A Hybrid Approach to Develop Freight Model from Commercial Vehicle Travel Surveys and Commodity Flow Data

Arun Kuppam*, Dan Beagan, Rachel Copperman, Ramesh Thammiraju – Cambridge Systematics, Inc.

Vladimir Livshits, Lavanya Vallabhaneni, Sreevatsa Nippani – Maricopa Association of Governments

*** Corresponding Author**

Arun Kuppam
Cambridge Systematics, Inc.
9015 Mountain Ridge Dr, Suite 210
Austin, TX 78759
Tel: 512 691 8503
Fax: 512 691 3289
E-mail: akuppam@camsys.com

Submitted for Presentation at the Innovations in Travel Modeling 2010 Conference, Tempe AZ, May 9-12, 2010

Introduction

The Maricopa Association of Governments (MAG), through a contract with Cambridge Systematics, Inc. (CS), has recently updated the Internal and External Truck Travel Models to a base year of 2007, to estimate Medium (single units) and Heavy (combination units) truck trips at the zonal level traveling within, from/to and through the three-county MAG region.

There are several unique and innovative features employed in this truck model update. These are:

- Collected O-D travel information from trucks that travel within the MAG region using different surveying techniques for different sectors;

- Innovative way of distributing trucks based on land use-to-land use interchanges rather than the traditional O-D based gravity model;

- Developed linear and log-linear relationships between external freight flows and underlying socio-economic activity at the zonal level;

- Integrated land use-based internal truck model and commodity flow-based external truck model into a "hybrid" truck model.

Approach

The "hybrid" approach involves blending commodity flow modeling techniques with freight truck modeling techniques. Commodity flow databases tend to be relatively accurate for inter-county flows, but undercount intra-county flows because commodity flow databases rely in part on economic input-output data that ultimately are based on financial transactions between producers and consumers of goods. However, in an urban area many truck moves are not easily traced to such transactions. Moves from warehouses and distribution centers, repositioning of fleets, drayage moves, parcel delivery, and the like are generally short-distance trips in which there may not be an economic exchange of the goods from one party to another. To compensate for the undercounting of the shorter distance trips, local truck trips are generated based on local employment and economic factors using trip generation rates. These trips are generated and distributed at the zone level, and are calibrated so that the truck traffic volumes that are generated from the combined commodity flow and locally generated truck trips match those from available truck counts. Several terms are used to refer to these two trip types, including commodity flow trips versus locally generated trips; external versus internal truck trips; and long-haul versus short-haul truck trips.

Hybrid models, which take advantage of the benefits of the commodity flow and local truck models, including freight and other non freight truck purposes, have proven to be a very effective modeling framework for MAG's regional freight planning and air conformity analyses. Long-haul truck trips are modeled using the commodity flow database, which can be adjusted over time based on economic factors. Short-haul truck trips are estimated as a function of local employment characteristics and population.

Data Collection and Analysis

Collecting truck travel data internal to a region is integral to updating regional truck travel models. However, internal truck travel surveys are too few and far between, and little is known as to what works and what does not when designing surveys and collecting data. That is, there is a lack of significant research into what increases the effectiveness of truck data collection. This paper provides an innovative approach for collecting data using a combination of methods and sampling techniques as described below. These are used primarily to estimate and model internal truck travel behavior, where both the origin and destination is within the MAG region.

- Truck trip diaries were collected for sectors that generate multi-stop tours that are short haul in nature. These surveys are designed to collect truck travel information that include origin and destination information, stop locations and land use types at stops, trip lengths and number of trips by truck type and sector, and time-of-day distributions of truck trips. The sectors that were focused on in these surveys are agriculture, mining, construction, retail trade, pickup and delivery, mail/parcel, and for-hire categories. Trip diaries provided origin/destination, trip lengths, and information for trip generation estimates from truck drivers on the day they completed the diaries. These diaries were mostly mailed back but some of them were picked up at their respective facilities.

