corresponding to approximately 3% of the study area. The records include information on firm characteristics, location, number of employees, commodity type, delivery frequency, origins and destinations of freight trips, truck sizes, etc. This is very detailed firm-based information that has no equivalence in California. However, it shows what can be achieved if such data is available.

Class G—vehicle touring models

The difference between the vehicle touring models in this class and the ones mentioned in class F is whether the unit of analysis is a commodity/shipment or a vehicle. Generally the commodity/shipment models focus on the agents' behaviors from the perspective of minimizing logistics costs. The vehicle-touring models here focus on capturing the movements of vehicles and decisions of carriers realistically for more accurate evaluation.

Truck tours can be divided into many types, as shown by Figliozzi (2007), and can be used to analyze impacts from congestion (Figliozzi et al. 2007) or technological changes (Figliozzi 2006). Like the urban logistics models in class F, the class G models can also be considered a part of the field of city logistics.

The vehicle touring models include the developments by Garrido and Mahmassani (2000) and the more activity-based vehicle touring models of Hunt and Stefan (2007) and Gliebe et al. (2007).

Space–time multinomial probit model

Garrido and Mahmassani's (2000) model forecasts the distribution of freight flows over space and time by linking the generation of the service demand to different time intervals and zones. As an econometric model, it explains the demand in terms of socioeconomic factors such as population, population density, average weekly wages, unemployment rate, and number of private vehicles. An autoregressive discrete choice model with a spatial lag operator is used, f_{21} .

$$\begin{aligned} \text{Demand Spatial – Temporal Choice} = f_{21}(\text{Population, Density, Wage,} \\ \text{Unemployment, No. of Private Vehicles}) \end{aligned} \quad (21)$$

In addition to the socio-economic data, a sample of records of load pickup and delivery from one or more carriers is necessary. While this model explains the demand in terms of socioeconomic factors, it does not consider the nature of truck tours nor the logistics choices for different distribution channels.

Truck tour-based microsimulation model

This model is perhaps most practically viable because they have actually been implemented for the city of Calgary in Canada (Hunt and Stefan 2007) and for the urban areas in the state of Ohio (Gliebe et al. 2007). The core of the model is the tour-based microsimulation, and the specific model described below refers to the one estimated for Calgary.

The model is capable of analyzing the following illustrative truck policies: increasing the cost per distance of operating vehicles (gas prices); increasing the travel time for all vehicles (congestion); changing truck route restrictions (accessibility); or cordon toll pricing for particular zones.

Before the number of tours can be generated, an accessibility measure for each O–D pair is needed. Zonal accessibility is based on a logit regression model, f_{22} .

$$\text{Zonal Accessibility} = f_{22}(\text{Travel Disutility}(\text{Travel Time}, \text{Travel Distance})) \quad (22)$$

The number of tours generated in each zone is obtained from an aggregate exponential regression model, f_{23} . The land uses include: Industrial, Wholesale, Retail, Transport, and Services.

$$\text{No. of Tours by Zone} = f_{23}(\text{Land Use}, \text{Employment}, \text{Zonal Employment Accessibility}) \quad (23)$$

To simulate the individual tours using Monte Carlo simulation, a number of discrete choice models are employed. A logit model is used to determine the time period that a tour is allocated to, assuming the carrier is trying to maximize utility by choosing time period, f_{24} .

$$\text{Tour Time Period} = f_{24}(\text{Zonal Employment Accessibility}, \text{Land Use by Time Period}) \quad (24)$$

After allocating the tours to time periods, each tour is jointly assigned a primary purpose and a vehicle type using a multinomial logit model, f_{25} . The primary purposes include Goods, Service, and Other, while the vehicle types include Light, Medium, and Heavy.

$$(\text{Purpose}, \text{Vehicle}) = f_{25}(\text{Land Use by Purpose}, \text{Land Use by Vehicle}) \quad (25)$$

The start time of each vehicle tour is simulated using Monte Carlo from a cumulative percentage distribution function obtained from observed samples for each time period, f_{26} . Note that these start times are not correlated with any other attribute so there is no behavioral aspect to this variable.

