INCORPORATING TRUCK FLOWS INTO THE STATE-WIDE PLANNING TRAFFIC MODEL

December 2004

JHR 04-299 Project 02-1

Lisa Aultman-Hall
Feng Guo
Christopher O'Brien
Patrycja Padlo
Brian Hogge

This research was sponsored by the Joint Highway Research Advisory Council (JHRAC) of the University of Connecticut and the Connecticut Department of Transportation and was carried out at the Connecticut Transportation Institute of the University of Connecticut.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
JHR 04-299	N/A	N/A

4. Title and Subtitle Incorporating Truck Flows into the State-wide Planning Traffic Model	5. Report Date December 2004 6. Performing Organization Code N/A
7. Author(s) Lisa Aultman-Hall, Feng Guo, Christopher O'Brien, Patrycja Padlo and Brian Hogge	8. Performing Organization Report No. JHR 04-299
9. Performing Organization Name and Address University of Connecticut Connecticut Transportation Institute Storrs, CT 06269-5202	10. Work Unit No. (TRAIS) N/A 11. Contract or Grant No. N/A
12. Sponsoring Agency Name and Address Connecticut Department of Transportation 280 West Street Rocky Hill, CT 06067-0207	13. Type of Report and Period Covered Final 14. Sponsoring Agency Code N/A

15. Supplementary Notes

N/A

16. Abstract

This report contains four categories of information: a summary of an evaluation of the potential for use of commodity flow data for the truck component of Connecticut's state wide traffic model; the development of new truck generation models using a town-based synthetic OD for the state; an evaluation of the use of satellite-imaged data for traffic prediction; and recommendations for future freight transportation planning needs for Connecticut.

Given the limited and non-comprehensive freight data sources available, most state level models that incorporate freight or truck traffic are usually not validated against real world observations. Moreover, a thorough comparison of freight and truck data sources has not been undertaken. In this report, five freight or truck data sources were compared and the methodologies needed for their use in the State of Connecticut's truck generation model were evaluated. The current Connecticut truck generation model, which is based on data originally collected in the early 1970s, was found inconsistent with observed truck counts and truck data estimated from other public and private freight data. The objective of this section of the project was to combine available information to develop a defensible updated truck generation model for the state. Ultimately, a town-based synthetic OD estimation was adopted. It was based on existing truck classification counts performed between the late 1980s and early 2000s. A seed OD matrix was developed from Reebie Associate's TRANSEARCH data. The estimated OD was then applied at the TAZ level using linear regression with zonal employment and population data.

17. Key Words	18. Distribution Statement
State wide planning model, freight, truck volume	No restrictions. This document is available to the public through the National Technical Information Service Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	7	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimetres	mm
ft	feet	0.305	metres	m
yd	yards	0.914	metres	m
mi	miles	1.61	kilometres	km
		AREA		
in^2	square inches	645.2	millimetres squared	mm^2
ft^2	square feet	0.093	metres squared	m^2
yd^2	square yards	0.836	metres squared	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	kilometres squared	km^2
		VOLUME		
fl oz	fluid ounces	29.57	mililitres	mL
gal	gallons	3.785	Litres	L
ft^3	cubic feet	0.028	metres cubed	m^3
yd^3	cubic yards	0.765	metres cubed	m^3

NOTE: Volumes greater than 1000 L shall be shown in m^3

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9	Celcius temperature	°C

* SI is the symbol for the International System of Measurement

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimetres	0.039	inches	in
m	metres	3.28	feet	ft
m	metres	1.09	yards	yd
km	kilometres	0.621	miles	mi
		AREA		
mm^2	millimetres squared	0.0016	square inches	in^2
m^2	metres squared	10.764	square feet	ft^2
ha	hectares	2.47	acres	ac
km^2	kilometres squared	0.386	square miles	mi^2
		VOLUME		
mL	millilitres	0.034	fluid ounces	fl oz
L	litres	0.264	gallons	gal
m^3	metres cubed	35.315	cubic feet	ft^3
m^3	metres cubed	1.308	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.102	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C+32	Fahrenheit temperature	°F

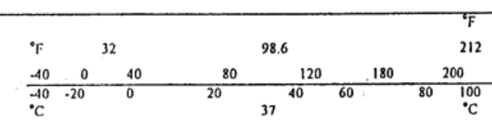


TABLE OF CONTENTS

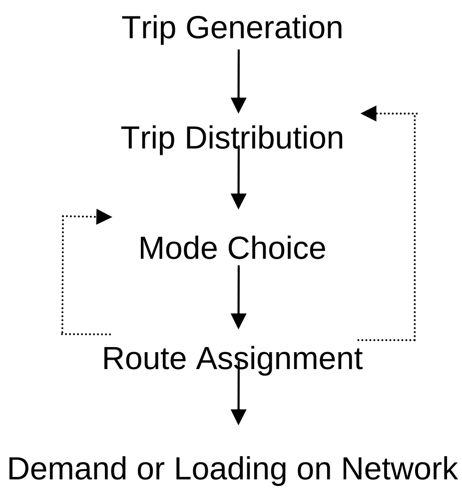
LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The Connecticut Department of Transportation, like many other states, maintains a state-wide traffic demand model for transportation planning purposes. Traditionally transportation demand modeling is undertaken in four stages as shown in Figure 1. The planning models involve trip generation (predicting the trips produced by or attracted to a traffic analysis zone), trip distribution (matching trip productions and attractions to determine trip interchanges or the number of trips between each pair of zones), mode choice (estimating the trips made by different modes such as auto or bus), and traffic route assignment (predicting the route as a sequence of roadway links to be used by each trip). These models require the study area to be divided into traffic analysis zones where the number of trips produced and attracted to each zone is predicted. The Connecticut model has just under 2000 zones. The overall planning model is used for infrastructure planning, evaluation of alternatives, emissions analysis and traffic management decision-making by the Department of Transportation.

Figure 1-1: Demand Modeling Stages



The original objective of this research was to evaluate the 1997 Commodity Flow Survey (CFS) and the Reebie Associates Transearch database for use in the state-wide traffic planning model and to develop truck model components that fit into the state traffic model context. More specifically, an appropriate new freight zone and network structure was envisioned. During the first year of the project, this objective was pursued. A significant amount of effort was invested in converting the ConnDOT planning model for use in the software TransCAD at the University of Connecticut. Moreover, the planning model truck generation levels were compared to various sources of freight commodity data being used frequently by others for freight planning model input.

The following is a summary of the major findings from Year 1 of research project.

1. The ConnDOT traffic planning model was successfully run in TRANSCAD and the truck component was evaluated and compared to 1997 USDOT Commodity Flow Survey, the 1996 Reebie Associates Transearch data, the FHWA Freight Analysis Framework Data, and traffic classification counts from ConnDOT.
2. The existing ConnDOT model was significantly overestimating truck traffic but the exact magnitude of the overestimate could not be determined without actual current data.
3. The inclusion of the parkways in the model significantly affects the truck traffic volumes estimated on different routes in the network and it is recommended that an approach be developed to prevent trucks from being assigned to the parkways in the model[1].
4. An updated truck component of the ConnDOT traffic model cannot be based on commodity freight data due to various data source limitations as applied to the Connecticut transportation network context and the nature of the state economy.
5. Traffic classification count data from ConnDOT are the most reliable and appropriately scaled data currently available to update the truck trip table in the model.
6. The Synthetic OD technique is a promising approach based on trial runs with a limited number of 1997 truck counts. However, its success is limited by the small number of traffic counts classified by vehicle type (100-200 unidirectional counts per year at 50-100 sites).
7. Collecting truck data via phone survey was considered by the research team and technical advisory committee. It was considered feasible but risky. Many questions on finding a representative sample and knowledgeable staff to interview were raised.
8. Getting more truck classification count data for use in the synthetic OD technique was recommended by the team and technical advisory committee as the most reasonable approach to improving the truck model in year 2 of this project. It was hoped that more truck volume data might be available from MPOs and cities, but a complete investigation including phone interviews uncovered none. An offer to conduct a limited number of additional counts with CTI research team assistance was offered by ConnDOT. The

[1] ConnDOT has developed a method to prohibit trucks being assigned to the parkways which will be used in the next model series.

UConn team used project resources, including limited additional funds, to obtain these additional counts in year 2.

9. Given the limited number of truck classification counts, the size of ConnDOT TAZs was deemed too small for the truck component of the model and the number of TAZs was deemed too large for applying the synthetic OD process. Therefore, it was proposed to use town level TAZs for the synthetic OD in the next stage of research, but to develop linear regression models appropriate for use at the TAZ level.

In light of these year 1 conclusions and recommendations the following adapted year 2 work plan was proposed by the research team and approved by JHRAC. The results of Task 1 through 9 are described in Chapter 2 of this report, while Task 10 is described in Chapter 3. Conclusions are contained at the end of each of these chapters but overall recommendations for future freight transportation research needs in Connecticut and the region are contained in Chapter 4.

Year 2 Task Plan:

1. As commodity-based data were not to be the data source in the next phase of this project, it was no longer desirable to use the 1996 ConnDOT model as used in year 1. Therefore, the 2000 ConnDOT planning model was converted to TransCAD (network, truck table and total traffic tables were needed) for use at UConn.
2. A new set of TAZ centroids corresponding to CT town centroids was developed. The external zone centroids were retained.
3. A new planning network was generated in Transcad with the town centroids and the ConnDOT 2000 planning highway network (travel on the parkways was prohibited for trucks).
4. All of the existing traffic classification counts for years 1999, 2000 and 2001 were aggregated and spatially located on the network using GIS programs. Counts in different locations between 1987 and 2002 were adjusted for year and also used.
5. Other agencies were contacted for additional existing truck volume counts but none were found (The contact list and documentation of the information requested are shown in Appendix A).

6. The spatial distribution of the existing counts were evaluated and a list of locations where counts would be desirable was formulated. Counts were obtained for the locations listed in Appendix B using the techniques outlined in Appendix C. This information was provided to ConnDOT as it was collected and therefore has been added to the overall inventory of traffic counts.
7. The existing 2000 truck trip table obtained from ConnDOT was aggregated to the town level.
8. A town based (169 x 169) synthetic OD was estimated using several different seed matrices including the ConnDOT 2000 truck OD matrix. The passenger traffic from the ConnDOT model was used as background traffic and was assigned in advance.
9. The town-level truck production/attraction from the synthetic OD was used as the dependent variable in a linear regression model. Town population and employment were used as possible predictor variables. Regression models did not have a constant term so that they would be directly transferable for use in the model TAZs.
10. In a related task, land use variables (obtained from satellite imaging) were evaluated as potential predictors of total traffic generation using the year 2000 TAZ polygons from the ConnDOT model.

CHAPTER 2: COMPARING AND INTEGRATING DATA SOURCES TO UPDATE THE TRUCK GENERATION MODEL

2.1 INTRODUCTION

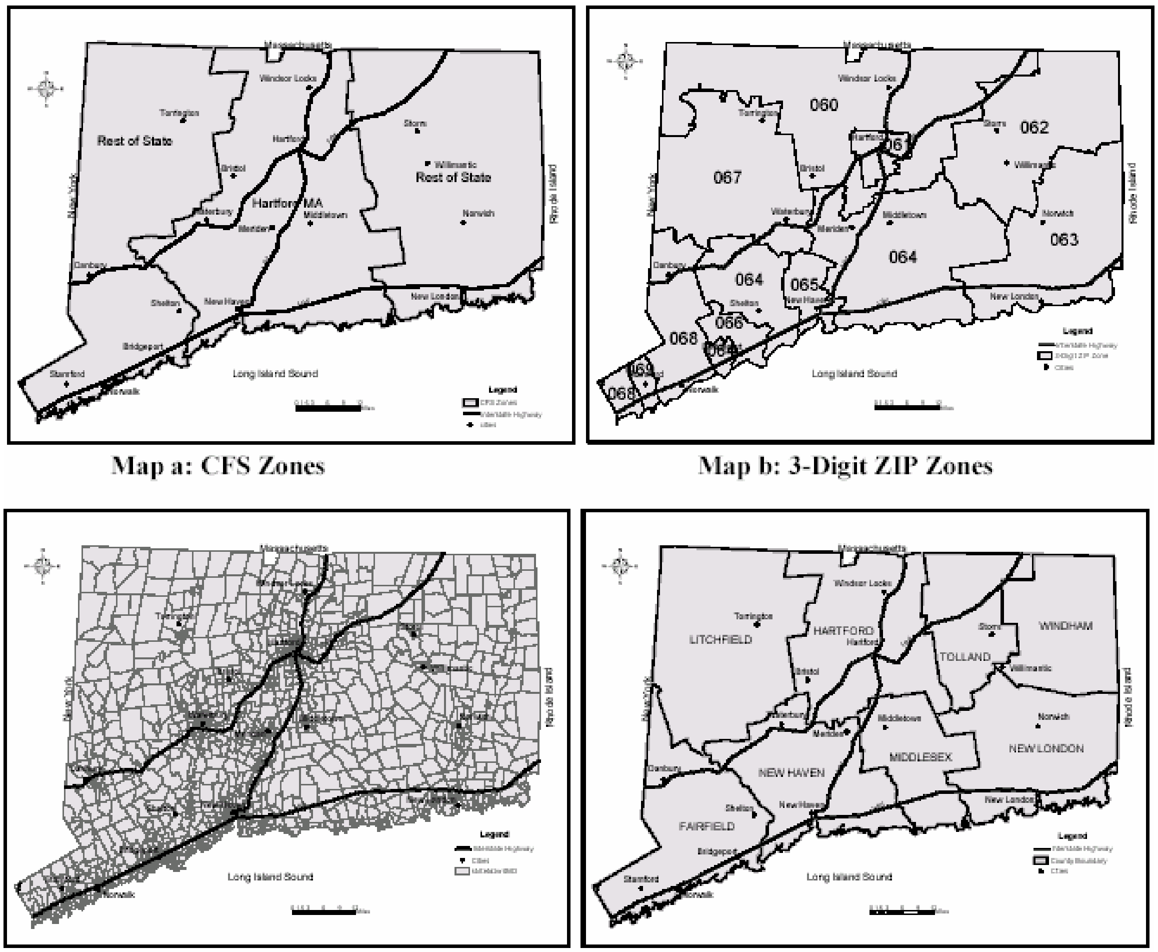
The Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA, 1991) and the Transportation Equity Act for the 21st Century (TEA21, 1998) required states to pursue freight transportation planning. However, the lack of appropriate freight data constitutes a major obstacle for the state planning efforts. The objective of the work described in this chapter was to update the current truck component of the Connecticut Department of Transportation statewide planning model. The state of Connecticut originally developed a four-stage transportation planning model in the 1960s. While the passenger component has been updated in the interim, no extensive new truck generation or other freight data had been collected since the 1970s. Consequently, the truck generation models, which consist of linear regression models by Traffic Analysis Zone (TAZ) as a function of population and total employment, have remained essentially the same over the past three decades. Given the amount of development and growth in Connecticut, as well as changes in the nature of freight transport, the economy, and the physical dimensions of trucks over the last 30 years, a new model based on updated data was needed.

Although there are several freight/truck data sources available, they cannot be readily used in state level planning because they lack geographic resolution and are not recorded using the same units. The TAZ is the base spatial unit used in transportation planning. Most stages of planning models, such as trip generation, trip distribution, and trip assignment, are associated with the TAZs. Map c in Figure 2-1 shows the 1996 version of the Connecticut transportation planning TAZ system. Unfortunately, the majority of freight commodity flow and socio-economic data are collected in geographic units that do not match the small size of the TAZs that are developed primarily for passenger traffic planning. The amount of detail that can be abstracted from freight data depends on the size of the zone in which they are collected. Although collecting data on the smallest scale possible is ideal, in that aggregation is easier than dis-aggregation, this is not practical for freight data. Although sometimes data are collected at high geographic resolution, they have to be aggregated to larger zones before public distribution in order to protect proprietary shipper information. Hence, in most cases only highly spatially aggregated freight data are available to public sector modelers.

