USING LARGE SAMPLE GPS DATA TO DEVELOP AN IMPROVED TRUCK TRIP TABLE FOR THE INDIANA STATEWIDE MODEL

Vincent L. Bernardin, Jr., Ph.D.
Bernardin, Lochmueller & Associates, Inc.
6200 Vogel Road
Evansville, IN 47715
VBernardin2@BLAinc.com
Ph: 812-479-6200
Fax: 812-479-6262

Jonathan Avner, PTP
Wilbur Smith Associates
9500 Arboretum Blvd., Suite 360
Austin, TX 78759
JAvner@WilburSmith.com
Ph: 512-592-3842

Jeffery Short
American Transportation Research Institute
2060 Franklin Way, Suite 201
Marietta, GA 30067
JShort@trucking.org
Ph: 770-432-0628

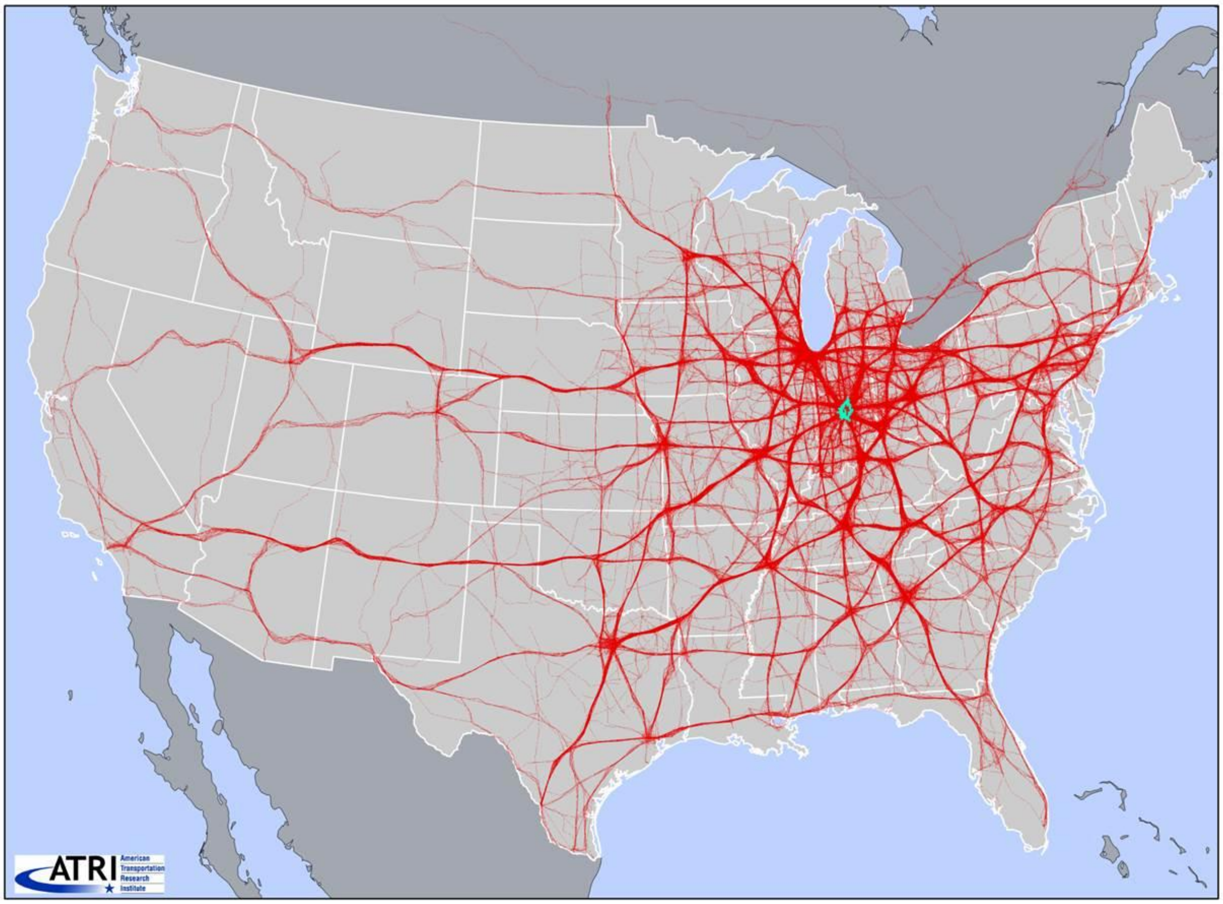
Laurence Brown, Roy Nunnally and Steve Smith, AICP
Indiana Department of Transportation (INDOT)
100 North Senate Avenue, Room N955
Indianapolis, IN 46204
lbrown@indot.IN.gov, rnunnally@indot.IN.gov, ssmith@indot.IN.gov
Ph: 317-232-5646
Fax: 317-234-1228

Submitted: November 16, 2011
1,774 Words, 2 Table, 3 Figures

Background

The Indiana Department of Transportation (INDOT) with the assistance of Wilbur Smith Associates (WSA) and Bernardin, Lochmueller & Associates (BLA) has undertaken in an innovative study to analyze and develop a truck trip table from a large sample of truck GPS data available from the American Trucking Research Institute (ATRI). The truck trip table developed will ultimately serve to enhance the Indiana Statewide Travel Demand Model (ISTDM) and improve its truck forecasts.