- Establishment surveys were used for sectors that generate truck traffic that are long haul in nature. These surveys were conducted by phone and were designed to collect information on the number of inbound and outbound truck trips at each facility or establishment, and the distribution of truck trips by trip distance and time of day. The sectors that were focused on in these surveys were primarily manufacturing facilities, wholesale trade, and warehouse/distribution centers. This form of surveying was done using Computer-Assisted Telephone Interviewing (CATI) and gathered information on truck's travel behavior on a typical day. These surveys provided production and attraction characteristics of different truck types by industry category, and trip length distributions by truck type.

As mentioned earlier, the above mentioned surveying methods are not suitable for obtaining external truck travel behavior which has at least one origin or destination or both outside the MAG region. TRANSEARCH commodity freight flow database was used to estimate the external truck travel for the region. The data base includes the annual volume of freight modes by all modes between the origins and destinations in the database. The units of geography in this database are zip codes within Arizona and states outside Arizona, and only freight moved by trucks was used from this database. Also included in this database is a set of fixed routing tables (paths) which can be used to assign this traffic. If the flow database is considered to be a trip table, that is it has volume flows between origins and destinations for specific purposes (commodities), and the routing tables can be considered to be an assignment, then the problem of converting the entire TRANSEARCH database into an external model trip table can be treated as if it were a classic subarea extraction problem. Using queries within Microsoft Access, the external stations in the MAG model were treated as boundary nodes on the segments (links) on the TRANSEARCH highway network and the trip table was windowed to the area inside and outside of that boundary. Effectively all zones outside of the boundary were transformed into movements at the external stations, while zones within the boundary were retained at the level of zip codes.

After this processing, TRANSEARCH was transformed into a trip table with zones which are zip codes for internal zones and highway stations for external trip ends, and with annual flows by truck by commodity. This trip table was treated as a survey database which was then used to estimate linear and log-linear regression models by commodity against zip code level employment to support production and attraction equations. The external portion of the trip table was used as a seed trip table constrained by observed truck volumes at external stations in a Iterative Proportional Fitting process (Fratar) to develop an external truck trip table. That trip table was then used to identify the portion of truck traffic at external stations which were considered as productions or attractions associated with the internal movements of freight. Additionally, the windowed TRANSEARCH trip table was considered to be the calibration table from which friction factors by commodity were developed for use in the external truck trip distribution model.

Internal Truck Model Development

Development of land use-based trip rates

The trip diaries and establishment surveys that collected information on the type of business at each stop, when expanded, resulted in data on the total number of trip ends at businesses in each industry category. Dividing by the number of employees in each industry category provided trip generation rates by truck type that was then applied to socioeconomic data to estimate trip ends by TAZ. With the available survey data, one set of trip rates for the productions and one set of rates for consumption are estimated. These rates are developed for each industry sector or land use and truck type. These industry sectors or land uses are analogous to “trip purposes” in the passenger model. The intercept is always forced to zero, because there should be no truck activity in or out of a zone without any related economic activity.

The land use categories considered for the truck trip productions and attractions are: employer (start and end point of any truck trip), retail, construction, farming, mining, households, governments, warehousing, transportation, office, manufacturing, and other. These estimated trip rates are derived after several iterations of the trip generation model. The results of the trip generation model are productions and attractions by land use and truck types, which are compared against the expanded survey database. The comparison of these results helped determine the final trip rates that need to be used for the truck-trip generation model.

Development of land use-to-land use distribution models

In typical trip distribution models, aggregating truck trip ends by purpose and then distributing those aggregated productions and attractions would link types of land use categories for which no activity was identified in the survey, and for which none is expected. Therefore, an innovative process was developed here that is based on the recognition that trip distribution is a connection between a land use category serving as a production and a land use category serving as an attraction. For example, in passenger modeling, Home-Based Work (HBW) passenger trips are those that occur between the Home land use production and the Work land use attraction. The HBW productions are the percentage of total home productions that will be distributed to work attractions, and the HBW attractions are the percentage of total work attractions that will be distributed to home productions. Using the same principle, separate gravity models for different land use interchanges were developed because the movement of truck trips from one land use to another are very distinct when compared against different land use exchanges. As there are 12 land uses identified from the surveys, the number of land use exchanges are 12 times 12 or 144 gravity models. This effectively retained the land use information throughout the trip distribution step. The friction factors between land uses, which are a function of average trip lengths, were updated using information from truck trip diary survey data collected.