$$\text{Start Time} = f_{26}(\text{Observed Start Times by Time Period}) \quad (26)$$

During each leg of the tour, a logit model is used to determine the purpose of the following leg, f_{27} . There are three general purposes: Business (continuing their primary purpose), Other (changing purpose), and Return. The Return alternative sends the vehicle back to the depot for the day. The tour continues simulating new stops until the Return alternative is chosen.

$$\begin{aligned} \text{Next Stop Purpose} = f_{27}(\text{No. of Previous Stops}, \\ \text{Previous Business by Purpose and by Vehicle}, \text{Total Time}, \\ \text{Travel Time}, \text{Zonal Employment Accessibility}) \end{aligned} \quad (27)$$

For each new stop, a location needs to be estimated using a logit model to select the zone that maximizes the utility function, f_{28} . The Direction factor refers to the factor of angle change in direction, such as going in the opposite direction for a vehicle in the tour.

The Transport Attraction refers to the relative attractiveness of the zone to Transport land uses.

$$\text{Next Stop Location} = f_{28}(\text{Land Use, Zonal Avg Household Income, Travel Disutility, Travel Disutility of Return, Zonal Population Accessibility, Zonal Employment Accessibility, Direction, Population, Employment, Transport Attraction}) \quad (27)$$

In the Calgary model, the stop duration is a static Monte Carlo simulation from observed stop times; however, in the Ohio model the stop duration is handled as an additional Next Stop Purpose alternative: Stay.

The data required for the Calgary model is a set of extensive interviews of commercial vehicle movements for over 3100 transport businesses in the study area, similar to household trip diary interviews. Records include origin, destination, purpose, fleet, and commodity type information. The result of the survey is choice behavior information for 64,000 commercial vehicle trips. In the Ohio statewide model, an establishment survey was conducted for 562 establishments with similar types of information as the survey for Calgary.

Referring back to Table 2, Table 3 shows that the class F and class G should be able to address needs (1), (2), and (11) with their primary outputs because they can be used to provide value to different network improvements. Modal diversion (3) and rail planning (8) are not considered by truck touring models but they are an integral part of logistics models that include carrier behavior and handling costs. Since the final output of the two types of models includes truck volumes, they can be used to address Pavement, bridge, and safety

Table 3 State needs versus new model classes

Policy and analytical needs		Type of tool	
		Logistics models	Truck touring models
1	State transportation planning	P	P
2	Project prioritization, STIP development	P	P
3	Modal diversion analysis	P	–
4	Pavement, bridge, and safety management	P	P
5	Policy studies	P	P
6	Needs analysis	P	P
7	Commodity flow analysis	P	S
8	Rail planning	P	–
9	Trade corridor and border planning	S	–
10	Operations, safety, security, truck size and weight issues, etc.	–	P
11	Project development or design needs; e.g. forecasts and loadings	P	P
12	Terminal access planning	P	S
13	Truck flow analysis and forecasting	P	P
14	Performance measurement/program evaluation	–	–
15	Bottleneck analysis	S	P

P primary, S secondary

management (4). Class F can be used to address policy studies (5) based on economics and trade such as commodity-based taxes and subsidies, whereas class G can be used for vehicle-based policies only, such as pricing strategies and restricted lanes. Like the other models, both F and G can be used for needs analysis (6) by identifying gaps in goods movement that need improvement. While F can directly model commodity flow analysis (7), G supports it only to some degree with broader commodity groups that are linked to the commercial vehicles. Both F and G models tend to use exogenous aggregate economic data, so they would not directly address trade corridor and border planning (9). However, depending on the scope and data availability of the logistics models there could be some sensitivity to external commodity flows. Truck operations (10) can be addressed with class G since the weight and type of truck at each leg of the tour would be known. Terminal access (12) is addressed by the logistics models with handling costs, and to some extent by truck touring models that include truck terminals as stops. Truck flow (13) and bottleneck analysis (15) can generally be addressed by both classes, although class G can provide much better resolution information. Like the other models, F and G do not directly address performance measurement (14), although new measures can be defined for these models compared to the A–E.

