MODELLING THE DETERMINANTS OF TRUCK TOURS WITHIN CANADIAN MARKETS

Kevin Gingerich, Hanna Maoh, Bill Anderson
University of Windsor

Introduction

The modern supply chain process relies on the transfer of goods between different entities such as manufacturers, distributors, and retailers within and between various regions. Trucks have remained one of the most predominant shipping methods since other modes of transport often have limited location accessibility and time constraints.

Freight transportation models have been historically based on aggregate models of trip generation and distribution between origin and destination zones. The National Cooperation Highway Research Program has categorized common freight models into five aggregate model types (TRB, 2008). However, these aggregate models are often criticized for their inability to model complex freight movements as they lack behavioral realism and adequate policy handles (Miller and Hatzopolou, 2008).

In response to concerns towards aggregate freight models, state of the art modelling practices, particularly within academia, have shifted towards activity based freight models that represent the microscopic movement of commercial goods (Ferguson et al. 2012). This predominantly includes two classes of models (Chow et al., 2010): logistics models used to track the movement of goods across a supply chain from the manufacturer to the final consumer (Raothanachonkun et al., 2008; Holguín-Veras and Thorson, 2003) and truck tour models (for example: Hunt and Stefan, 2007; Ferguson et al. 2012).

The work presented in this paper contributes to the latter paradigm by conceptualizing a new tour-based model (Figure 1) and developing two of its components: (1) Number of Stops, and (2) Type of Stops.

To our knowledge, the modeling approach presented here has not been conducted in previous studies.

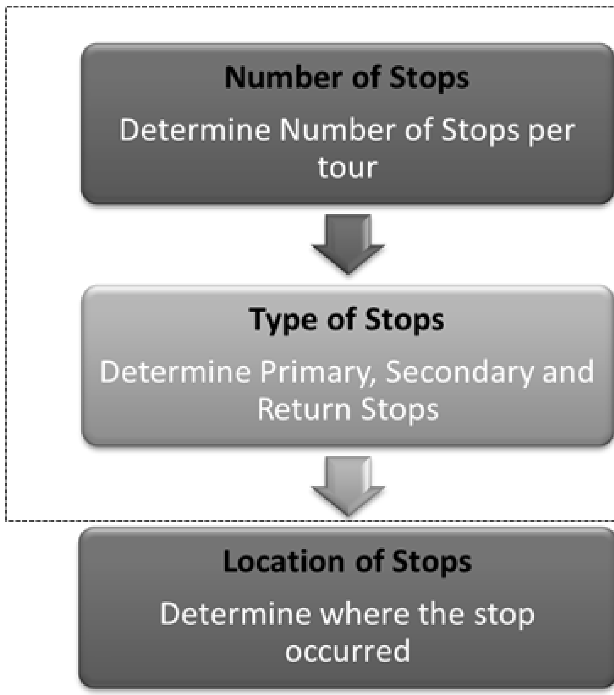

Figure 1: Tour-Based Modeling Framework

The rest of this paper begins with a background that emphasizes the nature of commercial tours. It also provides a brief review of previous studies that have used GPS data as an input for transportation models with a focus on truck tours. The implementation of GPS data used in this paper is then discussed along with the methods used to formulate the models. This includes an ordered logit model used to determine the frequency of stops for a given tour, and a multinomial logit model used to predict the type of stop that is most likely to occur among three alternatives: primary stop, secondary stop, and return to establishment. The subsequent section presents the statistical results from estimating the order logit and multinomial logit models. Finally, a conclusion section is provided at the end.

Background

A truck tour is typically defined as a round trip where a truck leaves the establishment to perform one or more stops (to transfer goods, provide services, or take a break) before returning to the establishment. A two-leg tour is performed if the truck only transfers goods or provides services at one intermediate stop. However, time is utilized more efficiently when a truck makes multiple intermediate stops before returning to the starting location in a process known as trip chaining. Figure 2 shows an example of a truck performing a multi-leg tour via trip chaining.