Three sets of freight zone structures used by freight data providers are shown in Figure 2-1: 1) the three-digit ZIP code zones in map b; and 2) the county zones in Map d, which are used by the Reebie TRANSEARCH® freight data set; and 3) the Hartford metropolitan area and "rest of state" zones in Map a, which are used by the 1997 Commodity Flow Survey (CFS). The CFS freight data contain just two zones for the whole state of Connecticut. TRANSEARCH provides data at two slightly better resolutions: eight counties or ten three-digit zip code zones. In contrast, the state level planning model has a highly refined TAZ system: the 1996 version of the Connecticut transportation planning model contains about 1300 TAZs and the 2000 version has more than 1800 TAZs. Ideally, the freight data would have similar geographic detail, but none of the available data do. Collecting freight data for this level of geographic detail is cost prohibitive and shipper confidentiality issues arise. At the starting point in this project, we assumed that freight or truck modeling would not necessarily need to be conducted with the same established zone structure as the passenger travel model. However, phone interviews conducted with 16 other states confirmed that this was common and desirable from the point of view of application of the freight and passenger models (the telephone survey used in this project is shown in Appendix D, while the results are described briefly in Appendix E). Therefore, a realistic solution was to disaggregate freight data to the TAZ scale based on employment and population.

A second challenge in the use of existing freight data for state planning models is the use of different measures of freight flow (tonnage/value versus truck volume) in different datasets. This makes it hard to compare or combine different freight data. The freight data used in this study can be categorized into two groups: truck-based or commodity-based. The TRANSEARCH data from Reebie Associates provide both commodity tonnage and truck volume. The commodity-based category also includes the 1997 CFS data and the Waterborne Commerce data. The truck-based data include 1) a 1970s CONNDOT truck survey that is used in the 1996 version of the Connecticut truck generation model; 2) the truck volume classification counts routinely collected by DOTs; and 3) the Freight Analysis Framework (FAF) data from the Federal Highway Administration (FHWA).

Figure 2-1 Freight Data Zoning Systems



Map a: CFS Zones

Map b: 3-Digit ZIP Zones

Map c: 1996 ConnDOT TAZ Zones

Map d: County Zones

While commodity-based data can be converted to truck volumes, it is not possible to estimate commodity tonnage accurately based on observed truck counts, which do not contain any commodity value or tonnage information. Therefore, all freight data had to be converted to truck loads for comparison. The conversion is complicated by lack of complete coverage of freight commodities data, as well as the lack of information on empty, service, or utility trucks. A careful tonnage to truck conversion is essential in order to compare the data sources in this study. Although none of the available freight data were ideal for the task, the high cost of freight/truck data collection makes it unrealistic to collect new data for planning purposes. The available data were evaluated and compared as an initial stage of this research. Current state-level freight/truck planning models typically use one data source exclusively and the validation of the models is not routinely possible. It is critical to understand the pros and cons of different data sources and their impacts on model accuracy before endeavoring on a model update as planned here.

The comparison of the data conducted here indicates inconsistency between the data sources. The observed truck counts are considered most accurate but of course apply only to specific locations which are limited in number. Given the incomplete coverage of commodity-based data and its aggregate nature, a synthetic OD approach was employed to estimate a town level truck OD matrix based on observed link truck counts. The TRANSEARCH commodity data were used to generate a seed OD matrix as the base input for synthetic OD estimation. Linear regression models were then developed based on the population and employment. The intercepts of the linear models were excluded so that the models can be directly applied at the TAZ level. The results of this study provide useful direct guidance for state level freight data selection and model updating.

2.2 BACKGROUND

The public sector interest in freight transportation modeling has increased in the last decade. Responding to this need, a series of guidebooks were published, and several states have developed comprehensive state-wide freight transportation planning models. The Quick Response Freight Manual (Cambridge Systematics Inc. 1996) is a general guideline for applying the four-stage transportation planning process to freight traffic at the state/regional level. The manual addresses the characteristics of freight transportation and the effect on different stages of planning. The manual also introduces several case studies at different levels. One common characteristic of many statewide freight planning efforts is the use of the four-stage planning method (Cooghan 1996). Most, but not all, of the states contacted as part of the background investigation for this project use the four stage planning process for freight. NCHRP Report 388 (Cambridge Systematics Inc. 1997) has been the most comprehensive study of modeling approaches in freight generation to date.

In considering the update of the truck component of the state-wide model, the research team considered several modeling approaches. Linear regression models, time series analysis, growth rates and input/output models, are the common techniques used for freight generation models (Oum 1989, Tadi 1994). Linear regression models are straightforward to estimate at the zonal level and the coefficients can be easily interpreted from a transportation engineering point of view for policy evaluation. The time series approach is used to predict the future volumes based on historic data trends. This method is appropriate to predict the freight generation for a particular site (Al-Deek 1998). However, it does not provide causal explanations for changes in freight generation which limits its use as a good policy tool. The method of using commodity-specific annual growth rates is easy to apply, but deciding the appropriate rate is difficult and cannot reflect the spatial reallocation of demographic and economic variables. Input-output (I/O) models simulate the flow of goods and services between different sectors in an economy based on certain internal and external parameters. There are some applications of using I/O models for freight generation (Vilain, 1999; Cascetta, 1996), but they represent a complete departure from the traditional four-stage model used for the state model in Connecticut. Changing the modeling paradigm might be worthwhile, but only if pursued as a large endeavor for both passenger and freight planning model components. Therefore, the team determined that the most useful approach in this case was to update the linear regression models for truck trip generation based on newer data sources.

The two freight commodity data sources, which were first considered in this research, are commonly used by others, including both academics and state agencies: the CFS data (Huang *et al.* 1999 and Krishna and Hancock 1998) and the Reebie Associates TRANSEARCH dataset (Huang *et al.* 1999, Aultman-Hall 2000). Some states and regions have conducted surveys along truck routes or of industries that ship goods. Many states calibrate their model results against classification volume counts taken along highways.

Another approach coming into wider use at the state level for freight planning is modeling commodity flows by commodity group rather than simply estimating truck flows (Tadi, 2000; Eatough 2000). The first advantage of modeling commodities is that mode choice or intermodal policy options can be more easily evaluated. Second, the volume-tonnage ratio and value-tonnage ratio could be substantially different from commodity to commodity, and therefore conversion to vehicle units is more accurate when commodities are known. Souleyrette *et al* (1998) proposed a “layered” architecture for modeling the statewide freight transportation demand by simulating traffic for one commodity at a time. In Souleyrette’s study, the model was applied to the state of Iowa to estimate the freight production/attraction for the meat industry and farm machinery. Although estimating freight demand by commodity type is potentially more accurate and better describes potential causal relationships in the freight generation, the layered model requires detailed freight and social-economic data and meaningful correspondence between categories of commodities and employment. This approach was considered impractical for this application and was not pursued.

In summary, despite recent improvements in the data sources, limitations in the available freight flow data, particularly by commodity group, remain a major obstacle for state-wide freight planning. Freight information is hard to collect because of the shipper confidentiality issues and the many individual vehicle units and companies involved. The existing freight data sources cannot satisfy all freight analysis requirements due to deficiencies in coverage of commodity types, transportation modes, or lack of geographic detail. Furthermore, it is difficult to merge different sources of freight data together to a common unit or to compare them across regions or categories. While more advanced approaches are available, coordinating the truck component of the planning model with existing passenger models necessitates use of truck-based linear regression models.

2.3 COMPARISON OF DATA SOURCES

As there are two types of data (commodity and truck counts) as well as various geographic scales being compared in this study, it was necessary to convert the data or apply it in the model to facilitate comparison. Two types of comparisons were used: one aggregated and one disaggregated. The first one compares total freight tonnage or total truck demand for the entire state. Second, a traffic assignment procedure was used in the state-wide model to obtain link truck volume from datasets where truck origin/destination matrices could be developed. These link volumes were then compared to actual classification counts or model volumes from other data sources. Comparisons were conducted with 1996 and 1997 data sources and the 1996 version of the state-wide model was used. The following sections describe each data source individually.

ConnDOT State-wide Planning Model (1996 version)

The existing ConnDOT truck generation models are based on a truck flow survey conducted in the 1970's. Linear regression models with population and employment as independent variables were used to estimate the truck production and attraction, which are assumed equal. A town specific split indicator was introduced to split the internal and external truck trips. The following regression models are embedded in the model to estimate truck trip generation by zone.

Internal Truck Trips

$P=A=1.244743*(0.097*POP+0.258*EMP)*(100-I_{Avg})/100$

External Truck Trips

$P=A=1.244743*(0.097*POP+0.258*EMP)*I_{Avg}/100$

Where P = trips produced in the zone
A = trips attracted to the zone
POP = population of zone
EMP = total employment in zone
I_{AVG} = zone constant

Socio-economic data are only available for 1995 and 2000 and truck trips are estimated for these two years. The truck demand is interpolated for 1996, which is the base year in this study. The results indicate a total truck freight generation of approximately 900,000 trucks per day. Note that this number is nearly an order of magnitude greater than other values presented below. At the time of original model development, trucks were smaller and all type of trucks, not just large single unit trucks and tractor trailers, were classified as trucks in this model. This included pick-up trucks and vans. This further emphasizes the need for the model to be updated.

The ConnDOT truck generation models were also used to estimate the link volume. In this procedure, the passenger traffic was assigned to the network first as background volume. The truck traffic was then assigned to the network using a standard user equilibrium traffic assignment method. The output link truck volumes were compared to other truck count data when possible.

1997 Commodity Flow Survey

The CFS is undertaken as a joint effort of the Bureau of the Census, the U.S. Department of Commerce, and the Bureau of Transportation Statistics (BTS). The two most recent CFS datasets were collected in 1993 and 1997. The 1997 CFS provides more geographic detail than the 1993 survey. The CFS collects commodity origin and destination data by value, weight, and mode of transportation from approximately 100,000 manufacturing, mining, wholesale, and select retail establishments from a universe of about 800,000. The selected establishments were required to report data for 25 outbound samples for a one-week period in each of the four quarters in the survey year. The establishments classified as farms, forestry, fisheries, governments, construction, transportation, foreign establishments, services, and most retail establishments are **not** included in the survey. The CFS survey is mandatory which ensures the CFS covers all commodity groups in the surveyed industries.

The geographic coverage of the 1997 CFS is nationally stratified by state and selected Metropolitan Area (MA), which are small study areas that have high levels of economic activity. These metropolitan areas and the remainder of the state areas called "rest of the state" zones are the smallest geographic coverage in the 1997 CFS. There are only two zones in Connecticut, the Hartford metropolitan area and "rest of the state", which are shown on map a in Figure 1. It should be noted that the "rest of the state" or non-MA zone consists of two non-adjacent areas, which could be a problem for determining the centroid of the zone as well as any dis-aggregation undertaken with the CFS data. The geographic detail of the CFS is a limiting factor to its use at the state wide planning level for Connecticut.

To make the CFS data comparable to the existing Connecticut truck generation model, the data for the truck mode were extracted. The parcel and postal service freight is reported as a separate mode and were included here because these shipments are assumed to be made by truck. Freight measured by weight was used so that conversion to trucks could be made. The data illustrate, not surprisingly, that the truck mode dominates freight transportation in Connecticut, accounting for 91.36% of the total freight.

The 1997 CFS truck data and 1997 CBP employment data were used to fit a truck freight generation regression model for the six New England states.

Freight production (Tonnage) = 0.0315* Total Employment
R-Squared: 0.838

The model was then applied to the 1300 zone year 1996 TAZs in the Connecticut model using zonal employment data provided by ConnDOT. The estimated freight production for the truck mode is 48,969 (1000 tons/year) which equates to 9795 trucks/day if one assumes 20 tons per trucks and 250 working days per year. This truck volume is about two magnitudes lower than that of the ConnDOT model. Note again that the CFS does not cover all types of establishments. This missing data can result in significant bias. Analysis of the employment by sector for Connecticut (Table 2-1) indicates that the retail trade, services, and transportation sector account for nearly 60% of the total employment. Since establishments from these economic sectors are not included in the 1997 CFS, it was assumed that the Connecticut CFS data is much lower than the actual freight generation and that the CFS data on its own could not be used to update the planning model.

Table 2-1 Employment by Class for Connecticut

SIC code	Industry	Employment	Employment Percentage
	Total	1,471,970	100.00%
07--	Agricultural Services, Forestry, and Fishing	8,478	0.58%
10--	Mining	1,426	0.10%
15--	Construction	57,965	3.94%
20--	Manufacturing	289,670	19.68%
40--	Transportation and Public Utilities	72,655	4.94%
50--	Wholesale Trade	90,544	6.15%
52--	Retail Trade	271,502	18.44%
60--	Finance, Insurance and Real Estate	144,734	9.83%
70--	Services	534,677	36.32%
99--	Unclassified	319	0.02%

Source: U.S. Census Bureau

Waterborne Commerce Data

The existing truck generation model and the CFS models above only consider the total employment and population in modeling freight/truck generation. However, zones with intermodal facilities could have different freight/truck generation patterns that are not explained using the employment/population variables because the intermodal facilities themselves are a production/attraction point with large freight flow and thus truck flows.

The Waterborne Commerce Data, which are maintained by the U.S. Army Corps of Engineers, consist of annual total inbound/outbound freight tonnages and values by commodity, harbors, waterway segment, direction, and type of movement for domestic waterways. These data are not origin or destination based and therefore, cannot be used directly for planning models including generation models.

There are four principle ports in Connecticut, which are located in Stamford, Bridgeport, New Haven, and New London respectively. The total tonnages for these ports are listed in Table 2-2, together with the approximate number of trucks converted from the reported tonnage. The last two columns of Table 2-2 indicate the number of trucks or tonnages predicted using the ConnDOT and CFS models. These data also suggest ConnDOT's truck data are much higher than other data for these TAZs. At the same time, comparison of the CFS data to the port data also confirms the limitation in coverage of the CFS data. It is interesting to note that the discrepancy here was greater than that found in similar work conducted recently for the state of Kentucky (Scott 2003). This suggests that the usefulness of the CFS data for freight planning are related to the nature of the economy within a state.

Table 2-2 Connecticut Principle Ports and the Annual Freight Tonnage (1997)

PORT NAME	1997 Total Tonnage (tons/year)	Ports Truck Flow (300day/year) (20tons/truck) (trucks/day)	TAZ ID	ConnDOT model (trucks/day)	CFS freight generation rate model (tons/year)
Stamford, CT	1,040,926	173	*322*	1426	65,945
Bridgeport, CT	5,340,257	890	*39*	2909	220,930
New Haven, CT	9,593,835	1599	*588*	1512	134,056
New London, CT	608,291	101	*1147*	1021	67,380

Source: US Army Corps of Engineers

TRANSEARCH

TRANSEARCH is a private freight database with nation-wide coverage produced by Reebie Associates in Connecticut. It provides integrated multimodal freight information by Standard Transportation Commodity Code (STCC) with volumes in terms of loads, tonnage, or value at the county, zip code, metropolitan area, state, or province level. The database is generated from a model that integrates many public freight data sources, a survey of manufacturers, and truck movement information from a data exchange program with motor carriers. The Reebie

TRANSEARCH data are collected at the 3-digit ZIP code zone level and then translated to county level. In some states, this represents an aggregation (California) while in others this represents a dis-aggregation (Kentucky). Connecticut has purchased the county level data for within the state as well as relatively disaggregated data for areas lying just beyond the state border. The data consist of commodity flows between 8 county zones within Connecticut and 13 external zones.

As indicated in Table 2-3, the number of truck trips from TRANSEARCH and those estimated from the CFS data are similar to each other. However, both are substantially less than the total truck trips predicted by the ConnDOT truck model. Overhead truck trips (external to external) make up almost half of the total Reebie truck trips. This illustrates another challenge for public sector truck freight planning in this state. The large proportion of external freight origins and destinations cannot be easily targeted by state efforts for modal substitution or other traffic management strategies. The truck mode is critical in Connecticut, carrying over 70 percent of the total commodity tonnage reported by Reebie. According to TRANSEARCH, the top commodities by truck trips include consumer and construction related goods. This suggests that databases that include primarily basic aggregate commodities are a poor choice for freight planning in Connecticut.

Table 2-3 Daily Truck Trips of ConnDOT Model, CFS, and Reebie Data

	ConnDOT Model (‘96)	CFS (’97)	Reebie (’96) (exclude External-External)	Reebie(’96) External-External
Truck Trips	907,244	9,768*	10,188**	9116**

*Computed from tonnage, the default truck load is 30 tons and 245 working days.
**Computed from annual truck trips, assuming 245 working days.