Other innovations

Several other pioneering models have been proposed, though some have not yet been implemented due to data limitations and lack of sufficient interest and resource availability from appropriate decision makers. Nagurney et al. (2002) proposed a multilevel network formulation for the conceptualization of supply-chain problems that examines a single product at a time. Their novel approach combines logistical, information and financial networks.

A number of interesting new insights and improvements have come about for models in practice as well as in areas related to freight demand modeling.

One related area is freight network simulation. The manner with which commodity flows are assigned to an infrastructure network can affect how the demand relates to freight traffic. Southworth and Peterson (2000) created a nation-wide freight simulation network using the commodity flow survey data on a geographical information system (GIS) platform. Xu et al. (2003) developed a freight traffic simulator as part of their TTMNet module that receives pseudo-real time information from a dynamic supply chain network equilibrium model. Mahmassani et al. (2007) and Zhang et al. (2008) developed an intermodal, dynamic freight network simulator that accounts for load scheduling and vehicle routing. Their system has been applied to an intermodal corridor in Europe to illustrate the analysis of several policies such as reducing technological/communicative barriers across national borders and improving infrastructure by increasing maximum rail speeds.

Sivakumar and Bhat (2002) developed an alternative approach to distributing commodity flows. Instead of assuming a gravity model derived primarily from trip-based passenger demand modeling, their fractional split-distribution model is structured to resemble the choice patterns of tours in a logistical distribution channel. Fractions of the commodity from an origin are estimated to be consumed at each destination using a multinomial logit form. The authors show that their model performs better than a standard gravity model, and has been used by Wisetjindawat et al. (2006). Nuzzolo et al. (2008) uses a partial share method of distribution similar to the fractional split method to estimate freight flows for an international road network.

Wang and Holguín-Veras (2009) developed an alternative freight distribution model using entropy maximization to assign truck volumes to tours. Unlike the truck touring models in Class G or the urban logistics models in Class F, this model uses aggregate data only. The benefit to this type of model is the reduced data required compared to the data-intensive F and G models, while still providing realistic touring elements in a freight demand model.

Ham et al. (2005) developed a combined model of interregional, multimodal commodity shipments, incorporating regional input–output relationships, and the associated transportation network flows. Their model, which presents an alternative to the traditional four-step travel forecasting procedure, falls into Class E, economic activity models.

Holguín-Veras and Thorson (2003) created an explicit model for empty truck movements that could be integrated with commodity flow models. The motivation for their study is the multidimensionality of freight demand, where vehicle-based models cannot capture economic factors but commodity-based models are unable to capture realistic truck movements such as empty trips. The total number of truck trips is split into the loaded trips and the empty return trips. This model was incorporated into de Jong and Ben-Akiva's (2007) model.

Giuliano et al. (2007) addressed the issue of data insufficiency by using secondary data sources to estimate commodity freight flows at a resolution from which a singular data source is not available. The resolution of existing public data such as the Commodity Flow Survey only goes down to the level of the aggregate metropolitan area. Using secondary data sources, Giuliano et al. were able to estimate the inter-county flows for the five counties in the Los Angeles metropolitan area. Reconciliation between multiple conflicting data sources for different years is conducted by using one source, IMPLAN, as the control.

Holguín-Veras and Patil (2008) developed an OD estimation method for freight that includes both loaded and empty truck trips. OD estimation is a method of estimating origins and destinations based on observed link counts. While the literature in OD estimation for passenger models is abundant, there is not as much found in freight literature. With this method, it would be possible to use truck link volume data in California such as their Freeway Performance Measurement Systems (PeMS) data to estimate intercity truck OD flows.

Lessons learned

Implementation issues

To date regional logistics models have not been applied by any U.S. state agencies because the private firm supply chain costs and operating behavior needed by these models are unavailable. However, this type of model has been developed and applied in some European countries such as the Netherlands and Germany.