The concept of retaining the land use distinctions for truck trips is at the heart of this innovative internal truck distribution concept. Prior to each of the 144 trip distributions, it would be necessary to determine what percentage of the total trips at the production land use are to each of the attraction land uses; and what percentage of the total trips at the attraction land use are from each of the production land uses. As applied to the productions for each zone for the production land use and the attractions for each zone for the attraction land use, this defines the P_i times A_j portion of the standard gravity model equation. The remaining term in the standard gravity model equation is the friction factor of distance between zones. Using separate gravity models for each land use to land use interchange, the average distance for each of the 144 land use interchanges was determined from the truck trip diary data.

This association of truck trips produced by a land use with trucks trips attracted by land use can also be considered to address some of desirable aspects of a trip chaining model. While the actual connection of truck trips into tours was not done as part of this truck model update, the distribution of trips between land uses does ensure that connections which would not exist in a trip chain will also not exist in this trip distribution. For example, if trip chaining would suggest that a retail to retail trip should be followed by another retail to retail trip and never by a retail to farm trip, the method described will ensure that the correct number of trips by purpose will result, just not in the right order. However the order would be irrelevant for purposes of trip table assignment.

External Truck Model Development

The external truck model is designed to estimate the daily generation of internal-external and external-internal trucks produced and attracted by each commodity type at the zonal level, with the exception of special generators, based on regression equations. The truck trips through the external stations are identified directly from the TRANSEARCH database.

The trip generation equations were developed through the estimation of linear and log-linear regressions using TRANSEARCH as the estimation database, and population and NAICS employment as the explanatory variables. The regression equations were developed at the zip code level for which the commodity, population, and NAICS employment data were available.

These relationships established at the zip code level were then applied to the zonal level data to produce productions and attractions of external truck trips for the MAG region.

The initial selection of appropriate employment and population variables to generate commodity volumes was guided by the U.S. Department of Commerce's Bureau of Economic Analysis Input-Output Tables. Those tables indicate the commodities made or used by various industries. The industries were matched up to the corresponding NAICS code. The explanatory variables tested within the regression models included population and employment by NAICS at 2- and 3-digit levels.

Special generators, which do not correlate well with the underlying employment and population, were identified while developing the regression models. These special generators were then removed from the estimation database so that the regression models fit well to the data. The special generator values for each commodity were determined by calculating the difference between the totals from the TRANSEARCH data and the estimated annual truck trips from the external truck model.

The production and attraction outputs from external truck trip generation were used as inputs into trip distribution. Trips were distributed using a doubly constrained gravity model with exponential decay. The average of internal-external and external-internal coefficients from original model were used for the friction factors in the impedance matrix. Trip tables obtained by different commodity type and time of day were added together and converted to daily truck trips.

Calibration/Validation of Truck Model

MAG collected truck classification counts on arterials through out the region but not on freeways. Therefore, screenlines for trucks were not available for validation of the truck model. Though truck counts on freeways and expressways were available from Arizona DOT, they were not classified in a manner that MAG model stratifies truck types. While the development of truck count screenlines to support validation of the truck model is preferred, given the limitations of the data, it was decided to validate only to the sum of medium and heavy trucks on arterials only. Also, internal and external trucks were included together and were validated to counts grouped by City in which the counts were located that served as a substitute to screenlines. It was however recognized that this is not ideal, since an error repeated on one route through a City would be added for all of the counts along that route. Therefore, the development of validation truck count screenlines was considered to be a next step in truck model validation once the data becomes available. The final validation results of the modeled daily truck volumes compared against observed counts indicate an overall observed error of seven percent.