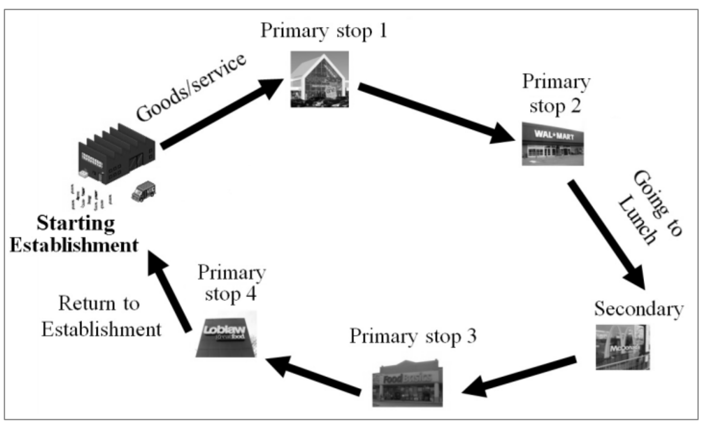

Figure 2: Sample Truck Tour

We classify the intermediate stops occurring throughout the truck tour as either primary or secondary. A primary stop occurs when goods are transferred between the truck and location. A secondary stop occurs when the vehicle dwells at a location for other utilitarian purposes such as fuel refills or a driver break. Finally, the last stop of a tour is a return to the original starting establishment thereby completing the tour.

The increasing availability of passive data derived from global positioning system (GPS) technology to observe the movement of freight trucks has led to a heightened interest in truck tour models. In the U.S., numerous studies have utilized GPS data purchased from the American Transportation Research Institute (ATRI) to analyze and model truck movements. Recent studies that have utilized ATRI data include truck tour models generated by Kuppam et al., 2014 and OD trip matrices developed for Florida (Zanjani et al., 2015), Iowa, and Tennessee (Bernardin et al., 2015).

The truck tour models created by both Kuppam et al. (2014) and Hunt and Stefan (2007) created similar models to those proposed for this project. However, the modelling efforts for this paper can be differentiated from the above studies in several ways. First, the secondary stop locations where goods are not transferred are identified in this model and separated from primary stops. In addition, the location of tour start/end is identified here as the shipping depot. This provides a more realistic starting point compared to other approaches such as the first location of a vehicle for a given day.

Methods of Analysis

This study utilizes GPS data obtained from Transport Canada for March, 2013. The data is derived from 750 Canadian owned carriers conducting trade throughout Canada and the U.S. These carriers have a combined total of 56,000 trucks that produced approximately 100 million GPS pings. Each GPS ping provides typical information including unique carrier and truck identification fields, the location (latitude and longitude), date, and time.

A tour was delineated here when a truck leaves a shipping depot belonging to the carrier and returns at a later time. The location of the primary shipping depot for each carrier in the dataset was determined using a density procedure outlined in Gingerich et al. (2014). Approximately 50,000 truck tours were generated for tours with a total duration between one to eight hours. The one hour lower bound was specified to remove short tours generated as a truck moves within or near the shipping depot location for issues such as vehicle maintenance. The eight hour upper bound was selected to limit the tour to a common work day for a single driver. Longer tours will be investigated in future studies.

Stop events for a given tour were identified when a truck was relatively stationary for 15 minutes or longer. Moreover, the type of stop was classified as a primary stop, secondary stop, or return stop. Primary stops occur when goods are transferred between the truck and a location or another truck. Secondary stop events occur when the truck stops at a location such as a rest stop to fulfill other needs such as a fuel refill or driver break. Secondary stops were identified in the dataset using an entropy classification method (Gingerich et al. (2014) and proximity to known secondary locations from a comprehensive firm database. Any intermediate stop not identified as secondary is given the primary stop label. Finally, the return stop is identified as the last stop for a tour when the vehicle returns to the shipping depot.

A random sample of 500 tours was drawn from the full set of 50,000 truck tours. However, only 267 valid tours remained after removing truck tours that had no visible primary stops in which goods were transferred. The proportion of stops by type for this remaining set of 267 valid tours includes 58% primary stops, 8% secondary stops, and 34% return stops as shown in Figure 3. The geographic distribution of these tours (by starting location) is shown in Figure 4 with the majority of tours occurring in the provinces of Ontario, Alberta, and Quebec.