Urban logistics models have seen implementation or demonstrations in other countries as well, including Tokyo, Japan and Groningen, Netherlands.

Vehicle touring models have been applied in several U.S. states and in Canada. Development efforts of this type are also underway in some European Countries and Australia.

Case study: statewide model implementation

A regional commodity-based model is necessary for statewide implementation. Using California as a representative case study here, we first point out that because the state includes several major metropolitan areas and some of the largest ports in the world, it is also important to incorporate vehicle-based truck touring models to accurately reflect all the movements. While a truck touring model such as the one implemented by Ohio may be considered, in such a large state, the regional metropolitan planning agencies tend to have a much stronger role in overseeing truck policies than does any single state agency. Because of that, only regional commodity-flow based models are appropriate in such a setting.

If alternatives were to be considered for a statewide freight demand model in California, the following could represent a well-rounded spectrum of state of the art developments depending on the investments the agency is willing to make for data.

“Traditional” aggregate commodity flow model

With the incorporation of recent developments, a four-step commodity flow or economic input–output model as described under the class D or E models may not be as “traditional” as one might think. Empty vehicles can be incorporated, and OD flows can even be estimated using only truck link volumes based on the results of Holguín-Veras and Patil (2008). The gravity model in the class D can be replaced with the fractional split model developed by Sivakumar and Bhat (2002) or the tour-based gravity model by Wang and Holguín-Veras (2009). Secondary source reconciliation methods shown by Giuliano et al. (2007) can be used to estimate flows at a higher resolution of detail such as county to county.

The resulting model would have more realistic truck movements than the Class D or E models, more realistic commodity flows, and also allow the use of an additional data source (truck link volumes) for estimating the flow distribution.

The additional data required with this model would include socioeconomic data for the distribution step. However, much of that data is already required at the commodity generation step. Of greater concern is the need for a higher resolution model; the public data sources currently available for California essentially drill down to the BEA zonal levels for major metropolitan areas, not for the county-to-county level. Depending on the degree of disaggregation desired, the use of secondary sources requires an abundant number of local socioeconomic data. Empty truck estimations would require survey data on truck loads, which are possibly available from weigh-in-motion (WIM) data at a regional level.

This hypothetical model is still an aggregate model, so it lacks much of the agent-based behavioral aspects found in the disaggregate models. For example, it is capable of analyzing a fuel cost policy in terms of shifts in empty truck trips, but it would still not be able to handle changes in port operating hours or warehouse real estate costs. In addition, the empty truck model portion would still not explain much of the service delivery truck movements found in urban areas, for which truck policies would have significant impact such as truck tolling or restrictions. For a statewide model, it can be argued that such visibility is best left to the metropolitan planning agencies.

Table 4 Alternative state-of-the-art models and data requirements

If data available...	California example(s)	Aggregate class D or E	Regional logistics	Urban logistics
Commodity flow	Commodity flow survey (CFS), freight analysis framework version 2 (FAF ²)	X	X	X
Employment	Bureau of economic analysis (BEA)	X	X	X
Truck load Data	Weigh-in-motion (WIM)	X	X	X
Make/use, I–O tables	BEA make/use by industry		X	X
Intermodal facilities	Caltrans intermodal traffic management system (ITMS)		X	X
Logistics costs	Distribution channel choice, transport costs		X	X
Firm shipment sizes and distributions	Supplier choice, no. employees		X	X
Truck activity diaries	Tour choice, times of activities, durations			X
Objectives achieved		Same as in NCHRP 606, plus more realistic truck counts	Same as Aggregate, plus distribution/ mode choice/shipment size behavior explained	Same as Regional Logistics, plus urban truck behavior explained

Disaggregate regional logistics model

At the middle of the spectrum, a Class F model such as de Jong and Ben-Akiva's would provide significantly more detailed analysis of more modern policy questions such as intermodal facility operations, investing in new modes or intermodal facilities, locations of distribution centers, etc. This class of models provides behavioral insights to: receivers' choice of suppliers and distributors, shippers' choice of mode and shipment size, as well as their transport and operating costs.