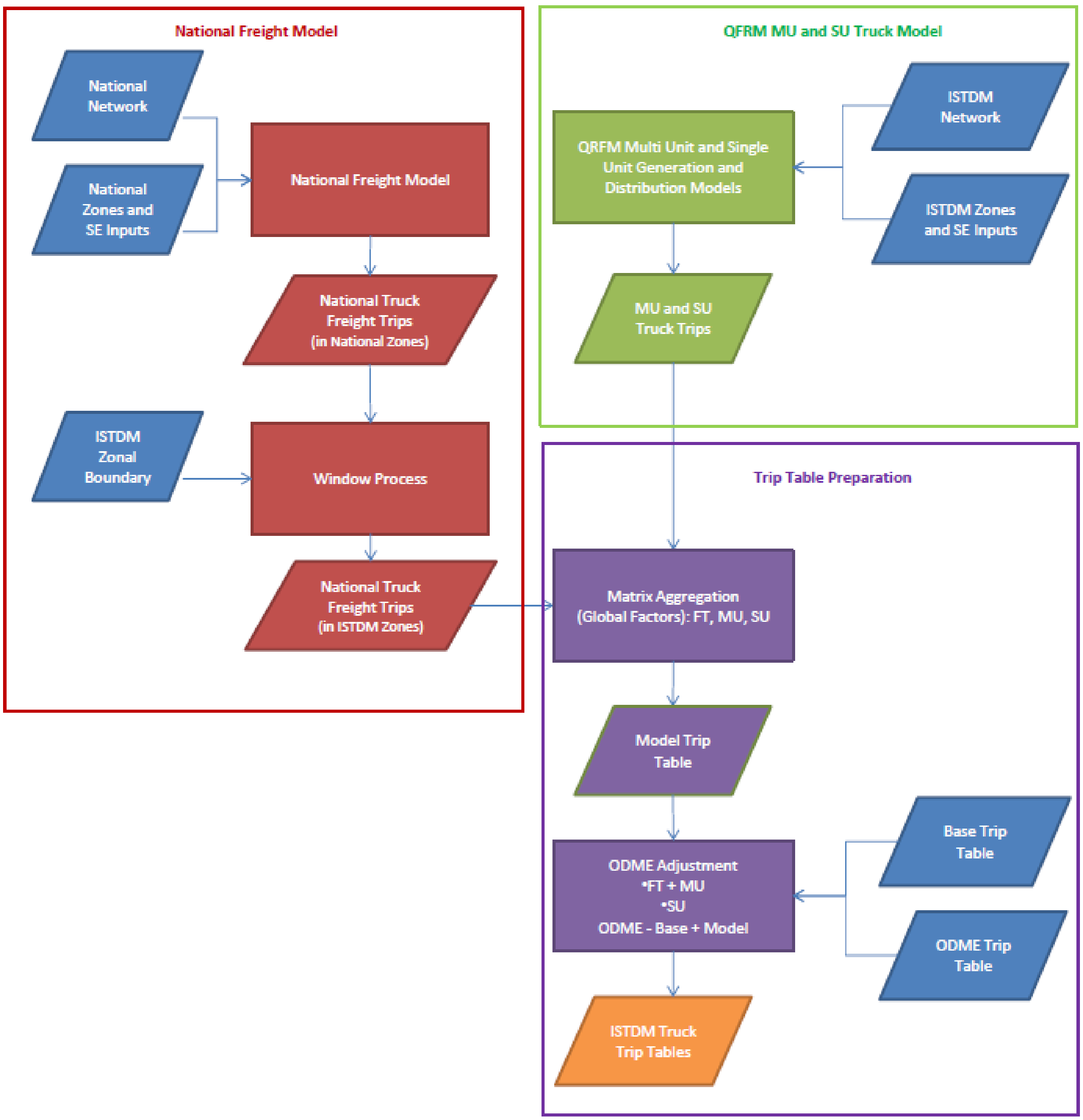
Figure 1 One Week of ATRI's GPS Traces of Trucks Traveling in or through Indiana



Recently, INDOT lead a four-state coalition including the Ohio, Indiana, Illinois and Missouri DOTs that engaged in a feasibility study of dedicated truck lanes on I-70 with funding from FHWA's Corridors of the Future program. As a part of this study, WSA and BLA used truck GPS data from ATRI to analyze truck flows on the I-70 corridor. The GPS data proved extremely helpful in understanding the length and pattern of truck movements utilizing the corridor as well as "ground-truthing" and improving travel demand model forecasts. The experience with the I-70 study lead INDOT to engage BLA and WSA to investigate the incorporation of ATRI's GPS truck data in the next update of the ISTDM.