Freight Analysis Framework Data for CT

The Freight Analysis Framework (FAF) is a national level freight planning model developed by the FHWA. The major data source for the FAF is the county-to-county level Reebie Associates data. Two kinds of truck link flows are provided publicly for the nationwide truck highway network: the FAF truck volume, which is the commodity-based truck volume generated from the planning model; and the total truck volume, which is the observed truck link volume (HPMS) as provided by state DOTs. In this study, Geographic Information System (GIS) spatial analysis was used to compare the FAF truck volumes to both the classification counts and the current

ConnDOT's planning model output. There are 40 locations where classification counts from 1998 matched the FAF network links. The bi-directional truck counts were aggregated because FAF provides only total two-directional volume. Details of the comparison are shown in Appendix F.

Within the FAF data itself, the average absolute difference between the FAF total truck volumes and the classification counts is 2251 trucks per day. Overall, the FAF volumes are much higher with an average difference of 1990 trucks per day. Given that the source of data for the total trucks in the FAF dataset is the HPMS, this difference is not expected. The absolute difference between the total FAF truck volumes and the ConnDOT model output is smaller than that between FAF and truck counts. Compared to the FAF, the ConnDOT model is underestimating trucks on the interstates while overestimating on the non-interstate routes.

The FAF commodity-based truck volumes are much smaller than the HPMS observed total truck volume reported in the same dataset. The national average FAF commodity-based truck volume to total truck volume ratio is 0.43, ranging from 0.1 to 0.7. This ratio for Connecticut is 0.22, which implies that the trucks estimated from commodity tonnage only represent 20 percent of the total trucks in Connecticut. These results confirm the analytical results of the Reebie Associates CT data described above: the commodity-based truck data do not sufficiently capture all the truck traffic in the state and are not useful on their own for traffic demand models.

Truck Classification Data

Truck classification count data were provided to the study team by the Connecticut Department of Transportation. The data set contains 2697 traffic count stations where observations were made between 1987 and 2002. Each count station contains a unique station ID, route number, milepost, direction, collected year, and traffic count by 12 vehicle classes. The count stations were associated with planning network links through GIS spatial analysis programs followed by a manual check. The output lookup tables connect the count stations and the ConnDOT planning network links for comparison. Given the year for other freight data sources (1996 for Reebie, 1997 for CFS, and 1998 for FAF), counts from 1996, 1997 and 1998 were included in this analysis.

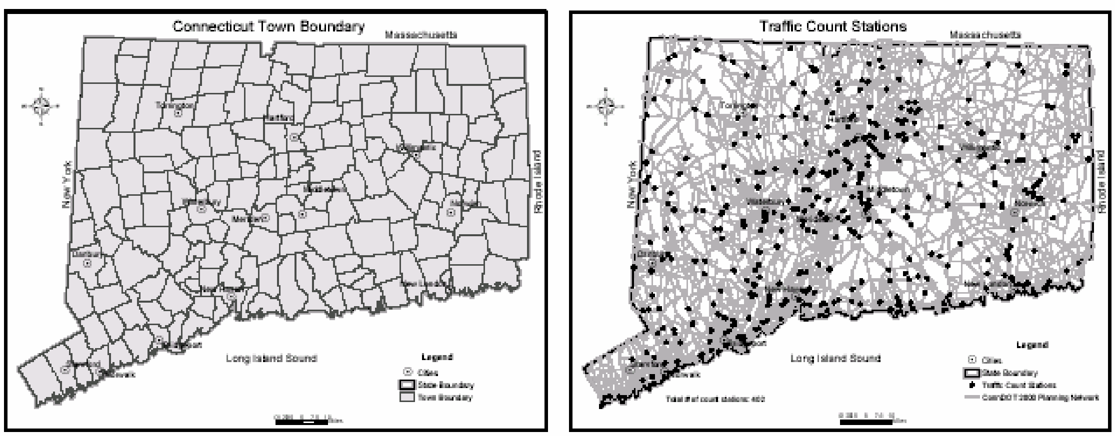
There are 449 count stations in 1996-1998. Coverage is spread throughout the state and on both interstate and non-interstate routes. The comparison of existing count data to the planning model (See Appendix G) indicates that the volumes from the model are higher than the truck counts. This again confirms that the existing model is overestimating trucks. This finding is consistent with results from the comparison of the ConnDOT model data to the Reebie and CFS data. The comparison includes both the absolute different and percentage different. Many of the large outliers occur at locations on or near freeways where the volumes are high and the percentage difference is not great. However, a large number of locations with high percent differences are located on lower volume highways where the actual truck counts are very low. Consequently, the percentage difference is great when any significant number of trucks is predicted by the model. In short, most of the modeled truck volumes are more than 100 trucks per day different from the actual counts and just over half of the volumes are more than 100% different. This further confirms that improvements could be made to the truck component of the state wide model. However, consideration of the commodity data in this project does not suggest its use alone would improve the model.

2.4 UPDATING THE MODEL: SYNTHETIC OD ESTIMATION

Given the inability of any existing dataset to fit the socioeconomic data or truck volume counts for the state of Connecticut, the project team turned to the creation of a synthetic OD to update the demand model. This decision was based in part on the fact that while classification counts are aggregate and spatially local, they are the most reliable source of data. The software package TRANSCAD, which integrated the OD estimation as well as four-stage planning functions, was employed. Synthetic OD estimation can be considered a "reversed" traffic assignment. A seed OD matrix and a certain number of observed link volumes are needed as input to the model. The output is an adjusted OD matrix, which produces a better fit for the observed link volumes when conducting the traffic assignment. The trip interchanges in the trip table are adjusted using multiplicative ratios until the error between the predicted and measured truck count volumes is minimized. In this case, the year 2000 CT model was used because it had updated zones and network structure. To maximize the number of traffic count stations, count stations from additional years were used when newer counts were not available. The data are converted to the base year 2000 using an annual growth rate of 5%. In addition, 52 additional 24 hour counts were performed in the spring of 2004 and adjusted for day of week and year. In total, 402 count locations were available as shown in Map b of Figure 2-2.

Although the number of the count stations is still limited compared to the number of TAZs in the state model, this is the maximum number of count stations that could realistically be obtained given time and budget constraints. To ensure the number of origins and destinations in the OD matrix was comparable to the number of count stations, a new set of TAZs based on 169 Connecticut towns were used. These are shown in Map a of Figure 2-2. The 52 external zones from the original network were used.

Figure 2-2 Synthetic OD Zones and Traffic Count Stations



Map a: Synthetic OD Zones(Town-based)

Map b: Traffic Count Stations and ConnDOT 2000 Planning Network

The analysis of the several datasets described above led the team to consider the TRANSEARCH commodity data as the best choice for construction of a seed OD. The TRANSEARCH truck flows from the 1996 dataset were converted to year 2000 using an annual growth rate of 5%. The TRANSEARCH annual truck flow was converted to daily truck volumes using 245 working days per year. A gravity model was used to convert the total truck production and attraction to an OD matrix at the town level. The total truck trip production and attraction in the seed OD matrix was 23,465 trucks/day. The ConnDOT year 2000 truck volumes produced and attracted to the external zones were used for the SEED. The passenger traffic was assigned to the network first and used as the preload traffic for the synthetic OD analysis. The output matrix from the synthetic OD procedure indicates a total truck generation of 114,066 trucks/day. The magnitude of this result agrees well with the CFS, TRANSEARCH and FAF data presented above if one considers the commodity coverage of each dataset. The difference between the assigned truck flows and the classification counts is minimized since this is the criterion in establishing the synthetic OD.

In order for the synthetic OD developed at the town level to be useful in the planning model context, TAZ-level linear regression models had to be established. Intercepts (constants), which represents the base truck production for each zone, are associated with the size of the zones and were not used in the models to allow for direct application of the regression models at the TAZ level. Three sets of models are shown in Table 2-3: one based on population, one based on total employment and one based on retail and non-retail employment. These are the predictor variables available in the ConnDOT demand model dataset. The total employment and population were not used together in one model (as they are in the original ConnDOT model) due to high correlation between them. All parameters are significant based on the t-test at the 0.05 level. The R-squared values indicate that the models using population as the predictor fit the data best and these are recommended as the updated truck generation models. The regression models were not estimated using the external zones nor should they be applied to the external zones. The trip rates for external zones are listed in Appendix H and growth rates should be applied to these rates for use in future version of the ConnDOT model.

Table 2-4: TAZ Regression Models Results

	TRUCK PRODUCTION			TRUCK ATTRACTION		
Population	**0.01822**	--	--	**0.01474**		--
Total employment	**--**	0.02944	--	**--**	0.03022	--
Retail Employment	**--**	--	0.10743	**--**	--	0.09297
Non-retail Employment	**--**	--	0.01790	**--**	--	0.02094
R-squared	**0.64**	0.56	0.59	**0.55**	0.50	0.52

2.5 CONCLUSIONS REGARDING THE UPDATED GENERATION MODELS

Given the small geographic size of the state of Connecticut, the service and consumer-based economy, and the large proportion of external-external truck traffic on the state highway network, the recently popular commodity-based freight datasets cannot be used directly for the state-wide planning model. However, they do provide the best source of data for an up-to-date SEED OD matrix to indicate the pattern in freight traffic for a synthetic OD procedure. The most accurate and routinely collected source of information for a truck model was concluded to be the truck classification counts. Unfortunately, the number of count stations was not sufficient compared to the number of TAZs in a typical state planning model. While the disaggregated small TAZs are appropriate for passenger models, they do not match well with typical freight data sources. However, a uniform TAZ system is desirable from a practical point of view. To facilitate the synthetic OD procedure, TAZs were aggregated to the town level in the state of Connecticut. The 402 two-directional truck counts were a reasonable match for the 169 town zones. The resultant synthetic OD consisted of approximately ten times the truck trips as that predicted by the CFS or TRANSEARCH data. Given the nature of the state economy, this generation level was deemed appropriate. This level of generation is only about 10% of that in the existing model. In order to apply the truck generation at the TAZ level, a town-based linear regression model with no intercept was estimated using the town employment and population data. These models are recommended for use in the state-wide model where the current version has approximately 1800 zones with employment and population estimates. It is expected that this process may be a cost-effective method to update truck trip demand models in other states.

CHAPTER 3: USING SATELLITE IMAGED LAND USE DATA TO PREDICT TRIP GENERATION[2]

3.1 INTRODUCTION AND BACKGROUND

This portion of the research investigates whether trip generation (the number of trips starting or ending in a zone) can be predicted based on the type of land use in a zone as measured by satellite imagery. Typically trip generation is modeled with population and employment, but satellite imagery is readily available and has not been extensively considered as an alternative predictor. The total number of trips by zone from the Connecticut state-planning model are predicted in this study using linear regression as a function of zone population and total employment, and then subsequently compared to models based on the amounts of different types of land use in the zone. While it is ultimately of interest in the project to find additional predictors for freight trip generation, this investigation started with total traffic generation because this portion of the state wide model was considered more accurate than the current truck portion. A total of 1806 zones in the state of Connecticut averaging 2.75 square miles are used. The results of this study indicate promise that a new predictor of trip generation may be feasible and could be even more effective than measured here considering that satellite image datum are now more readily available with improved accuracy.

Trip generation is used to link land use and the travel associated with it on a zonal basis. Trip generation represents an estimate of the number of trips attracted to or produced by a traffic zone within an area. Both trip attractions and trip productions are referred to as trip ends because the final destination of the traveler is not part of consideration. The zones used in this study, traffic analysis zones (TAZ), are areas that are delineated, based on Census information by the state or planning agency in order to conduct transportation studies. The 1806 Connecticut TAZ areas, used by the Department of Transportation for planning purposes, are used in this study of land use and trip generation.

The typical variables used for trip generation models are population and employment for each traffic analysis zone. These variables are the typical standard for models that predict trip generation. However, in this case, land use determined using satellite imagery will be used to

[2] This work was conducted primarily by Mr. Chris O'Brien as an independent undergraduate research project and is also published as his undergraduate thesis. Mr. O'Brien was advised by Dr. Aultman-Hall and provided data by Mr. Guo and Ms. Padlo.

predict total trip generation (productions and attractions) as a function of land use. The models will be compared to those for the population and employment to see if the satellite imaged data shows promise for further study.

There have been other studies conducted that use satellite data for transportation purposes. For example, there has been research conducted on estimating average traffic flow from satellite imagery [Coifman, 2003]. However, in that study, the types of land were not considered, just the amount of traffic passing through each zone. The traffic volume information obtained from the satellite imagery was then cross-examined with concurrent ground-study information. This study concluded that the estimations based on the satellite imagery were more accurate than previously assumed, and when considered with the ground-study, the estimations were very accurate.

Satellite imagery has also been used to track urban growth development [Glutch, 2002]. This investigation of urban development used Landsat mapping merged with SPOT-P data and developed a map that defined “urban” versus “non-urban” pixels. The results from the study showed “growth” pixels with 92% accuracy when the “growth” pixels were compared with GIS analysis on environmental constraints to growth. When compared, the “growth” pixels were shown to coincide with areas that had actually become more urbanized, indicating that the information obtained from the satellite imagery is a reliable way of predicting urban growth. This data is intended to be used to predict, identify, map, and monitor urban growth and change for future uses.

3.2 DATA

Four types of data are needed for this study: land use, total trips per zone (productions and attractions), population and employment data. The land use data were collected using satellite imagery from the Landsat Thematic Mapper (TM) and the Multispectral Scanner satellites. The information was gathered and compiled in 1995. Twenty-three land types are classified by these satellites that lie in 1806 TAZ in Connecticut (see Table 3-1). The area of each type of land in each TAZ comes from GIS overlaying and this information has been provided to the planning personnel at ConnDOT for their use. The most common land use types are Soil/Grass/Hay, Grass/Hay/Pasture, Coniferous Forests, Wetlands and Forests, and Major Roads. All of the satellite information is stored in the Mapping and Geographic Information

Center located in the University of Connecticut library. The satellite imagery, which provides the land types has a high resolution of pixel images (30 x 30 m^2) and contains very limited areas among the 1806 zones in which it cannot distinguish the type of land use for that particular area. The classification algorithms are established by Landsat – TM.

Table 3-1: Summary Table of Land Type Linear Regression Models

Independant Variable	R^2	Intercept	t-Test p-value	Slope	t-Test p-value	Category
Surface-Impervious	0.240	8330.6	<0.001	0.003	<0.001	1
Residential/Commercial – High Density	0.150	8290.7	<0.001	0.002	<0.001	1
Residential – Medium Density	0.107	7769.1	<0.001	0.000	<0.001	1
Surface-Roof	0.117	9775.8	<0.001	0.021	<0.001	1
Road-Pavement	0.001	10605.1	<0.001	0.001	0.120	2
Turf/Grass	0.012	10104.1	<0.001	0.001	<0.001	2
Barren Land	0.035	9586.4	<0.001	0.001	<0.001	2
Bare Soil	0.008	10307.9	<0.001	0.000	<0.001	2
Soil/Grass/Hay	0.026	11332.8	<0.001	0.000	<0.001	3
Grass/Hay/Past	0.017	11163.4	<0.001	0.000	<0.001	3
Soil/Corn	0.007	10803.0	<0.001	0.000	<0.001	3
Grass/Corn	0.009	10838.9	<0.001	-0.001	<0.001	3
Soil/Tobacco	<0.001	10627.6	<0.001	0.000	0.845	3
Grass/Tobacco	<0.001	10627.1	<0.001	0.001	0.803	3
Forest - Deciduous	0.070	11918.3	<0.001	0.000	<0.001	3
Forest - Coniferous	0.031	10999.7	<0.001	0.000	<0.001	3
Water Deep	0.012	10896.1	<0.001	0.000	<0.001	3
Water Shallow	0.023	11267.1	<0.001	0.000	<0.001	3
Wetland - Non Forest	0.019	10950.9	<0.001	-0.005	<0.001	3
Wetland - Forest	0.036	11241.2	<0.001	0.000	<0.001	3
Low Coast Marsh	<0.001	10635.1	<0.001	-0.003	0.400	3
High Coast Marsh	<0.001	10621.4	<0.001	0.000	0.749	3
Major Road	0.003	10962.0	<0.001	0.000	0.031	3

The trip generation information was provided by the Connecticut Department of Transportation (ConnDOT). The data were modeled using population and employment figures gathered in the year 2000. The trip productions and trip attractions were provided individually, but for the purpose of this study the total of attractions and productions will be used as the dependant variable in the regression models. The average number of daily trips per zone is 10,630 with a standard deviation of 7,091. This large standard deviation is one indication of how much the number of trips varies from zone to zone.