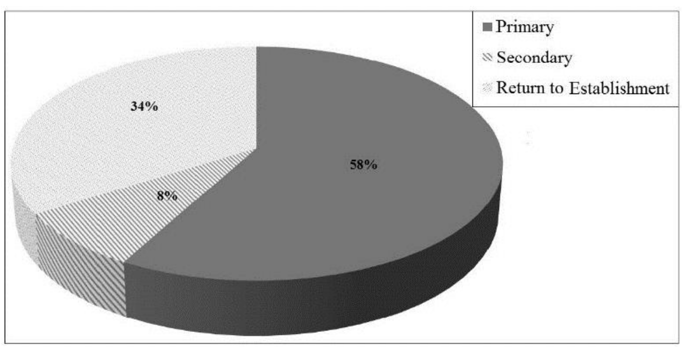

Figure 3: Proportion of stops by type

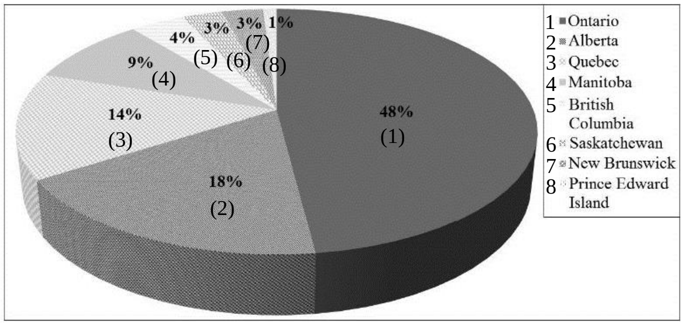

Figure 4: Proportion of truck tours by Canadian Province

Stop Frequency Model

Two ordered logit models were estimated to investigate the determinants influencing the number of stops for a given tour. This includes one model with the total number of intermediate stops (primary or secondary) and another model consisting of only primary stops. The 267 valid truck tours exhibit a range of intermediate stops from one to seven as shown in Figure 5. A negative relationship can be seen with the majority of tours conducting only one or two intermediate stops.

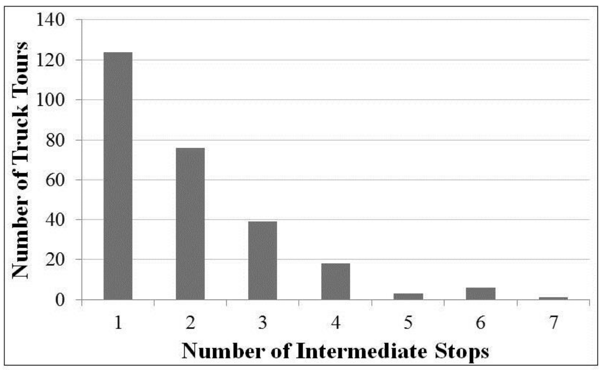

Figure 5: Frequency of tours classified by the number of intermediate stops

Due to the small frequency of tours with stops greater than four, the possible values for the ordinal dependant variable y was reclassified as 1, 2, 3, or 4+ stops. The utility function for the ordered logit can be described as follows:

$$y^* = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \varepsilon \quad (1)$$

Where y^*, an unobserved continuous measure, is related to the number of stops, β_0 is the constant intercept value, β_i is a set of parameters weighting the effects of the x_i explanatory variables, and ε is the stochastic error component assumed to follow an iid logistic distribution.

The discrete observed number of stops (y) can be related to the continuous unobserved value (y^*) as follows:

$$\begin{aligned} y &= 1 \quad ; if\ y^* \leq 0 \\ y &= 2 \quad ; if\ 0 < y^* \leq \mu_1 \\ y &= 3 \quad ; if\ \mu_1 < y^* \leq \mu_2 \\ y &= 4+ \ ; if\ y^* > \mu_2 \end{aligned} \quad (2)$$

where y is the number of intermediate stops, μ_1 and μ_2 are threshold values.

Four types of explanatory variables were included in the ordered logit models to explain the frequency of observed stops for a given tour, as shown in Table 1. The total time that it takes the truck to complete a tour (TOUR DURATION) is expected to be positive since tours with a longer duration tend to make a larger number of stops.

Table 1: Explanatory variables for the stop frequency model

Variable	**Description**
TOUR DURATION	Cumulative time (minutes) taken to complete the tour
INDUSTRY$_n$	1 if tour pertains to industry type *n*; 0 otherwise
METROPOLITAN$_k$	1 if tour started in metropolitan area *k*; 0 otherwise
RURAL	1 if tour started in a rural area; 0 otherwise

Dummy variables for specific industry types (INDUSTRY$_n$) were included to control for possible truck behaviour differences arising from varying industry needs. The industry for the nearest firm to each primary stop was typically assigned to the given truck tour. Subsequently, a tour with multiple primary stops may have multiple industries assigned to it. The firm dataset was purchased from InfoCanada.