Data requirements would be costlier because a survey of shippers and/or receiver firms would be needed. Shipper survey data by itself would provide insight on the commodity distributions in terms of logistical movements, but firm survey data is necessary to understand the choice of suppliers for behavioral commodity productions and attractions (such as the urban freight logistics models in class F). In California, this type of information could be extremely hard to get because of privacy issues in the U.S.

By itself, a class F model does not do a good job of describing truck movements in urban areas, such as delivery trucks and empty vehicles. Like the model described in the previous section, urban truck movements would not be captured, but for a statewide model it might not be necessary.

Regional logistics plus urban truck touring model

Alternatively, California can include both a class F and a class G model to analyze logistics policies as well as truck policies. The commodity distributions that end up in metropolitan areas can be used as the productions for intra-urban truck touring simulation similar to the model by Hunt and Stefan (2007).

The data requirements of such a model would be the most intensive. Not only would the agency need to obtain shipper and receiver surveys, but they would also need to get carrier/trucking company surveys that are similar to household activity diaries for developing an activity-based model. For such an alternative in California, the daily activity surveys may be conducted for each of the four primary metropolitan areas listed as separate zones in the BEA.

With the costly data requirements comes the most comprehensive model in the spectrum. This hybrid model would be able to handle the widest range of policy questions and support investment decisions ranging from public and private infrastructure to vehicle technologies.

The road ahead

Table 4 summarizes the data needed for the alternatives discussed in the preceding section along with the objectives that could be achieved using each model.

While significant progress has been made in the last 8 years since Regan and Garrido (2001), improvements and extensions are still needed. We find that further developments including dynamic shipper-carrier interaction are needed and that because of limited data availability and tight public sector resources for additional data gathering, that advances in applying data mining techniques to available or easily developed data sources would be a huge benefit to researchers and planners alike. In addition, the development of hybrid models, for example progress in the integration of regional logistics models with urban

truck touring models, will introduce new problems such as reconciling the outputs of multiple models for consistency.