The information received for the population and employment figures was obtained from ConnDOT but originally from the U.S. Census. The totals were tabulated by TAZ areas by the ConnDOT division of planning. The average population per zone is 1886 with a standard deviation of 1110. The average total employment per zone is 915 with a standard deviation of 1428. These numbers suggest, as is the case for land types, that there is a large variance in the population and employment among zones.

3.3 METHODOLOGY

Before any analysis was undertaken, the first step in this analysis was to combine and re-organize the large datasets for ease of model development. The land use data were not structured in an appropriate way for statistical analysis. First, the land type data were reorganized into a format where the amount of each type of land in each zone was in a single row of a tabular dataset. Second, the population, employment, and trip data were imported into the same data file using the M.S. Excel spreadsheet program. The final dataset had the following format: one row per zone with subsequent columns of the amount of each land type, population, employment and total trips.

The method used to model trips as a function of land use, population and employment was linear regression. Linear regression develops the relationship between independent and dependent variables and shows the correlation between them. The coefficients in the liner equation are estimated to minimize the sum of the squares of error between the observed points and the predicted dependent variable. The linear regression models were estimated in the M.S. Excel spreadsheet program. For this analysis, the dependent variable in all of the regression models will be the total trips produced and attracted to each of the 1806 zones specified by TAZ. The independent variables will be total zonal employment, population, and the area of different land uses for the zone.

In addition to the main research question as to whether land use can be used to predict trip generation, the quality of the land use models was compared to those based on the more traditionally used variables, employment and population using two goodness of fit measures for each regression model. One test is the R^2 statistic, which measures the scatter of the data around the predicted line. The closer the R^2 is to a value of 1, the less scatter exists, meaning the linear equation captures more of the variability in the data set. The second measure is the t-test for the coefficient (slope) values. A t-test is performed to determine if the coefficient is statistically different from zero, in other words if the slope is truly positive or negative.

The first linear regression models developed were for population and employment individually. These models are considered the base cases in this study as trip generation is traditionally predicted based on one or both of these variables. Once this task was accomplished, individual land use linear regression models were developed. There were twenty-three different land types and all were tested against total trips individually.

These 23 models were classified into three categories based on quality on the R^2 statistic and t-test results: 1.) good predictor of trips with average scatter of data, 2.) good predictor of trips with high scatter of data, and 3.) poor predictor of trips. After all the land types had been categorized only the best land use variables were examined for correlation and used in a combination model of the “good” land types. The land use variables were evaluated for correlation because if two variables are strongly correlated then the linear regression model that is developed would violate assumptions and the parameter coefficients are not interpretable.

The final step in the analysis procedure was to compare the combined land type model against the population and employment models to determine if using a combination of land types is a better predictor than one of the standard models typically used.

3.4 RESULTS

The first models developed were those for employment and population versus the total trips of a zone. These results are shown in Figures 3-1 and 3.2. The results from these models are relatively good and it is clear why population and employment are normally used to predict trip generation. However, the R^2 for the population model is 0.327 and the R^2 for the employment model is 0.595 indicating significant scatter. Also, the P-value from both of these models fell well below 0.05, indicating positive slopes not equal to zero. While population and employment are good predictors, based on scatter, there is still an opportunity to improve and have better and more reliable models for trip generation prediction.

Figure 3-1: Employment Linear Regression Model

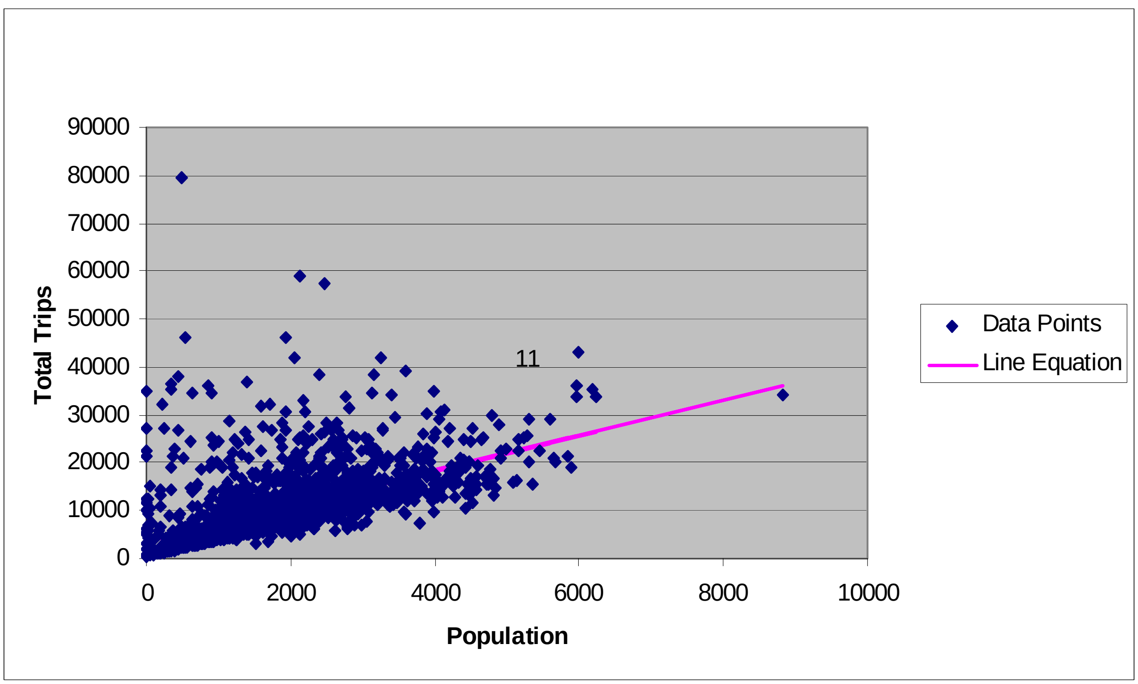


Figure 3-2: Population Linear Regression Model

All twenty-three land types (cited in Table 3-1) were tested individually with the total trips in the 1806 TAZ zones to determine if there were any correlations between the trip volume and the land use. A summary of the resulting linear regression models is shown in Table 3-1 with goodness of fit statistics. The category column refers to how good the variable is for trip generation prediction. This classification, which is based on the R^2 value and the t-tests is shown on the far right in Table 3-1. Category 1 is very good, category 2 is average, and category 3 is poor. The land types that fell into the first category are: Surface-Impervious, High Density Residential/Commercial, Medium Density Residential, and Roof Surface.

Figures 3-3 through 3-6 illustrate the linear regression models based on these category 1 land types. Visual inspection of Figures 3-3 through 3-6 indicate that as the amount of these four land types increase, so do the number of trip attractions and productions. In all four of the above land type figures, the scatter of data points is much higher and consequently the R^2 statistics are much lower than those for the employment and population models. But, the R^2 for these models, although low, is still much higher than any of the other land use models. All of these four models' t-tests yielded very low P-values suggesting the overall relationship between land use and travel is significant and valid. The next step was to examine a combination of land use variables. As a result, the land use models on an individual basis do not justify using satellite imagery as a means of predicting trip generation. For these several reasons, linear regression models combining these variables were pursued.

Before pursuing multivariate regression models, the correlation of these independent variables including population and employment was calculated. The results are shown in Table 3-2. Based on the correlation matrix, it was concluded that none of the land use variables were strongly correlated and therefore could be used in one multi-variable regression model. Also, the population and employment data for each zone were not strongly correlated with land use so a multi-variable regression model with these two variables was also possible. One last observation showed that none of the six variables in Table 3-2 were strongly correlated, so a final six-way multi-variable regression model was formed from all of these variables.