Location specific dummy variables were also included in the ordered logit models based on the shipping depot site ($METROPOLITAN_k$). This includes various metropolitan areas such as Winnipeg and Toronto. In addition, a dummy variable for rural areas (RURAL) was introduced into the model to capture tours beginning outside of any major metropolitan areas.

Stop Type Model

While the two stop frequency models discussed above can be used to estimate the number of generated stops for a given truck tour, those stops vary in their type. Therefore a multinomial logit model (MNL) was estimated to model the type of stop that is most likely to occur during a tour. The utility function for the MNL model is similar to equation 1 except for the dependent variable.

The dependant variable in this model is the stop type consisting of three possible alternatives. The first type is a primary stop (P) where goods are transferred between the stop location and given truck. The second option is a secondary stop (S) where other needs are met such as fuel refills. These two stop types represent intermediate stops that can occur throughout the truck tour. Finally, we have included a third stop type representing the return (R) to starting point. The occurrence of this stop type marks the completion of the truck tour. 790 observations are used in the MNL model that represents stops conducted during the 267 truck tours.

A condition is initially imposed on the tours to make sure they have at least one primary stop where goods are transferred. A result of this condition is that the return to starting point is not possible until a primary stop has occurred. Subsequently, the choice set for some tours in the MNL model consisted of two alternatives (i.e. **P** or **R**) while it contained the three types of stops (**P**, **S** or **R**) for multi-leg tours. It should be noted that the third stop type representing a return to the starting point becomes available after at least one primary stop has occurred.

The explanatory variables for the MNL stop type model are provided in Table 2 along with the alternative stop type (**P**, **S** or **R**), description, and expected sign. A truck that has already completed several primary stops is expected to be more likely to perform a secondary stop. Therefore the number of previous primary stops for the given truck tour (PRVPS) is included in the model. Similarly, the number of previous secondary stops (PRVSS) is expected to increase the likelihood of the current stop occurring for primary purposes. However, these variables were not statistically significant to 90% in the model.

Table 2: Explanatory variables for the stop type model

Variable	Alt.	Description	Sign
PRVPS	S	Number of previous primary stops	+
PRVSS	P	Number of previous secondary stops	+
CTIME	R	Cumulative time (minutes) from start of tour until beginning of observed stop	+
NIGHT	R	1 if tour started at night (9pm - 6am); 0 otherwise	+
$INDUSTRY_n$		1 if tour pertains to industry type *n*; 0 otherwise	
$METROPOLITAN_k$		1 if tour started in metropolitan *k*; 0 otherwise	

Time is also expected to play a major role in the type of stop that will occur. If a truck tour has already spent considerable time on the road, it is more likely to perform a return to the establishment (i.e. starting point) to complete the truck tour. The cumulative time (CTIME) since the start of the tour is used to capture this phenomenon. We also expect those tours which begin at night to conduct fewer intermediate stops. Therefore a night time dummy variable (NIGHT) was introduced with an expectation that the return stop for these truck tours will be more likely to occur compared to intermediate stops. Finally, variables similar to the ordered logit model were included to capture any potential heterogeneity in the preferences by industry type ($INDUSTRY_n$) or metropolitan area ($METROPOLITAN_k$).

Results

The results for the **Number of Stops Model** (i.e. total stop and primary stop frequency) are given in Table 3. The two models behaved similarly with only a few minor differences. The parameter for tour duration is positive as expected. This suggests that a tour that takes longer to complete will typically result in more stops, other things being equal.