Acknowledgements The authors would like to thank the three anonymous reviewers for their very helpful comments on an earlier draft of the paper. Opinions expressed are those of the authors alone.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Boerkamps, J.H.K., van Binsbergen, A.J., Bovy, P.H.L.: Modeling behavioral aspects of urban freight movement in supply chains. *Transp. Res. Rec.* **1725**, 17–25 (2000)
- de Jong, G., Ben-Akiva, M.: A micro-simulation model of shipment size and transport chain choice. *Transp. Res. B* **41**(9), 950–965 (2007)
- de Jong, G., Gunn, H., Walker, W.: National and international freight models: an overview and ideas for further development. *Transp. Rev.* **24**(1), 103–124 (2004)
- Federal Highway Administration: Freight Analysis Framework Overview. Publication FHWA, FHWA-OP-03-006 (2002)
- Federal Highway Administration: Quick Response Freight Manual I and II. Freight Management and Operations. U.S. Department of Transportation, FHWA-HOP-08-010 (2007)
- Figliozzi, M.A.: Modeling the impact of technological changes on urban commercial trips by commercial activity routing type. *Transp. Res. Rec.* **1964**, 118–126 (2006)
- Figliozzi, M.A.: Analysis of the efficiency of urban commercial vehicle tours: data collection, methodology, and policy implications. *Transp. Res. B* **41**(9), 1014–1032 (2007)
- Figliozzi, M.A., Kingdon, L., Wilkitzki, A.: Analysis of freight tours in a congested urban area using disaggregated data: characteristics and data collection challenges. In: Proceedings from 2nd Annual National Urban Freight Conference, Long Beach, CA (2007)
- Fischer, M.J., Outwater, M.L., Cheng, L.L., Ahanotu, D.N., Calix, R.: Innovative framework for modeling freight transportation in Los Angeles County, California. *Transp. Res. Rec.* **1906**, 105–112 (2005)
- Friesz, T.L., Holguín-Veras, J.: Dynamic game-theoretic models of urban freight: formulation and solution approach. In: *Methods and Models in Transport and Telecommunications, Advances in Spatial Science*. Springer, Berlin (2005)
- Garrido, R.A., Mahmassani, H.S.: Forecasting freight transportation demand with the space-time multinomial probit model. *Transp. Res. B* **34**(5), 403–418 (2000)
- Giuliano, G., Gordon, P., Pan, Q., Park, J.Y., Wang, L.L.: Estimating freight flows for metropolitan area highway networks using secondary data sources. *Netw. Spatial Econ.* (2007) (Online First)
- Gliebe, J., Cohen, O., Hunt, J.D.: Dynamic choice model of urban commercial activity patterns of vehicles and people. *Transp. Res. Rec.* **2003**, 17–26 (2007)
- Ham, H., Kim, T.J., Boyce, D.: Implementation and estimation of a combined model of interregional, multi-modal commodity shipments and transportation network flows. *Transp. Res. B* **39**(1), 65–79 (2005)
- Hensher, D.A., Puckett, S.M.: Refocusing the modeling of freight distribution: development of an economic-based framework to evaluate supply chain behavior in response to congestion charging. *Transportation* **32**(6), 573–602 (2005)
- Hensher, D., Figliozzi, M.A.: Behavioural insights into the modeling of freight transportation and distribution systems. *Transp. Res. B* **41**(9), 921–923 (2007)
- Hesse, M., Rodrigue, J.P.: The transport geography of logistics and freight distribution. *J. Transp. Geogr.* **12**(3), 171–184 (2004)
- Holguín-Veras, J., Thorson, E.: Modeling commercial vehicle empty trips with a first order trip chain model. *Transp. Res. B* **37**(2), 129–148 (2003)
- Holguín-Veras, J., Patil, G.R.: A multicommodity integrated freight origin-destination synthesis model. *Netw. Spatial Econ.* **8**(2–3), 309–326 (2008)
- Hunt, J.D., Donnelly, R., Abraham, J.E., Batten, C., Freedman, J., Hicks, J., Costinett, P.J., Upton, W.J.: Design of a statewide land use transport interaction model for Oregon. In: Proceedings from World Conference on Transportation Research, Seoul, South Korea, p. 41 (2001)
- Hunt, J.D., Stefan, K.J.: Tour-based microsimulation of urban commercial movements. *Transp. Res. B* **41**(9), 981–1013 (2007)