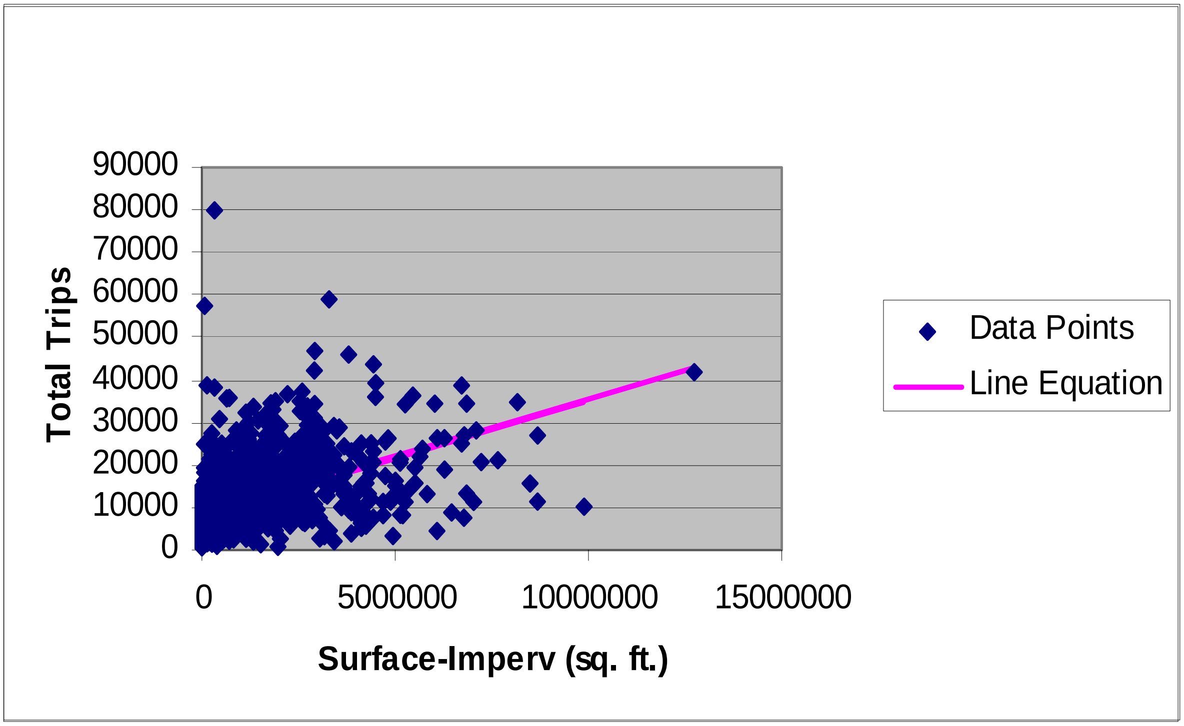


Figure 3-3: Impervious Surface Linear Regression Model

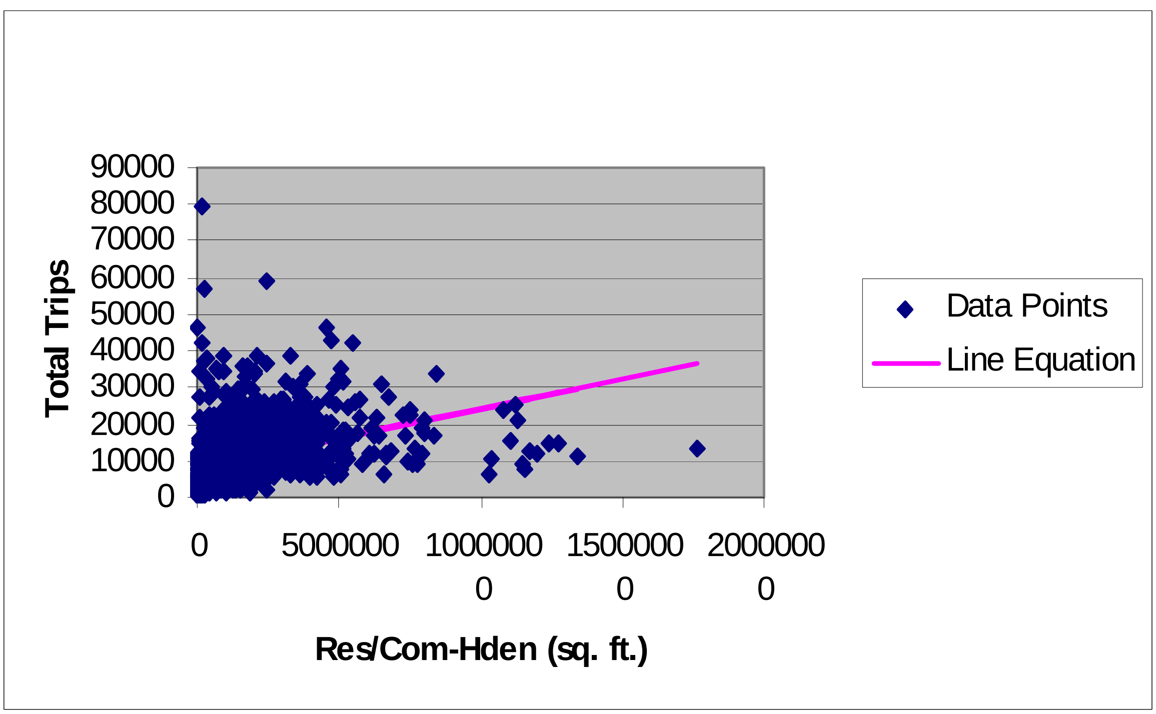


Figure 3-4: High Density Residential/Commercial Linear Regression Model

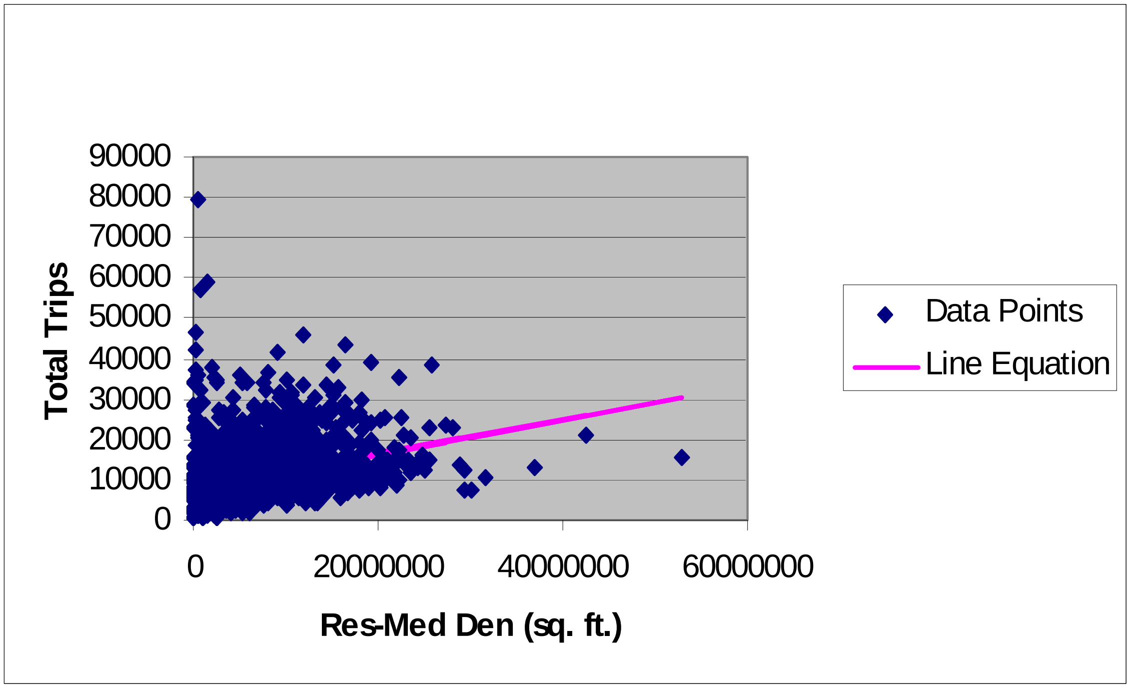


Figure 3-5: Medium Density Residential –Linear Regression Model

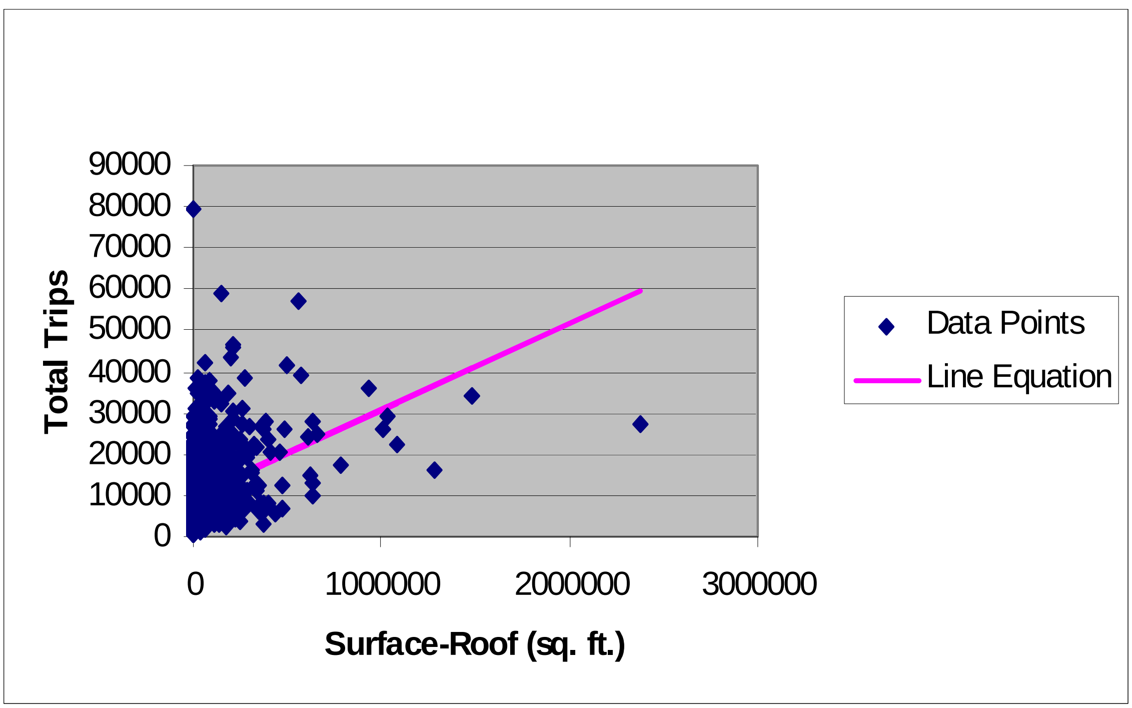


Figure 3-6: Roof Surface Linear Regression Model

Table 3-2: Correlation Matrix

	Surface Impervious	Residential/ Commertial High Density	Residential Medium Density	Surface-Roof	Population	Total Empl.
Surface Impervious	1					
Residential/Commercial High Density	0.254	1				
Residential Medium Density	-0.091	0.382	1			
Surface - Roof	0.496	0.082	0.012	1		
Population	0.224	0.441	0.466	-0.016	1	
Total Empl.	0.520	0.169	0.007	0.404	0.108	1

For the first combination model of the four land types, the R^2 statistic was 0.404 and the P-values were all well below 0.05. This is a key finding, because the R^2 statistic for the employment model was 0.595, while the R^2 statistic for the population model was 0.327. This shows that this combination of land types can provide a more accurate prediction of trip generation than one of the standard models used for trip generation prediction.

The second combination model was that of combined population and employment. For this model, the resulting R^2 statistic was 0.837, and the P-values were again well below 0.05. The R^2 statistic is very high and represents the best model with this data. This is the standard for comparison for the new models.

The final combination model developed was that of all four land types, population, and employment. The R^2 statistic for this test was 0.852, an improvement upon the second combination model. However, for this model the P-value for the impervious surface variable was 0.601, which indicates that it is indeterminate whether the slope is negative or positive. To account for this, one last regression model was developed including population, employment, and the category 1-land types excluding the impervious surface. In this model, the R^2 statistic was again 0.852 and this time no P-value exceeded 0.05. As a result, this model was deemed the most accurate predictor of trip generation, although the improvement over the model of population and employment is very limited suggesting land use can predict as well as one of population or employment but does not improve these models by adding additional information.

3.5 CONCLUSIONS REGARDING SATELLITE-BASED LAND USE DATA FOR TRIP GENERATION

The regression models developed here suggest promise for using satellite imagery to predict trip generation. The results of this study suggest that land use data can be used as a predictor of trip generation. The combination model of the four land types proved to be more accurate than use of the model based on only population. Moreover, when combined with population and employment, certain land use variables still improve the prediction of trip generation (although minimally).

This study suggests that because the satellite imagery is readily available it could be beneficial to transportation engineering and planning. Future studies should consider use of improved satellite image databases or focus on metropolitan areas' models where zones are smaller and more uniform. There are many reasons why trip generation increases as certain land types (those which produce or attract trips) increase regardless of other factors, and it is reasonable that land use is a causal predictor of trip generation.

CHAPTER 4: RECOMMENDATIONS FOR FUTURE FREIGHT TRANSPORTATION RESEARCH IN CONNECTICUT

This research project has provided several tangible products. First, updated truck trip generation models for the Connecticut state-wide traffic model were developed. Second, the use of freight commodity data for the model was evaluated, and although it is used by other states for direct planning model input, is deemed appropriate only for validation given the Connecticut context. Third, a survey of 16 other states revealed that most states lag behind Connecticut in that they do not consider truck or freight flows in their state wide models. Many expressed similar research needs as those being faced in Connecticut. Fourth, additional traffic classification counts were conducted throughout the state to complement existing counts and provide a more geographically dispersed set of data. This data were collected in cooperation with ConnDOT and can now be used for other planning purposes. Finally, the use of readily available satellite imaged land use data were evaluated as a predictor of trip generation at the TAZ level. These measures show promise for practical use in forecasting.

This chapter summarizes some outstanding issues in freight demand modeling that require research. These issues were encountered during this project, but were beyond the specific scope.

4.1 REGIONAL AND NATION-WIDE MODELS

The nature of freight transportation demands that the public sector efforts to model freight be undertaken at the national and regional scale in addition to the state level. However, most current freight planning and research focus on the state and metropolitan levels (Apffel et al. 1996; Aultman-Hall et al. 2000; Eatough et al. 2000; Huang and Smith 1999; Sorratini and Smith 2000). The freight trip length is usually longer than that of the personal trip. In the United States, for example, the 1997 Commodity Flow Survey (CFS) (US Census Bureau, 1997) found that the average trip length for all modes was 472 miles, however the average other modes is longer than truck-only trips (rail: 769; water, shallow draft: 177, great lakes: 204, deep draft: 1024; air: 1380). The long freight trip length leads to a significant proportion of the traffic having an out-of-state or country origin/destination. A state level planning model cannot sufficiently address the large number of external freight trips. The trip length is also one of the key factors in deciding the mode of shipment. The lack of a continental-level or regional scope for modeling is particularly problematic for evaluating new optimal intermodal facility sites and for studies of modal substitution. Furthermore, most freight corridors span several regions, states or countries, thus requiring multi-region models. When freight has nation-wide or global origins and destinations, it is difficult to use the state boundary for modeling. National level or regional models are a necessary complement to smaller area models in that they cover the national, international and intermodal freight trips and larger facility networks. Therefore, it is recommended that in addition to further development of the Connecticut state model that ConnDOT consider the merits of pursuing joint regional freight modeling efforts with other states in the region. This is not to suggest that state or regional model are not important. Indeed, the types of trips of interest in an MPO model (distribution, parcel delivery, or garbage collection for example) are important to traffic management and air quality issues and would not be included in a regional model.

4.2 FREIGHT DATA CONSIDERATIONS

Freight transportation planning and demand models were not extensively studied (at least in the public sector) until the 1990s. Many of these freight modelling efforts have involved trying to add freight into the existing passenger demand models using the same techniques. Many times, as found here for the state of Connecticut, the freight data have been insufficient and the study area boundaries inappropriate for freight transportation. Efforts to fill major gaps in publicly available freight transportation data are needed.

In many cases, businesses and industries that rely on freight transportation have the data and resources to develop extensive freight logistic models that benefit their planning and policy decision making. No reasonable argument can be made that these businesses should have to share these proprietary data or models with the public sector for planning purposes. However, the argument can be made that public sector entities such as the departments of transportation and planning agencies must have a reasonable level of information about freight flows to properly manage the transportation infrastructure system for which they are usually responsible. Decisions on new infrastructure investment, air quality modeling, funding mechanisms such as tax structures and transportation safety require freight data that are simply not available publicly. The public sector needs freight transport models to support their role as system managers and/or coordinators between private systems. One approach to filling this freight transportation data gap might be the pursuit of a greater number of geographically located traffic classification counts for use in synthetic OD procedures as undertaken in this project. Alternatively, the following paragraphs argue for a simplified criteria for more uniform data for regional demand model development.

A major obstacle for freight planning models is the lack of consistent freight flow data. Freight information is hard to collect partly because of the confidentiality issues. However, we argue that using a larger regional scale offers the advantage of aggregating by commodity and geography such that confidentiality can be better protected than it would be within smaller areas. The existing freight data sources cannot satisfy freight analysis requirements due to deficiencies in coverage of commodity types, transportation modes, or lack of geographic details; and it is hard to merge different sources together to a common unit or to compare different data sources across regions or categories. Furthermore, international shipments are often well documented but difficult to aggregate in a form useful for transportation planning models. Public sector transportation planners have relied on multiple sources of data collected by other agencies to try and quantify the freight flow problem. These sources are not sufficient and it is timely to outline exactly what data requirements would appropriately serve planning model needs so that jurisdictions can collaborate to collect these data (relatively simple and limited in scope) across jurisdictions to be able to meet the freight modelling requirements.

The data needed to build a comprehensive demand model are relatively simple. With the extensive use of Geographic Information Systems (GIS), the multimodal network can be aggregated. The poor state of GIS databases for local or minor roads for example are not a barrier. The regional networks of interest are the best developed GIS databases. The zone system for regional models must consist of relatively large cities or regional areas which are not only relatively uniform in terms of economy, and therefore freight generation and attraction, but whose boundaries fall along jurisdictional boundaries so that demographic predictor variables such as population and employment can easily be obtained. Spatial overlay can allow these predictor variables to be combined. It is also ideal if zones share common access to the transportation network. These large geographic zones protect the proprietary nature of the shipment information and are sufficient for regional level modelling which is more relevant for policy questions such as intermodal facility placement and modal substitution.

In terms of freight shipment data, we believe commodity details need only being provided in 5-10 categories at most. Many current data collection efforts involve many more categories specifying exact commodity types that are simply not necessary for transportation planning purposes. Commodity detail needs only be disaggregated to the point where the mode of transportation can be modelled (i.e. is this a bulk good) but not so refined that collection is cumbersome or confidentiality is compromised.

The amount of a commodity shipped should be measured in volume for transportation planning purposes as it can more easily be used for conversion to vehicle units such as trucks or barges. The value or ton-miles of freight are also commonly used units but while it would be ideal to have all measures, it is necessary to simplify collection so that data can be more routinely obtained. This is even more true when one considers that it will be necessary for multiple jurisdictions to collect data for regional models. Many challenges suggest focusing on one measure such as tonnage is best. Tonnage is most appropriate for transportation planning models and less proprietary than value measures.

In addition to the total amount of commodity shipped, it is ideal to have both the origin and destination of the shipment defined in the large geographic units discussed above. Mode and route might be considered desirable quantities by some modellers. However, we argue that these factors can be modelled when origin, destination and commodity are known, if counts of vehicles along the network links are used for model calibration. Route and mode data are partially cumbersome to collect especially if the locations of intermodal transfers are desired. Coordination of collection efforts between many jurisdictions might make data collection infeasible if mode and route were also collected.

In summary, we propose that the data necessary to support a regional freight planning model consists of 1) the volume of shipped commodity (crude categories), and 2) both its origin zone and destination zone (large full city or region zones). No matter the format or scale, opportunities to collect additional data relevant to freight trip generation in the state of Connecticut and the surrounding region should be pursued whenever possible. Such data might include disaggregate employment information, manufacturing surveys, import and export documentation or classification counts. But in general, better models will be possible only if more and better data are available.

4.3 FORMULATION OF FREIGHT GENERATION MODELS

In freight modeling, social economic variables such as employment by sector and population are commonly used predictors. In general, using multiple predictors in a linear regression model would improve model performance in terms of total variance explained. However, highly correlated predictors can lead to problems such as unstable model parameters and false significance test results. Unfortunately, there is a high correlation between the population and employment by sector in some zone systems.

In related work (Guo and Aultman-Hall 2004), we have demonstrated that satisfactory models can be estimated even for commodity specific generation using social economic data, especially employment by sector. However, proper transformations are essential to generate better models. Use of spatial variables and spatial autocorrelation analysis confirms that freight generation is related to its spatial characteristics. Furthermore, use of spatial regression techniques improves the models significantly. For the commodity specific freight generation models needed to consider mode choice, the large number of zeros in zone-based freight production constitutes a major obstacle for model development. Two preliminary techniques to overcome this problem have been demonstrated: use of the employment by sector data that corresponds to the specific commodities as well as use of logistic regression (Guo and Aultman-Hall). The former resulted in the better model. Overall, we have demonstrated that even with the limited categorical data publicly available at this time, progress towards better freight generation models is possible with standard improvement to the typical linear regression model.

Therefore, while the ease of TAZ-based basic linear regression model for the Connecticut state wide traffic model supports its continued use, in order for modal substitution or intermodal planning to be considered within the model, research into more sophisticated freight generation models may be necessary.

ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of the Connecticut Department of Transportation's Bureau of Policy and Planning for providing the raw data for this analysis. Without the frequent provision of data and information by Mr. Paul Buckley of ConnDOT, this project would not have been possible. The authors also extend their thanks to Mr. David Kilpatrick and Mr. Joseph Cristalli for the provision of traffic volume data. The assistance of ConnDOT personnel and equipment for classification counts was essential. The project's technical advisory committee included Andy Asaro, Barbara Breslin, Paul Buckley, Gerald Jennings, Stu Leland, Tom Maziarz, and John Valengavich. This group was particularly active and reviewed two interim reports and directed development of new project objectives and scope as needed. The advice of Mr. James Sime of the ConnDOT Division of Research was appreciated throughout the project. Finally, advice and data from Reebie Associates is gratefully acknowledged.

REFERENCES CITED

1. Al-Deek, H. M., and G. Johnson (2000) "Truck Trip Generation Models for Seaports with Container and Trailer operation." *Transportation Research Record* 1719 Transportation Research Board, National Research Council, pp 1-9.

2. Apffel, C., J. Jayawardana, A. Ashar, K. Horn, R. McLaughlin, and A. Hochstein (1996) "Freight Components in Louisiana's Statewide Intermodal Transportation Plan," *Transportation Research Record* 1552, Transportation Research Board, National Research Council, pp 32-41.

3. Aultman-Hall, L., B. Johnson, and B. Aldridge (2000) "Assessing Potential for Modal Substitution from Statewide Freight Commodity Flow Data," *Transportation Research Record* 1719, Transportation Research Board, National Research Council, pp 10-16.

4. Cambridge Systematics Inc. (1996) *Quick Response Freight Manual*, U.S. Department of Transportation, Washington DC.

5. Cambridge Systematics Inc. (1997) *A Guidebook for Forecasting Freight Transportation Demand*, National Cooperative Highway Research Program, Report 388, Transportation Research Board, National Academy of Science, Washington DC.

6. Coifman, B., P.K. Goel, M.R. McCord, Z. Jiang, Y. Yang. (2003) "Estimating Annual Average Daily Traffic From Satellite Imagery and Air Photos," *Transportation Research Record*, p. 136-142.

7. Eatough, C., J. Brich, C. Stephen, and M. J. Demestsky (2000) "A Statewide Intermodal Freight Transportation Planning Methodology" *Journal of Transportation Research Forum*, Vol. 39, No.1, pp 145-155.

8. Glutch, R. (2002) Urban Growth Detection Using Texture Analysis on Merged Landsat TM and SPOT-T Data. *Photogrammetric Engineering and Remote Sensing*, 2002, p. 6.

9. *Guidebook on Statewide Travel Forecasting.* FHWA, U.S. Department of Transportation, 1999.

10. Feng Guo and Lisa Aultman-Hall, "Alternative Nationwide Freight Generation Models" submitted to the Transportation Research Board, July 2004.

11. Huang, W. J. and R. L. JR. Smith (1999) "Using Commodity Flow Survey Data to Develop a Truck Travel-Demand Model for Wisconsin" *Transportation Research Record* 1685, Transportation Research Board, National Research Council, pp1-6.