Table 3: Stop frequency model results

Variable	**Total Stops**		**Primary Stops**	
	Value	**T-stat**	**Value**	**T-stat**
CONSTANT	-3.62	-7.76	-4.12	-7.88
TOUR DURATION	0.58	7.45	0.52	6.40
TRANSPORTATION	1.32	3.91	1.32	3.82
MANUFACTURING	1.07	2.82	1.33	3.51
RETAIL	1.19	3.43	1.39	3.88
SERVICES	0.71	1.93	0.82	2.19
FIRE	2.30	2.86	2.46	3.40
MINING	3.91	2.48	4.19	2.63
WINNIPEG	-1.34	-1.62	-0.65	-0.78
TORONTO	1.13	3.31	0.49	1.35
PRINCE GEORGE	1.57	2.10	2.17	2.81
CALGARY	1.06	2.06	1.11	2.01
WINDSOR	1.36	2.03	1.99	2.95
RURAL	0.76	2.09	1.20	3.09
Threshold μ_1	1.94	11.37	1.95	10.35
Threshold μ_2	3.61	13.74	3.72	11.71
Observations	267		267	
Chi-squared	159.55		137.23	
Degrees of freedom	13		13	
Prob[Chi^2 > value]	0.000		0.000	
ρ^2	0.243		0.235	

Tours associated with certain industries exhibit a higher number of stops for a given tour. Based on the value of the parameters, truck tours associated with the mining industry have the largest number of intermediate stops. This is followed by tours pertaining to the real estate industry.

Among the location variables introduced in the models, Winnipeg was the only metropolitan area with a negative parameter indicating that fewer stops occur for tours starting in this area. However, the Winnipeg parameter was slightly under 90% statistical significance in the total stop frequency model and much less significant in the primary stop frequency model. Several other location parameters were positive and statistically significant. The parameter for Prince George, BC, has the largest value suggesting that more stops occur there compared to any other locations for our sample of truck tours. Interestingly, while the Toronto location variable increase the propensity of generating total stops this effect diminishes in the primary stop model.

The results obtained for the **Stop Type Model** (i.e. multinomial logit model) are provided in Table 4. The constants included for primary stops and secondary stops are positive and negative, respectively. This fits with the overall proportion of stops as seen in Figure 3 due to the high proportion of primary stops.

Table 4: Stop type model results

Parameter	Utility	Coefficient	T-stat
CONSTANT	P	1.28	4.97
PRVSS	P	0.29	1.50
RETAIL	P	0.86	3.83
MANUFACTURING	P	0.49	2.04
TRANSPORTATION	P	0.99	4.04
CONSTANT	S	-1.19	-3.68
PRVPS	S	0.19	1.54
TRANSPORTATION	S	0.90	2.44
TORONTO	S	1.46	5.05
CTIME	R	0.48	7.94
WINNIPEG	R	1.51	2.43
NIGHT	R	0.40	1.49
Number of observations		790	
Log-Likelihood (constants only)		-554.78	
Log-Likelihood (Final)		-486.93	
ρ^2		0.122	

The number of previous secondary stops (PRVSS) for a given tour has a positive impact on the primary stop type. In addition, the number of previous primary stops (PRVPS) for a given tour has a positive impact on the secondary stop type. This suggests an equilibrium effect for primary and secondary stops over the course of a given truck tour.

Heterogenous behaviour across various industry types was captured through various industry dummy variables. The results suggest that retail and manufacturing industries are predisposed towards more primary stops compared to other industries. Moreover, truck tours associated with the transportation industry had positive parameters for both primary and secondary stops, suggesting that they make more stops of both types before returning to the starting point compared to other industries. It should be noted that the transportation industry simply indicates that the truck made one or more intermediate stops at the location of a transportation company such as another shipping depot. Finally, tours starting in Toronto were more likely to perform secondary stops, all other things being equal. This finding is in line with the results of the ordered model, which show that the Toronto tours are more secondary in nature.

The return to establishment stop type also has several statistically significant parameters. The positive parameter for the cumulative time (CTIME) from the start of a given tour to the current stop reflects the limits of work hours on the driver. As the tour time gets longer, the tour completion marked by a return to establishment becomes more likely. Another parameter that was found to be positive when connected to the return stop type was the location variable for tours starting in Winnipeg. This suggests that tours in this area tend to make fewer intermediate stops (primary and secondary) compared to other locations. That is, most of the tours in this region pertain to two-leg tours. A third parameter applied to the return stop type was night time tours. This parameter is positive suggesting that night tours tend to have fewer intermediate stops but the statistical significance is below 90%.