- Iowa Department of Transportation and Iowa State University Center for Transportation Research and Education, accessed 2008, http://ctre.iastate.edu/Research/statemod/dev_guid.pdf/
- Liedtke, G., Schepperle, H.: Segmentation of the transportation market with regard to activity-based freight transport modelling. *Int. J. Logist. Res. Appl.* **7**(3), 199–218 (2004)
- Mahmassani, H.S., Zhang, K., Dong, J., Lu, C.C., Arcot, V.C., Miller-Hooks, E.: Dynamic network simulation-assignment platform for multiproduct intermodal freight transportation analysis. *Transp. Res. Rec.* **2032**, 9–16 (2007)
- McFadden, D., Winston, C., Boersh-Supan, A.: Joint estimation of freight transportation decisions under nonrandom sampling. In: Daughety, A.F. (ed.) *Analytical Studies in Transport Economics*, pp. 137–157. Cambridge University Press, Cambridge (1985)
- Nagurney, A., Ke, K., Cruz, J., Hancock, K., Southworth, F.: Dynamics of supply chains: a multilevel (logistical/informational/financial) network perspective. *Environ. Plan. B* **29**, 795–818 (2002)
- National Cooperative Highway Research Program Report 606: Forecasting Statewide Freight Toolkits. Transportation Research Board, Washington D.C. (2008)
- Nuzzolo, A., Crisalli, U., Comi, A.: A demand model for international freight transport by road. *Eur. Transp. Res. Rev.* (2008) (Online First)
- Ohio Department of Transportation, Office of Urban and Corridor Planning, http://www.dot.state.oh.us/Planning/Freight/freight_impacts_report.htm/
- Pennsylvania Department of Transportation: Development of the Pennsylvania Statewide Commodity-Based Freight Model. In: 11th TRB National Transportation Planning Applications Conference (2007)
- Prousaloglou, K., Popuri, D., Tempesta, D., Kasturirangan, K.: Wisconsin passenger and freight statewide model: case study in statewide model validation. *Transp. Res. Rec.* **2003**, 120–129 (2007)
- RAMON Classification Detail List, Statistical Classification of Products by Activity in the European Economic Community, 2008 version, http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC
- Regan, A.C., Garrido, R.A.: Modelling freight demand and shipper behaviour: state of the art, future directions. In: Hensher, D. (ed.) *Travel Behaviour Research*. Pergamon-Elsevier Science, Amsterdam (2001)
- Regional Multipliers: A User Handbook for the Regional Input-Output Modeling System (RIMS II), 3rd edn. Bureau of Economic Analysis, U.S. Department of Commerce (1997)
- Research and Innovative Technology Administration: Transportation Vision 2030. U.S. Department of Transportation (2008)
- Sivakumar, A., Bhat, C.: Fractional split-distribution model for statewide commodity-flow analysis. *Transp. Res. Rec.* **1790**, 80–88 (2002)
- Southworth, F., Peterson, B.E.: Intermodal and international freight network modeling. *Transp. Res. C* **8**, 147–166 (2000)
- Taniguchi, E., Thompson, R.G., Yamada, T., van Duin, R.: *City Logistics: Network Modelling and Intelligent Transport Systems*. Pergamon, UK (2001)
- Tavasszy, L.A., Smeenk, B., Ruijgrok, C.J.: A DSS for modeling logistic chains in freight transport policy analysis. *Intell. Transp. Oper. Res.* **5**(6), 447–459 (1998)
- United States Government Accountability Office: Freight Transportation National Policy and Strategies Can Help Improve Freight Mobility. Publication GAO, GAO-08-287 (2007)
- Wang, Q., Holguín-Veras, J.: Tour-based entropy maximization formulations of urban freight demand. In: *Proceedings of 88th Transportation Research Board* (2009)
- Wisetjindawat, W., Sano, K.: A behavioral modeling in micro-simulation for urban freight transportation. *J. East. Asia Soc. Transp. Stud.* **5**, 2193–2208 (2003)
- Wisetjindawat, W., Sano, K., Matsumoto, S.: Commodity distribution model incorporating spatial interactions for urban freight movement. *Transp. Res. Rec.* **1966**, 41–50 (2006)
- Xu, J., Hancock, K.L., Southworth, F.: Simulation of regional freight movement with trade and transportation multinetworks. *Transp. Res. Rec.* **1854**, 152–161 (2003)
- Zhang, K., Nair, R., Mahmassani, H.S., Miller-Hooks, E.D., Arcot, V.C., Kuo, A., Dong, J., Lu, C.C.: Application and validation of dynamic freight simulation-assignment model to large-scale intermodal rail network: Pan-European case. *Transp. Res. Rec.* **2006**, 9–20 (2008)

Author Biographies

Joseph Y. J. Chow is a post doctoral researcher at the Institute of Transportation Studies at the University of California, Irvine and lecturer at the University of Southern California. He obtained his Master's and Bachelor's from Cornell and his PhD from UCI. His research interests include network modeling, transport economics, and urban operations research.

Choon Heon Yang is a senior researcher with the Korea Institute of Construction Technology. He completed his PhD at the University of California, Irvine, where his research focused on the development of multicriteria decision making methods for prioritizing alternative truck management strategies. His research interests include all aspects of freight transportation.

Amelia C. Regan is a professor of computer science and civil (transportation systems) engineering at the University of California, Irvine. Her research interests include optimization of freight transportation and telecommunication and computing networks, optimal contracting, network design and network design and security issues for ad hoc vehicular networks (VANets).