12. Krishna, V, and K. Hancock (1998) “Highway Freight Flow Assignment in Massachusetts Using Geographic Information Systems.” *Transportation Research Record* 1625, Transportation Research Board, National Research Council, pp. 156-164.

13. Scott, M. D. (2003) “Assessment of the 1997 Commodity Flow Survey for State-Level Freight Transportation Planning,” *Report TTC-03-11/MSC-97-1F*, Kentucky Transportation Center.

14. Souleyrette, R., T. H. Maze, T. Strauss, D. Preissig, and A. Smadi, (1998) “Freight Planning Typology” *Transportation Research Record* 1613 Transportation Research Board, National Research Council, pp 12-19.

15. Sorratini, J. A. and R. L. Jr. Smith (2000) “Development of a Statewide Truck Trip Forecasting Model Based on Commodity Flows and Input-Output Coefficients.” In *Transportation Research Record* 1707, Transportation Research Board, National Research Council, pp. 49-56.

Appendices

Appendix A: Regional Planning Organizations Contacts for Additional Truck Volume Counts

This task was undertaken during a three-day period from November 17th, 2003 to November 19th, 2003.

	CT Planning Region	tel. #	Avail of info. requested
CCRPA	Central Connecticut Regional Planning Agency	860-589-7820	additional contact via email -no real info
COGCNV	Council of Governments of the Central Naugatuck Valley	203-757-0535	ConnDOT data only, counters available
CRCOG	Capitol Region Council of Governments	860-522-2217	no additional information, referred to ConnDOT
CRERPA	Connecticut River Estuary Regional Planning Agency	860-388-3497	no additional information suggested to Use CTI equipment
GBRPA	Greater Bridgeport Regional Planning Agency	203-366-5405	(Mike) no additional information
HVCEO	Housatonic Valley Council of Elected Officials	203-775-6256	Andy- (two days phone tag) no additional information
LHCEO	Litchfield Hills Council of Elected Officials	860-491-9884	no additional information, referred to ConnDOT
MRPA	Midstate Regional Planning Agency	860-347-7214	by vol. Traffic counts on local and state rd. (has the equip.)
NECOG	Northeastern Connecticut Council of Governments	860-774-1253	No additional information, referred to ConnDOT, also point out that needs ConnDOT permission to perform any work on state roads
NWCOG	Northwestern Connecticut Council of Governments	860-868-7341	Dan McGuinness, (phone tag), rely on ConnDOT data
SCRCOG	South Central Regional Council of Governments	203-234-7555	No additional information, refered to ConnDOT
SECOG	Southeastern Connecticut Council of Governments	860-889-2324	Richard, only information from ConnDOT, local roads only by volume. (In Disagreement with ConnDOT-old Estimates vs new conditions)
SWRPA	South Western Regional Planning Agency	203-316-5190	Mr. Suposi, no additional information
VCOG	Valley Council of Governments	203-735-8688	Talked with Kevin Omera from accident stats you can determine the classify. Traffic counts? Received fax- no additional information for this study use
WRCG	Windham Region Council of Governments	860-456-2221	no additional information, referred to ConnDOT

Appendix B: Additional Truck Volume Counts

Summary:

63	Sites requested
49	Sites where data were gathered in both directions
5	Sites where data were gathered in one direction only
3	Sites that had bad data in both directions
6	Sites that were not collected due to time constraints

	TOWN	LATITUDE	LONGITUDE	ROUTE	DIR	ADT	SINGLE UNIT TRUCKS	SINGLE-TRAILER TRUCKS	24 HR COUNT
1	BURLINGTON	41.77	72.97	4	EAST	6900	246	91	6657
					WEST	6800	239	139	6549
2	PLYMOUTH	41.66	73.05	262	NORTH	1200	27	2	1215
					SOUTH	1100	28	2	1082
3	BETHLEHEM	41.62	73.24	132	EAST	600	12	1	571
					WEST	550	10	2	539
4	OXFORD	41.44	73.13	67	EAST	6600	166	86	6405
					WEST	6300	148	89	6158
5	WOLCOTT	41.59	72.95	322	NORTH	2900	62	44	2921
					SOUTH	3000	50	27	2963
6	WOODBURY	41.55	73.21	47	NORTH	2400	60	17	2328
					SOUTH	2600	54	11	2520
7	ANDOVER	41.74	72.37	316	NORTH	1300	32	2	1318
					SOUTH	1200	25	2	1235
8	BLOOMFIELD	41.90	72.76	189	NORTH	6800	110	18	6823
					SOUTH	6700	87	77	6688
9	BOLTON	41.78	72.46	85	NORTH	3700	52	10	3776
					SOUTH	3400	30	21	3386
10	CANTON	41.90	72.89	179	NORTH	1500	34	6	1480
					SOUTH	1500	40	9	1432
11	EAST GRANBY	41.92	72.76	189	NORTH	TUBES UP	X	X	X
					SOUTH	5400	47	35	5269
12	ENFIELD	41.97	72.59	5	NORTH	6600	104	71	6641
					SOUTH	6800	101	85	6785
13	HEBRON	41.72	72.44	85	NORTH	2500	19	12	2527
					SOUTH	2400	42	19	2465
14	SIMSBURY	41.87	72.81	167	EAST	6400	112	108	6353
					WEST	6500	113	104	6477
15	SOMERS	41.99	72.45	167	EAST	6900	136	202	6691
					WEST	6900	139	112	6744
16	SUFFIELD	41.98	72.64	513	EAST	4600	64	9	4629
					WEST	4900	55	14	4886
17	NO DATA	X	X	X	X	X	X	X	X

18	WEST HARTFORD	41.80	72.72	185	NORTH	4800	51	12	4766
					SOUTH	4900	48	14	4899
19	DEEP RIVER	41.37	72.46	80	EAST	2700	74	37	2608
					WEST	2700	62	25	2577
20	ESSEX	41.36	72.42	154	NORTH	3300	61	43	3172
					SOUTH	3600	95	30	3462
21	KILLINGWORTH	41.36	72.56	81	NORTH	4700	84	26	4585
					SOUTH	4800	109	16	4700
22	OLD LYME	41.31	72.34	156	NORTH	3400	39	13	3342
					SOUTH	3400	67	12	3256
23	WESTBROOK	41.27	72.48	1	EAST	3800	50	17	3719
					WEST	4100	29	26	3964
24	EASTON	41.23	73.32	136	NORTH	3100	25	7	2974
					SOUTH	3300	34	10	3247
25	MONROE	41.33	73.21	111	NORTH	7200	103	89	7228
					SOUTH	7100	115	45	7084
26	BRIDGEWATER	41.50	73.36	133	NORTH	2300	15	7	2263
					SOUTH	2300	17	4	2263
27	NEWTOWN	41.41	73.31	25	NORTH	TUBES UP	X	X	X
					SOUTH	7200	137	125	7150
28	SHERMAN	41.58	73.50	39	NORTH	1800	36	11	1701
					SOUTH	1900	34	9	1803
29	BARKHAMSTED	41.92	72.94	219	NORTH	3000	62	44	2921
					SOUTH	3100	50	27	2963
30	COLEBROOK	41.97	73.05	8	NORTH	1300	36	155	1239
					SOUTH	1200	38	119	1179
31	HARWINTON	41.77	73.06	4	EAST	6200	75	79	6037
					WEST	6200	67	57	5977
32	LITCHFIELD	41.75	73.16	118	EAST	3300	73	29	3199
					WEST	3300	71	26	3226
33	MORRIS	41.68	73.20	109	EAST	2000	37	12	1949
					WEST	2000	40	10	1955
34	NEW HARTFORD	41.83	72.98	202	EAST	5300	68	54	5105
					WEST	5300	68	46	5099
35	TORRINGTON	41.82	73.08	183	NORTH	4400	129	63	4442
					SOUTH	4600	43	64	4598
36	EAST HADDAM	41.43	72.41	82	NORTH	1200	38	15	1191
					SOUTH	1200	38	13	1174
37	PORTLAND	41.60	72.60	17	NORTH	3400	124	13	3269
					SOUTH	3600	116	31	3496
38	NO DATA	X	X	X	X	X	X	X	X
39	NO DATA	X	X	X	X	X	X	X	X

40	POMFRET	41.86	71.96	169	NORTH	2000	47	16	2042
					SOUTH	1900	45	11	1957
41	NO DATA	X	X	X	X	X	X	X	X
42	STERLING	41.70	71.80	14	EAST	550	17	9	565
					WEST	600	21	12	581
43	NO DATA	X	X	X	X	X	X	X	X
44	UNION	41.98	72.18	190	EAST	950	15	13	964
					WEST	950	26	16	939
45	STAFFORD	41.98	72.29	319	NORTH	TUBES UP	X	X	X
					SOUTH	1900	26	6	1939
46	CANAAN	41.93	73.28	63	NORTH	1200	35	30	1155
					SOUTH	1200	33	28	1178
47	ROXBURY	41.56	73.31	317	EAST	1200	19	4	1154
					WEST	1100	13	3	1089
48	EAST HAVEN	41.29	72.87	100	NORTH	6800	193	55	6837
					SOUTH	6700	175	118	6724
49	GRISWOLD	41.58	71.90	201	NORTH	TUBES UP	X	X	X
					SOUTH	800	19	1	796
50	GROTON	41.35	72.04	1	EAST	7800	126	116	7816
					WEST	7900	128	129	7903
51	NO DATA	X	X	X	X	X	X	X	X
52	NO DATA	X	X	X	X	X	X	X	X
53	NO DATA	X	X	X	X	X	X	X	X
54	NEW CANAAN	41.13	73.46	123	NORTH	8800	191	156	8759
					SOUTH	9300	202	57	9252
55	STAMFORD	41.13	73.54	137	NORTH	6400	92	30	6415
					SOUTH	6100	76	151	6081
56	WESTON	41.21	73.39	53	NORTH	4000	75	12	3891
					SOUTH	4300	97	8	4167
57	ANSONIA	41.35	73.09	334	NORTH	3300	59	52	3271
					SOUTH	3000	42	41	2994
58	NO DATA	X	X	X	X	X	X	X	X
59	CHAPLIN	41.75	72.15	6	EAST	7200	145	212	7233
					WEST	6300	153	176	6354
60	COLUMBIA	41.70	72.30	87	NORTH	1900	38	26	1947
					SOUTH	1900	28	16	1887
61	COVENTRY	41.80	72.37	44	EAST	5000	73	46	5034
					WEST	4800	64	57	4872

62	HAMPTON	41.77	72.06	97	NORTH	400	17	0	405
					SOUTH	450	21	1	452
63	WILLINGTON	41.87	72.26	74	EAST	TUBES UP	X	X	X
					WEST	3300	97	113	3312

Appendix C: Classification Count Procedures

LOCATION SELECTION PROCESS

To select the locations to be counted, towns were first selected that had no previous counts. Selection preference went to towns that had adjacent towns with no previous counts as well. Once the towns were selected, major 2-lane routes in the North-South or East-West direction were looked at and selected. 4-lane routes were excluded for safety. Maps were prepared to direct counters.

EQUIPMENT USED

- 12 Pegasus vehicle traffic counters
- 12 chains and locks
- 24 O-shaped tubes at 30' length
- 2-1/2" MAG nails (16 per location)
- Metal clamps
- 3-lb hammer
- Large crowbar
- Safety tools – knee pads, safety vests and hats, cones, etc.

SETUP

- Traveled to approximate location on map
- First looked for a location with good driver line-of-sight for safety
- Next needed to find something near road to chain machines e.g. sign, tree, pole, guard rail
- Parked truck on shoulder, but very close to (sometimes in) roadway to slow down, but not stop, traffic
- Placed cones along center line to warn drivers to slow down
- Set up counter by naming 8-digit file based on direction, town number, urban/rural area, and Pat's ID number
- Placed two counters, one in each direction, with each counter containing two tubes at 12' spacing
- Confirmed that machines were properly classifying vehicles (real-time)
- Obtained GPS coordinates of site location
- Setup took between 20-45 minutes

PICKUP

- 24 hours later we could pick up the machines and tubes
- Pickup took between 10-15 minutes
- Immediately downloaded data from counter to a laptop computer using TrafMan v. 6.06 software

- Confirmed that the machines classified vehicles for 24 hours, viewing in hourly classification totals
- This download process creates a BIN file for each direction

POST-PROCESS

- Data process begins by converting BIN files into CLA (classification) files and creating one DAT file
- Using computers in ConnDOT's Planning section, the DAT file was FTP'd to the ConnDOT mainframe
- DOT mainframe prints a single page – called a traffic recorder data sheet - for each direction of each location
- Using a DOS program called *timclass*, the CLA files are used to print 3 pages – hourly vehicle classification data, including a summary – for both directions of each location

Data from these printed pages, along with site longitudes and latitudes, were then entered into an Excel Worksheet and submitted for entry into a database

MEETINGS WITH DOT STAFF ON THE FOLLOWING DATES: February 17th and February 23rd 2004

OFFICE TRAINING: March 1st -2nd 2004

FIELD TRAINING: March 3rd through March 11th 2004

OBTAINED COUNTING EQUIPMENT FROM DOT: March 12th 2004

During the week of March 8th, 2004, traffic counters were installed and recovered in the following towns:

Simsbury
Colebrook
New Hartford
Barkhamstead
Torrington
Plymouth
Wolcott

Week of March 15th:

Harwinton
Litchfield
Morris
Bethlehem

Week of March 22nd:

Canaan
Sherman
Roxbury
Woodbury
Monroe
Ansonia
Newtown
Bridgewater
Oxford
Stamford
New Canaan
Weston
Easton
West Hartford
Bloomfield
East Granby
Suffield

Enfield
Somers

Week of March 29th:

Killingworth
Westbrook
Old Lyme
Groton
Portland
East Haven
Essex
Deep River
East Haddam

Week of April 5th:

Ledyard
Griswold
Sterling
Hampton
Ashford
Pomfret

Week of April 12th:

Columbia
Hebron
Coventry
Willington
Union
Chaplin
Andover
Bolton
Stafford
Brooklyn

OFFICE POSTPROCESSING: April 1st, April 16th and April 28th 2004

EQUIPMENT RETURNED: April 15th 2004

Appendix D:
Trucks/Freight in State-Wide Planning Traffic Models – Phone Survey Fall 2002

Hello my name is Lisa Aultman-Hall and I am in the Civil Engineering Department at the University of Connecticut. I am working on a research project for the Connecticut DOT that includes consideration of how freight or trucks should be incorporated into our state-wide traffic model. I got your name from a list of participants at a state-wide travel demand forecasting workshop held in Irvine CA in 1998. Are you the most appropriate person to discuss the truck modeling in the statewide forecasting model or is there someone else in your state you would recommend I talk to.

Do you have a state-wide traffic model that is used for planning and forecasting? How many zones? What class of roads are included in the network?

What software package is your model in?

Are trucks included in your model?

Are any other modes of freight transportation included?

How is intermodal planning policy evaluation conducted?

Is the zone structure the same for trucks?

Is the network the same?

Are impedance factors for traffic assignment and trip distribution the same for trucks?

What data sources were used to create your truck models – the trip generation models in particular? What year was the data collected?

Have you considered using the data available from Reebie Associates (Transearch)?

Have you considered using the Commodity Flow Survey data?

Appendix E: Summary of Phone Survey Results

Contact names for the personnel responsible for state-wide modeling in 25 states were obtained from ConnDOT on a participants' list from an 1998 conference on travel demand forecasting held in Irvine, California. The intention was to contact states and ask what efforts, if any, they had undertaken to include freight or truck trips in their state wide traffic model. Many of these contacts were no longer with the agencies or more appropriate individuals were recommended for the particular combination of modeling and freight planning questions of interest. A total of 16 of these states have been interviewed by telephone about their procedures for including trucks and/or freight in their state-wide planning models. Eight of these 16 states do not have a state-wide traffic model for planning purposes, while four had a traffic model but did not have trucks or freight incorporated into that model. In some of these states, certain MPOs have models even when the state does not and some MPOs had included truck trips in their model. The state-wide planning efforts for freight in these states without models or truck models usually involves forecasts of future truck demand by road class based on growth rates applied to classification counts.

Two of the states contacted, Kentucky and Virginia, had four-step planning models that included truck components based on Reebie Associates data. The same zonal structure and network is used for passenger and freight components of the model. In both cases, some calibration against traffic classification counts had been conducted or was planned.