Conclusions

This paper outlined current efforts underway to convert GPS data observing freight movements into models truck tours. A valid tour is observed when a truck leaves a shipping depot belonging to the carrier and conducts at least one primary stop before returning to the starting location. In addition, only tours with time duration ranging from one to eight hours were utilized since they represent a tour consistent with a single day of work. Future work may also look at the determinants of truck tours occurring for longer periods of time.

The frequency of stops occurring for a given truck tour was analyzed using an ordered logit model. The results of this model confirm the intuitive behaviour that a longer truck tour is more likely to perform more stops. The models found some variance by industry with mining and FIRE industries performing the most stops compared to all other industries. There was also some variation between geographic starting locations, with more stops occurring for locations such as Prince George, British Columbia and Windsor, Ontario.

The type of stop occurring throughout a truck tour was also analyzed using three stop types: primary stops, secondary stops, and a return to establishment. The number and type of previous stops for a given tour have some impact on the choice of stop type, though these parameters are below a 90% statistical significance. In addition, a longer cumulative time for a given tour makes it more likely that the stop type would be a return to establishment. Finally, some variation occurs between industry types and the starting geographic location of the tour.

These models provide some insight into the determinants of truck tours in the Canadian context. However, their real value lies in the combination of these models along with others (such as the destination choice of stops and stop duration) to provide a more complete truck tour model. This would allow the model to be used as a microscopic tool suitable for the simulation of truck tours. Future research will focus on developing destination choice of stops and stop duration models to implement and test an operational truck tour model for a one working day tours.

Acknowledgements
The authors wish to thank Transport Canada for providing the GPS dataset used in this paper. We are also thankful to FedDev Ontario for their financial support.

References

Bernardin, V., Trevino, S., Short, J., 2015. Expanding truck GPS-based passive origin-destination data in Iowa and Tennessee. *Proceedings of the 2015 Transportation Research Board Annual Conference*, Washington, DC.

Chow, J., Yang, C., Regan, A., 2010. State of the art of freight forecast modeling: lessons learned and the road ahead, *Transportation*, vol. 37, pp. 1011-1030.

Ferguson, M., Maoh, H., Ryan, J., Kanaroglou, P., Rashidi, T. 2012. Transferability and enhancement of a microsimulation model for estimating urban commercial vehicle movements. *Journal of Transport Geography*, 24, pp. 358-369.

Gingerich, K., Maoh., H., Anderson, W., 2014. Mining freight transportation activity location & type using large-scale passive GPS datasets. Proceedings of the 49th annual CTRF Conference, pp. 156-170.

Holguín-Veras, J., Sánchez, I., González-Calderón, C., Sarmiento, I., Thorson, E. 2011. Time-dependent effects on parameters of freight demand models, *Transportation Research Record: Journal of the Transportation Research Board*, no. 2224, pp. 42-50.

Hunt, J.D., Stefan, K.J., 2007. Tour-based microsimulation of urban commercial movements. *Transportation Research Part B*, 41, pp. 981-1013.

Kuppam, A., Lemp, J., Beagan, D., Livshits, V., Vallabhaneni, L., Nippani, S., 2014. Development of a tour-based truck travel demand model using truck GPS data. *Proceedings of the 93rd Annual Transportation Research Board Conference*, Washington, DC.

Miller, E. Hatzopoulou, M. 2008. *Organizational structures for regional travel demand modelling*, Report submitted to Metrolinx, [online] Accessed on October 17, 2013, Available at:<http://www.citiescentre.utoronto.ca/Assets/Cities+Centre+2013+Digital+Assets/Cities+Centre/Cities+Centre+Digital+Assets/TMG+files/TMG+-Organizational-Structure-Report_2008.pdf>

Raothanachonkun, P., Sano, K., Wisetjindawat, W., Matsumoto, S., 2007. Truck trips origin destination using commodity based model combined with an empty trip model, *Transportation Research Record: Journal of the Transportation Research Board*, no. 2008, pp. 43-50.

Transportation Research Board (TRB), 2008. *National Cooperative Highway Research Program Report 606: forecasting statewide freight toolkits*, Transportation Research Board of the National Academies, Washington, DC. Accessed on Sept. 19, 2014, <http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_606.pdf>.

Zanjani, A., Pinjari, A., Kamali, M., 2015. Estimation of statewide origin-destination truck flows using large streams of GPS data: Application for Florida statewide model. *Proceedings of the 94th Annual Transportation Research Board Conference*, Washington, DC.