Two states, Ohio and Oregon, had very comprehensive and unique models. In Ohio, special roadside surveys as well as Reebie data were being used to create a commodity-based model that included a microsimulation of the economy by zone using input/output models. Oregon's model is also economy-based and generates freight by dollar value for conversion to trucks. Intercept surveys were performed on highways to provide data for model development. In the Oregon case, land use is also being modeled at the same time as the transportation demand that results from the land use. Their weight distance tax also provides an opportunity for additional data collection at special weigh stations that are moved throughout the state. Both these modeling efforts are beyond what would fit within the existing model being used at ConnDOT.

Some general conclusions can be drawn from the telephone surveys. First, comprehensive modeling of freight or truck flows using a four stage planning process at the state

-wide level is by no means common. Many states are simply not using this approach for freight or intermodal planning. Some states are making use of the comprehensive Reebie Associates data but few have made any large effort to validate the data with anything beyond limited classification volume counts where available. This lack of validation is for the most part because no reasonable means to collect freight data is available to the states. The Commodity Flow Survey data is not being used in general by states for planning models. While some states have gone to tremendous effort to collect freight data and to development new models beyond the four stage framework; this is rare.

Appendix F: Comparison of FAF to Classification Counts and ConnDOT Planning Model (1996)

Procedure:

The Freight Analysis Framework (FAF) is a national level freight planning model developed by FHWA. The major data source is the county to county level Reebie Associates data. Two kinds of truck link flows are provided publicly for the nationwide truck highway network: FAF truck volume, which is the truck volume from the planning model; and total truck volume, which is the observed truck link volume (HPMS) as provided by state DOTs. Using the lookup tables described in Part IV of this summary the FAF truck volumes were compared to both the classification counts and the planning model output. There were a total of 40 locations where counts from 1998 matched the FAF network links. Note that truck counts were aggregated because FAF provides only total two-directional volume.

Comments:

The average absolute difference between the FAF total truck volumes and the classification counts obtained for this study is 2251 trucks per day (full data shown in Table F-1). Overall the FAF volumes are much higher (the average difference is 1990 trucks per day). Given that the source of data for the total trucks in the FAF dataset is the HPMS, this difference is not expected.

The FAF and planning model volumes overlap over the entire length of routes within the FAF. Therefore the comparison shown in Table F-2 is averaged in both directions for full routes. The first two columns of the table illustrate again that the FAF or commodity based trucks are very low compared to the model, suggesting commodity data is not complete enough for CT use for forecasting models. The absolute difference between the total FAF truck volumes and the model output is better but the pattern in columns 3 and 4 of the table suggest that model is underestimating trucks on the interstates while overestimating on the non-interstate routes. This could be addressed by changing the parameters in the assignment model in the software. These conclusions are supported by the patterns illustrated in the maps of Figures F-1 and F-2.

Table F-1 Comparison of Truck Classification Count Data and FAF

Truck classification data (1998)				FAF 1998		Difference (Total Truck-AADTT98
RteNo	Mile Post	SU Trucks (per day)	Trailer Trucks (per day)	Total Trucks 1998	FAF Trucks (per day) 1998	
2	6	956	984	2550	281	-610
2	38	263	157	2070	742	-1650
6	0	83	85	9110	2576	-8942
6	15	282	490	10050	2480	-9278
6	80	599	570	660	198	509
6	90	502	442	8820	151	-7876
6	93	690	504	8820	151	-7626
6	107	285	335	8820	3879	-8200
7	14	625	256	1180	84	-299
7	18	602	583	1180	84	5
7	26	802	979	1000	84	781
7	26	796	964	1000	84	760
8	9	1461	1201	3290	197	-628
8	29	1642	1432	3290	310	-216
8	46	520	812	2370	310	-1038
8	47	628	843	2370	310	-900
10	20	483	663	1190	26	-44
10	41	254	213	370	0	97
10	47	241	123	540	2	-176
10	48	253	140	540	2	-147
10	48	190	105	400	2	-105
15	73	456	366	3510	588	-2689
25	10	709	501	2050	65	-840
32	0	193	90	590	0	-307
66	3	442	516	800	712	158
84	0	1806	8351	9110	2576	1047
84	34	1842	7830	10050	2414	-378
84	52	2621	6834	8820	2397	635
84	93	986	6631	8820	151	-1203
91	52	2750	6637	11130	6035	-1743
95	1	3587	13684	16560	4679	711
95	2	3486	13584	16560	4679	510
95	55	1519	6847	15500	5404	-7134
95	56	1563	7010	15500	5404	-6927
95	62	1354	6293	15500	5404	-7853
95	63	1410	6252	14440	5404	-6778
384	8	703	770	1470	88	3
395	9	1356	2617	3990	370	-17
395	35	410	1539	2960	941	-1011
695	1	88	240	560	343	-232

Table F-2 Comparison of FAF and ConnDOT Model Link Volume by Route

HWY	Difference Between FAF truck and ConnDOT Model	Percentage of FAF Truck to ConnDOT Model	Difference Between FAF Total Truck and ConnDOT Model	Percentage of FAF Total truck to ConnDOT Model
I291	-2528	5%	-173	107%
I384	-1809	5%	-427	83%
I395	-1124	51%	1453	263%
I691	-831	153%	572	238%
I84	-4324	88%	2348	339%
I91	-4083	82%	3978	274%
I95	-4451	46%	4002	202%
CT10	-1376	3%	-661	64%
CT11	-523	36%	-216	112%
CT12	-1484	6%	-940	47%
CT15	535		3956	
CT17	-2858	8%	-2095	33%
CT2	-1962	13%	-691	81%
CT20	-3087	0%	-334	89%
CT25	-1400	6%	-66	114%
CT3	-2109	13%	-945	65%
CT32	-1467	0%	-851	52%
CT34	-2264	7%	-965	62%
CT40	-3047	7%	-2042	57%
CT66	-570	48%	-355	73%
CT695	175	204%	399	337%
CT72	-1748	6%	-879	54%
CT78	56		570	
CT8	-3731	8%	-1494	77%
CT82	-228	39%	-228	39%
CT85	-287	59%	110	203%
CT9	-3098	11%	-808	85%
US1	76	193%	207	341%
US202	-852	15%	-618	47%
US44	-1127	5%	-510	83%
US5	-1160	4%	-151	110%
US6	-835	18%	-243	94%
US7	-2354	17%	-1162	364%
Ave	-1693	37%	22.5	135%

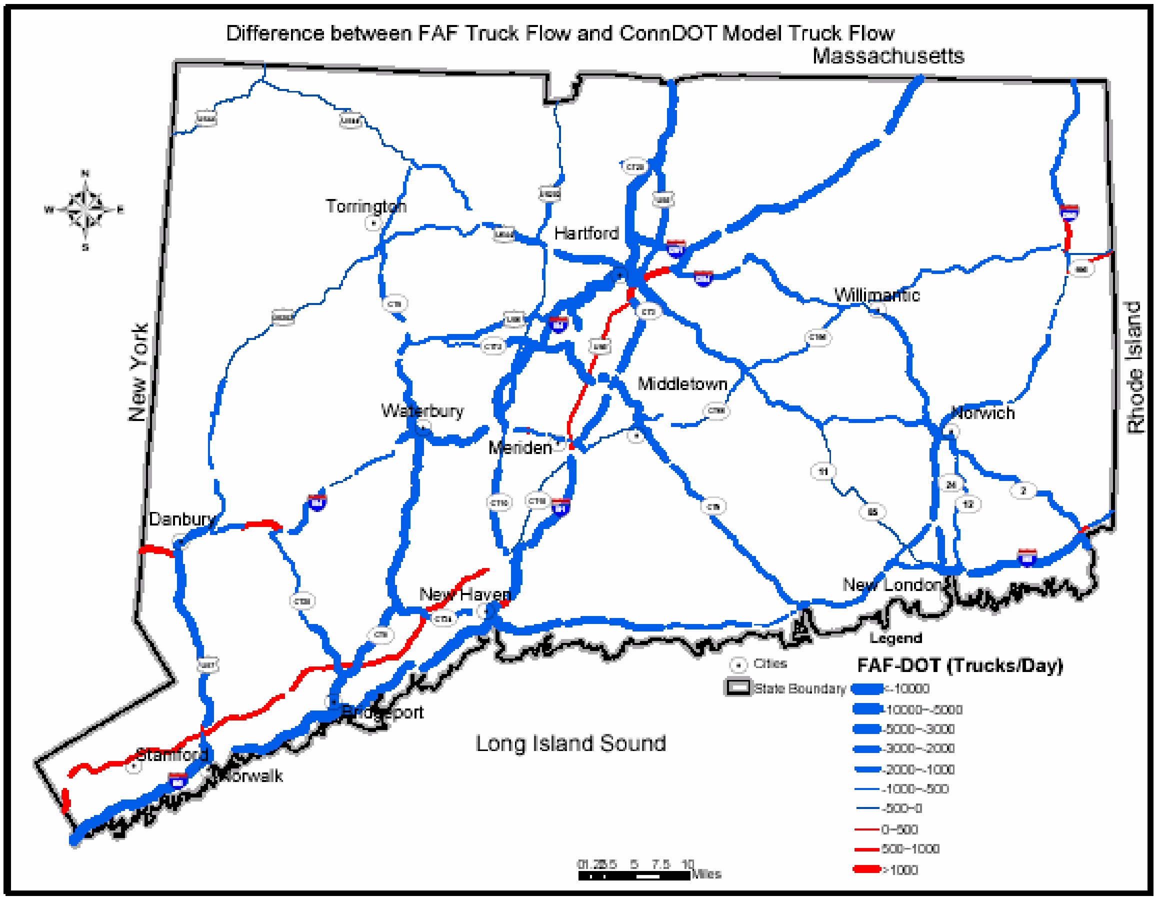


Figure F-1 Difference between FAF trucks and ConnDOT Model Trucks

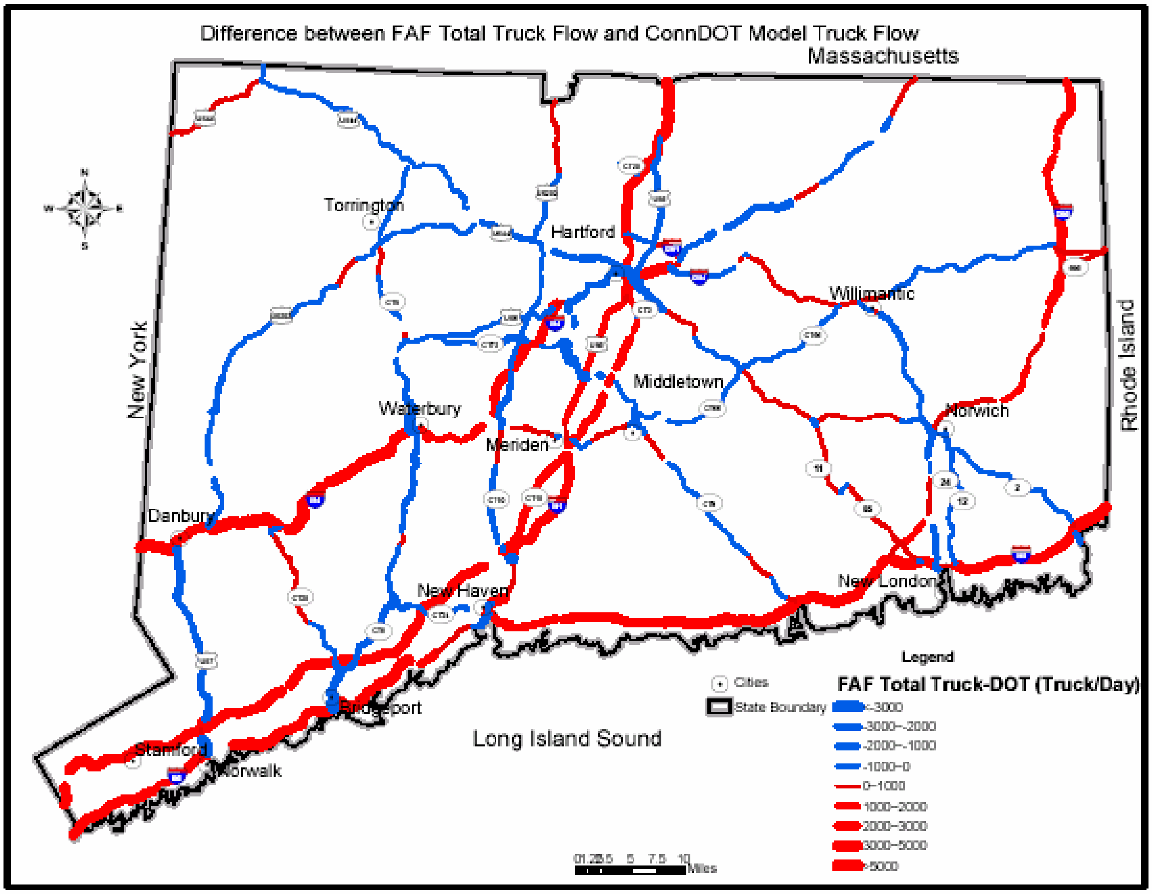


Figure F-2 Difference between FAF AADTT and ConnDOT Model Trucks

Appendix G: Comparison of ConnDOT Model and Existing Truck Classification Count Data

Procedure:

Using GIS lookup tables, comparisons were undertaken between the 1996 truck classification count volumes and the truck volumes predicted by the planning model. These results are tabulated in Table G-1. In many cases the differences are substantial. Figure G-1 illustrates the location of difference. Figure G-2 and G-3 illustrate histograms of the differences.

Comments:

Overall, the volumes from the model are higher than the counts, suggesting the model is overestimating trucks. This finding is consistent with results from the Reebie data and CFS data. This may be due to a difference in classifying trucks (the counts only include trucks with six tires or more). Many of the large outliers occur at locations on or near freeways where the volumes are high and the percentage difference is not great. However, many of the locations with high percentage differences are located on lower volume highways, where the actual trucks counts are very low and the percentage difference is therefore great when any significant number of trucks is output by the model. The large errors make either absolute and percentage different measures inappropriate in some cases and thus both are presented here. The higher assigned truck volume in the model on low volume roads is sometimes due to the location of a zonal connector. In short, most of the modeled truck volumes are more than 100 trucks per day different than the actual counts and just over half of the volumes are more than 100% different. This suggests that improvements could be made to the truck component of the state wide model. However, consideration of the commodity data in the project does not suggest its use would improve the model.

Table G-3 Comparison of the Link Volume of Classification Count and ConnDOT Truck Model

Truck Classification Data						CONNDOT MODEL Trucks per day	DIFFER*	PERCENT DIFF**
Route Number	DIR	MILEPOST	SU TRUCKS (per day)	TRAILERS TRUCK (per day)	TOTAL TRUCKS (per day)			
1	E	0.88	13	4	17	704	687	4041
1	W	0.88	14	5	19	585	566	2979
1	E	93.43	64	100	164	98	-66	-40
1	W	93.43	35	52	87	100	13	15
2	E	6.19	365	557	922	1768	846	92
2	W	6.19	360	693	1053	2884	1831	174
2	E	42.95	185	50	235	604	369	157
2	W	42.95	181	68	249	886	637	256
2	W	50.34	202	138	340	931	591	174
2	E	50.34	179	110	289	649	360	125
5	N	46.5	300	278	578	160	-418	-72
5	S	46.5	474	280	754	158	-596	-79
6	E	38.67	76	58	134	449	315	235
6	W	38.67	34	15	49	463	414	845
6	E	43.67	137	139	276	273	-3	-1
6	W	43.67	153	94	247	278	31	13
6	E	43.67	170	216	386	273	-113	-29
6	W	43.67	205	173	378	278	-100	-26
6	E	92.41	318	228	546	5117	4571	837
6	W	92.41	220	255	475	4965	4490	945
6	E	79.25	207	239	446	683	237	53
6	W	79.25	334	269	603	629	26	4
7	N	51.67	32	21	53	1016	963	1817
7	S	51.67	27	23	50	1165	1115	2230
8	N	66.19	29	166	195	627	432	222
8	S	66.19	29	129	158	802	644	408
8	N	19.19	629	732	1361	2240	879	65
8	S	19.19	581	736	1317	2431	1114	85
9	N	27.37	525	927	1452	1201	-251	-17
9	S	27.37	581	577	1158	961	-197	-17
9	N	16.66	274	344	618	1158	540	87
9	N	16.66	274	344	618	1158	540	87
10	N	47.86	104	82	186	612	426	229

* DIFFER=CONNDOT MODEL TRUCK-TOTAL TRUCK

** PERCENT DIFF=(DIFFER/TOTAL TRUCK)*100

Truck Classification Data						CONNDOT MODEL Trucks per day	DIFFER*	PERCENT DIFF**
Route Number	DIR	MILEPOST	SU TRUCKS (per day)	TRAILERS TRUCK (per day)	TOTAL TRUCKS (per day)			
10	S	47.86	117	59	176	376	200	114
10	N	23.02	47	42	89	944	855	961
10	S	23.02	48	30	78	1012	934	1197
10	N	12.17	259	211	470	803	333	71
10	S	12.17	316	201	517	830	313	61
12	N	6.9	109	95	204	355	151	74
12	S	6.9	127	129	256	369	113	44
12	N	4.21	95	34	129	744	615	477
12	S	4.21	85	38	123	755	632	514
12	N	35.97	104	45	149	29	-120	-81
12	S	35.97	128	43	171	39	-132	-77
30	N	0.5	57	16	73	308	235	322
32	N	16.5	164	151	315	227	-88	-28
32	S	16.5	104	181	285	233	-52	-18
32	S	47.09	80	91	171	172	1	1
32	N	47.09	91	72	163	212	49	30
44	E	38.55	124	79	203	105	-98	-48
44	W	38.55	120	70	190	61	-129	-68
45	N	3.77	27	9	36	40	4	11
45	S	3.77	26	8	34	48	14	41
53	N	10.61	23	10	33	65	32	97
53	S	10.61	26	8	34	78	44	129
61	N	0.08	44	8	52	40	-12	-23
61	S	0.08	38	22	60	57	-3	-5
63	S	25.77	39	15	54	188	134	248
63	N	25.77	46	17	63	168	105	167
63	N	13.96	162	213	375	91	-284	-76
63	S	13.96	102	138	240	103	-137	-57
67	W	3.56	43	28	71	194	123	173
67	E	3.56	52	27	79	264	185	234
67	E	27.54	54	76	130	285	155	119
67	W	27.54	66	116	182	209	27	15
69	N	28.22	35	26	61	744	683	1120
69	S	28.22	34	29	63	703	640	1016
69	N	17.66	55	80	135	332	197	146
69	S	17.66	49	85	134	402	268	200
69	N	19	133	63	196	332	136	69
69	S	19	138	65	203	402	199	98
71	N	10.69	22	13	35	259	224	640
71	S	10.69	10	21	31	252	221	713

Truck Classification Data						CONNDOT MODEL Trucks per day	DIFFER*	PERCENT DIFF**
Route Number	DIR	MILEPOST	SU TRUCKS (per day)	TRAILERS TRUCK (per day)	TOTAL TRUCKS (per day)			
71	N	12.38	50	34	84	257	173	206
ts71	S	12.38	71	53	124	252	128	103
72	N	7	134	64	198	1711	1513	764
72	S	7	171	74	245	2544	2299	938
82	E	25	86	26	112	29	-83	-74
82	W	25	93	21	114	40	-74	-65
83	N	16.05	167	73	240	721	481	200
83	S	16.05	130	76	206	787	581	282
83	N	24.48	19	18	37	343	306	827
83	S	24.48	22	15	37	331	294	795
83	N	20.92	95	45	140	396	256	183
83	S	20.92	72	54	126	383	257	204
84	E	94.5	465	3551	4016	5465	1449	36
84	W	97.35	576	3649	4225	6085	1860	44
85	N	23.75	88	10	98	40	-58	-59
85	S	23.75	70	9	79	33	-46	-58
87	N	4.05	74	9	83	149	66	80
87	S	4.05	7	1	8	113	105	1313
95	N	110	340	2630	2970	1302	-1668	-56
95	S	110	435	2278	2713	1633	-1080	-40
106	N	12.17	70	67	137	266	129	94
106	S	12.17	72	99	171	264	93	54
124	N	0	53	31	84	228	144	171
124	S	0	52	39	91	241	150	165
136	S	5.94	73	40	113	117	4	4
136	N	5.94	64	65	129	117	-12	-9
151	N	3.04	22	11	33	27	-6	-18
151	S	3.04	16	8	24	27	3	13
156	E	4.41	27	4	31	128	97	313
156	W	4.41	30	6	36	131	95	264
160	E	5.55	4	1	5	300	295	5900
160	W	5.55	7	0	7	299	292	4171
165	E	1.3	146	32	178	166	-12	-7
165	E	10.81	35	16	51	91	40	78
165	W	10.81	20	28	48	110	62	129
202	N	35.05	68	43	111	211	100	90
202	S	35.05	58	31	89	240	151	170
202	N	31.29	65	31	96	573	477	497
202	S	31.29	63	38	101	656	555	550
202	N	50.81	98	56	154	214	60	39

Truck Classification Data						CONNDOT MODEL Trucks per day	DIFFER*	PERCENT DIFF**
Route Number	DIR	MILEPOST	SU TRUCKS (per day)	TRAILERS TRUCK (per day)	TOTAL TRUCKS (per day)			
202	S	50.81	88	42	130	199	69	53
202	N	14	292	710	1002	361	-641	-64
202	S	14	217	456	673	392	-281	-42
286	N	0	70	14	84	278	194	231
286	S	0	74	12	86	292	206	240
341	E	4.5	23	6	29	76	47	162
341	W	4.5	24	6	30	46	16	53
395	N	33.51	428	1147	1575	797	-778	-49
395	S	33.51	460	1028	1488	996	-492	-33
395	N	53.78	171	861	1032	676	-356	-34
395	S	53.78	161	863	1024	761	-263	-26
641	N	1.02	45	88	133	1012	879	661
641	S	1.02	107	72	179	1084	905	506
660	N	1.64	19	49	68	41	-27	-40
660	S	1.64	17	65	82	42	-40	-49
691	W	7.52	436	931	1367	1675	308	23
691	E	7.52	443	888	1331	1839	508	38

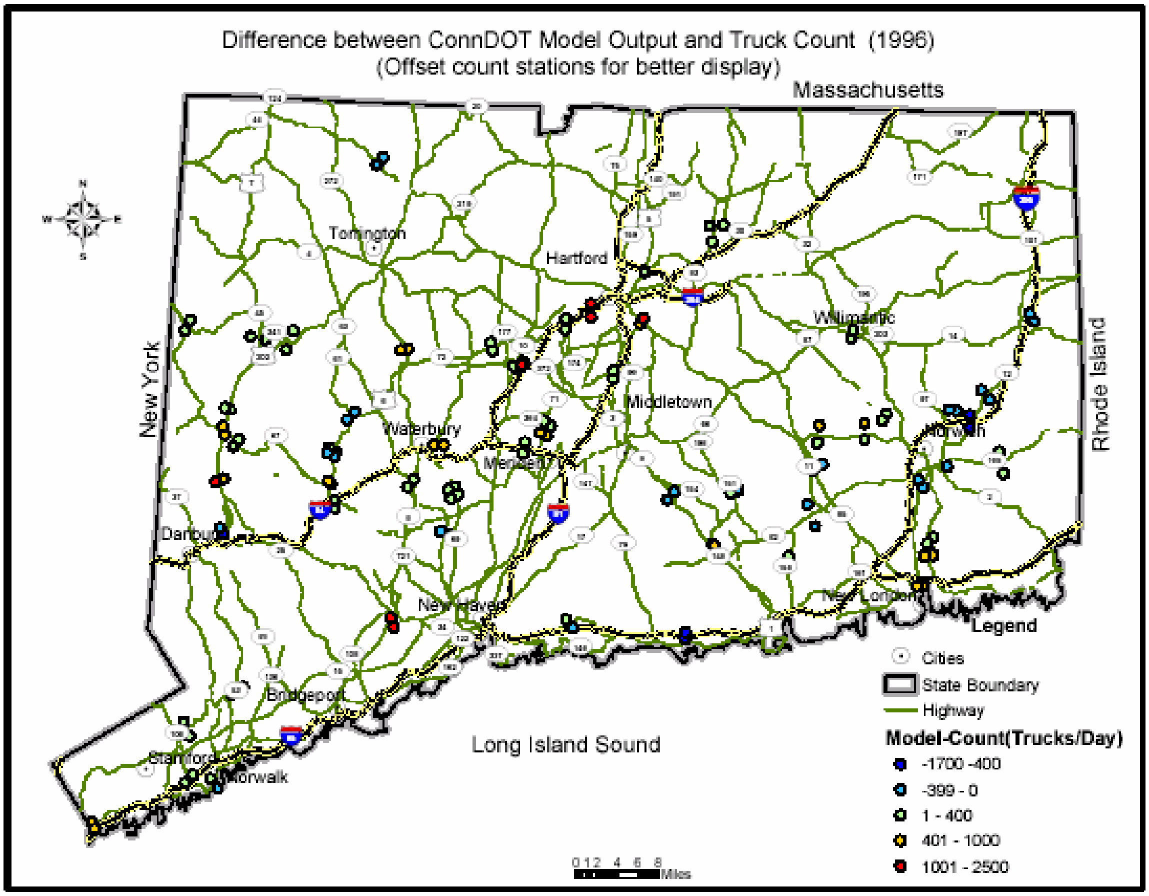


Figure G-1 Difference between ConnDOT Model Output and Truck Count (1996)

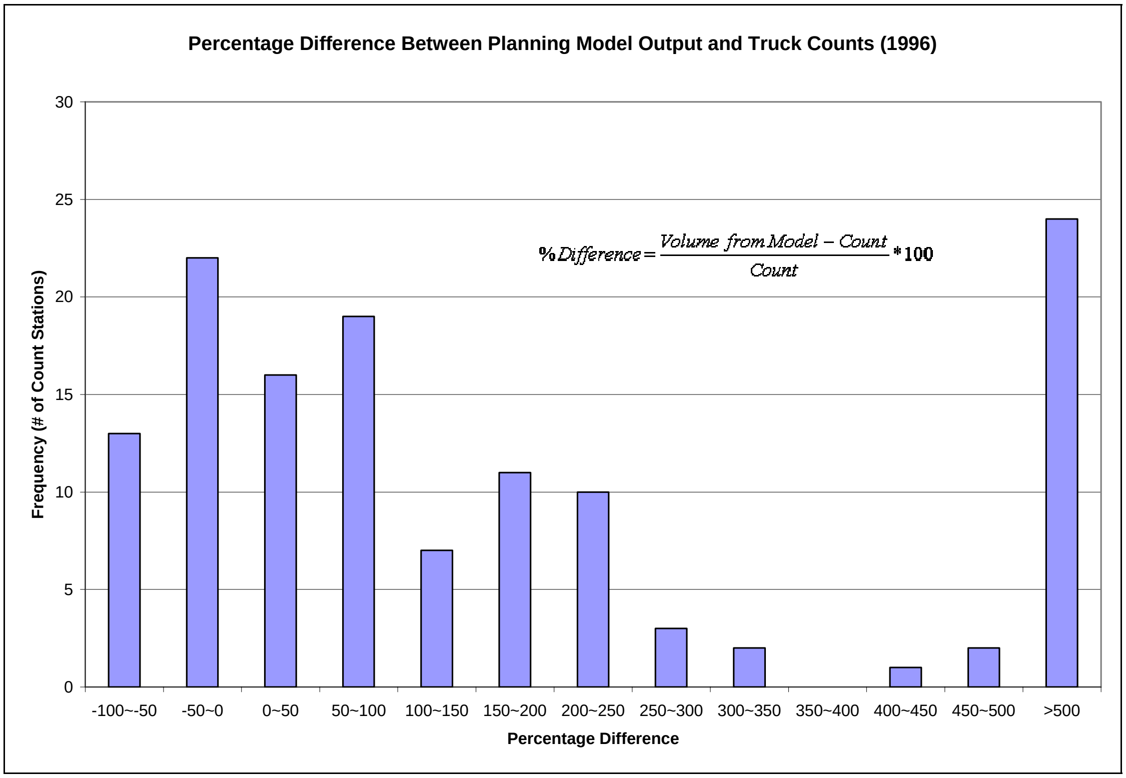


Figure G-2 Percentage Difference Between Planning Model Output and Truck Counts (1996)

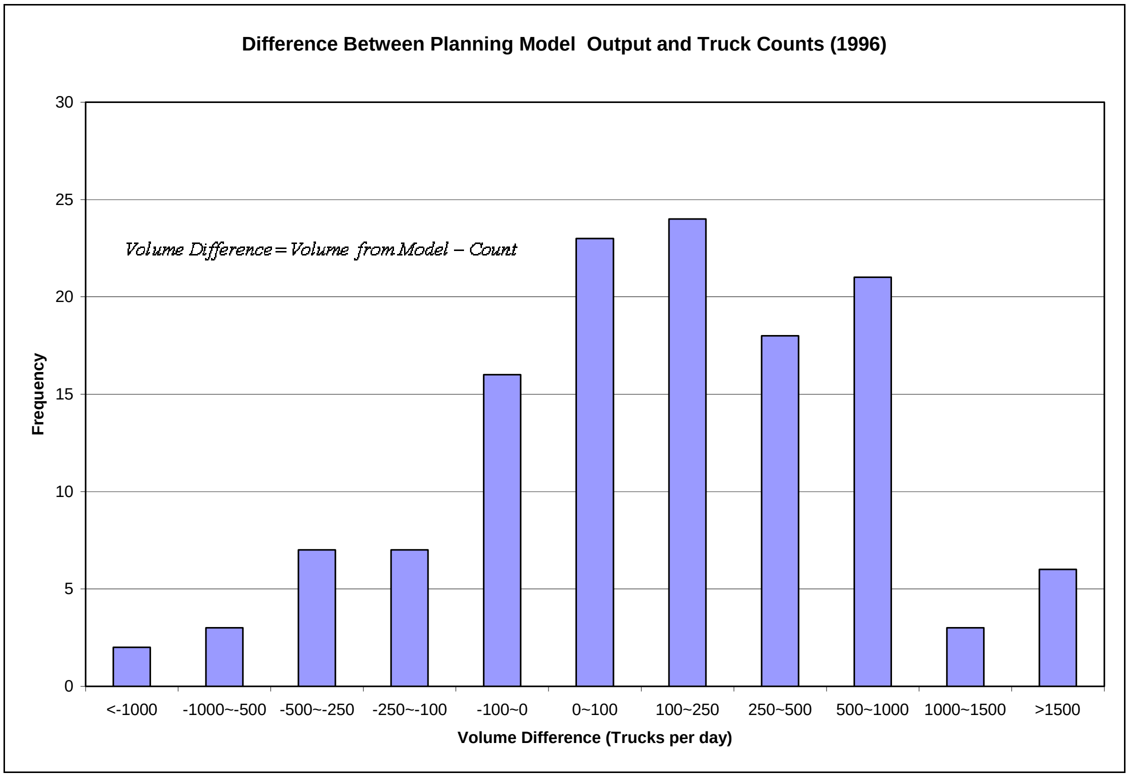


Figure G-3 Difference Between Planning Model Output and Truck Counts (1996)

Appendix H: Productions and Attraction for External Zones (based on synthetic OD model)

EXTERNAL TAZ_ID	TRUCK PRODUCTION	TRUCK ATTRACTION
1807	9190	8971
1808	411	956
1809	686	824
1810	0	0
1811	82	257
1812	35	105
1813	40	48
1814	213	291
1815	306	383
1816	450	580
1817	889	568
1818	272	597
1819	256	185
1820	3041	3092
1821	1081	397
1822	2505	708
1823	27	57
1824	124	87
1825	33	26
1826	16	11
1827	50	55
1828	120	144
1829	25	38
1830	64	73
1831	15	14
1832	216	202
1833	60	68
1834	9	12
1835	40	64
1836	151	186
1837	15	27
1838	60	3340
1839	541	193
1840	589	296
1841	7544	4980
1842	7516	189
1843	253	125
1844	35	62
1845	1033	1643
1846	163	192
1847	147	235

1848	938	908
1849	42	39
1850	37	39
1851	124	116
1852	446	482
1853	33	28
1854	116	65
1855	64	50
1856	16	5
1857	1186	812
1858	1838	1243