The first generation of the ISTDM was developed in the mid-1990's. Since that time the model has been through five major updates and been used continuously for numerous corridor studies throughout the state as well as in support of the development of statewide transportation plans. It was showcased as an example of best practice in *NCHRP Synthesis 358: Statewide Forecasting Models*. It currently includes a traditional four-step model of passenger travel with a long distance mode choice model as well as a truck model.

Figure 2 Indiana Truck Model Components



The existing structure of the ISTDM's truck model includes three elements. The components are a national commodity freight flow component, a local truck model based on the *Quick Response Freight*

Manual and finally matrix preparation and adjustment steps in which growth from the commodity flow and local truck models are used to pivot off of a base year truck trip table previously developed using origin-destination matrix estimation (ODME).

Based on the initial understanding of the ATRI data, the assumption was made that the truck trips included in the data were best described as long distance commodity flow truck movements. Therefore in the above structure, the ATRI trips could be used to develop more accurate seed and resulting ODME trip tables which serve as the basis for the model, while the commodity flow and QRFM would continue to supply the future growth for truck forecasts.

Data and Methodology

ATRI's database of GPS traces of trucks includes nearly 4 billion truck positions annually which provide empirical evidence of where and when real truck activity occurs. ATRI obtains this information through several private sector data-sharing partnerships and compiles it as part of the Freight Performance Measures Initiative, an effort sponsored in part by the Federal Highway Administration. The primary source of these data is on-board communications equipment installed on large commercial trucks (e.g. tractor-trailer combinations). Prior to processing, each truck GPS data record included in this research included an anonymous, unique truck identification number, latitude, longitude and a time stamp.

Sample datasets were compiled for eight one-week time periods drawn from February, May, July and October of 2010. Although the ISTDM's commodity flow model covers the lower 48 states, the study area of the ISTDM's truck model is focused on Indiana, extending roughly 50 miles beyond the state borders. The GPS data were therefore filtered spatially using GIS, to capture recorded truck positions in the study area or in external station catchment areas immediately adjacent.

Truck position records within the study area were then associated with the model's traffic analysis zones (TAZ) based on their latitude and longitude using GIS functionality. In addition to allowing the ultimate development of a trip table, this process also offered some benefit of further ensuring the anonymity of the data by associating truck positions with geographic areas far more generalized than a latitude/longitude position.

Within the study area, a set of polygons was also produced for large truck stops. This set of polygons was used as a filter to remove instances where trucks stopped for refueling, compliance with Hours of Service Regulations, or similar purposes. The final filtered dataset containing TAZ numbers as the key geographical component was exported from the GIS software.

A data management and analysis software package was used to further prepare the dataset for integration into the truck trip table. Truck positions for each unique vehicle were sorted into a time series, and within each series each truck position was matched with the subsequent truck position to produce a set of truck position pairs. The geodetic distance between the first and second truck positions for each of the truck position pairs was then calculated. This distance was used later for the identification of truck origins and destinations (stops) and the latitude and longitude values were dropped from the dataset to ensure anonymity. The resulting dataset contained the following variables for each set of truck position pairs: truckid, 1st date and time, 1st TAZ, 2nd date and time, 2nd TAZ, geodesic distance travelled.

From the list of records made by each truck, the trips were filtered to define individual trip records. The first step was to identify the records where the truck was considered in motion. This was based on a minimum travel speed based on the geodesic distance and time stamps. A minimum distance could not be used because the reporting interval varies based on the equipment used by the individual truck. Once the movements and stops were identified, the records were processed to identify the origin and destination for each sequence of moving records. When a stop record was found in the list, it signified the end of the trip.

As part of the ongoing model development, additional research is being done on the filtering or definition of moving records versus stops. In addition to a minimum travel speed criteria, a minimum travel time is being considered as well. The intent of adding a second criteria is to reduce the number of false stops in a sequence of trips. False origins or destinations can result from trucks being stopped at long traffic signals or in traffic jams or from noise in GPS position readings

Table 1 below reports the number of records, unique trucks, movement records, and trips for the datasets used.

Table 1 Summary of Processed GPS Data

		Raw	Seasonal	Adjusted
MAY	# of Records	6,802,370	0.971	6,605,101
	# of Unique TruckIDs	102,640		99,663
	# of Moving Records	3,130,121		3,039,347
	# of Trips	668,778		649,383
	Average Trip Length	67.4		67.4
OCT	# of Records	6,888,083	0.961	6,619,448
	# of Unique TruckIDs	72,751		69,914
	# of Moving Records	3,420,513		3,287,113
	# of Trips	660,007		634,267
	Average Trip Length	61.4		61.4
JUL	# of Records	6,559,498	0.944	6,192,166
	# of Unique TruckIDs	104,814		98,944
	# of Moving Records	3,054,682		2,883,620
	# of Trips	635,601		600,007
	Average Trip Length	66.0		66.0
FEB	# of Records	5,282,667	1.128	5,958,848
	# of Unique TruckIDs	107,408		121,156
	# of Moving Records	2,239,184		2,525,800
	# of Trips	528,894		596,592
	Average Trip Length	74.3		74.3

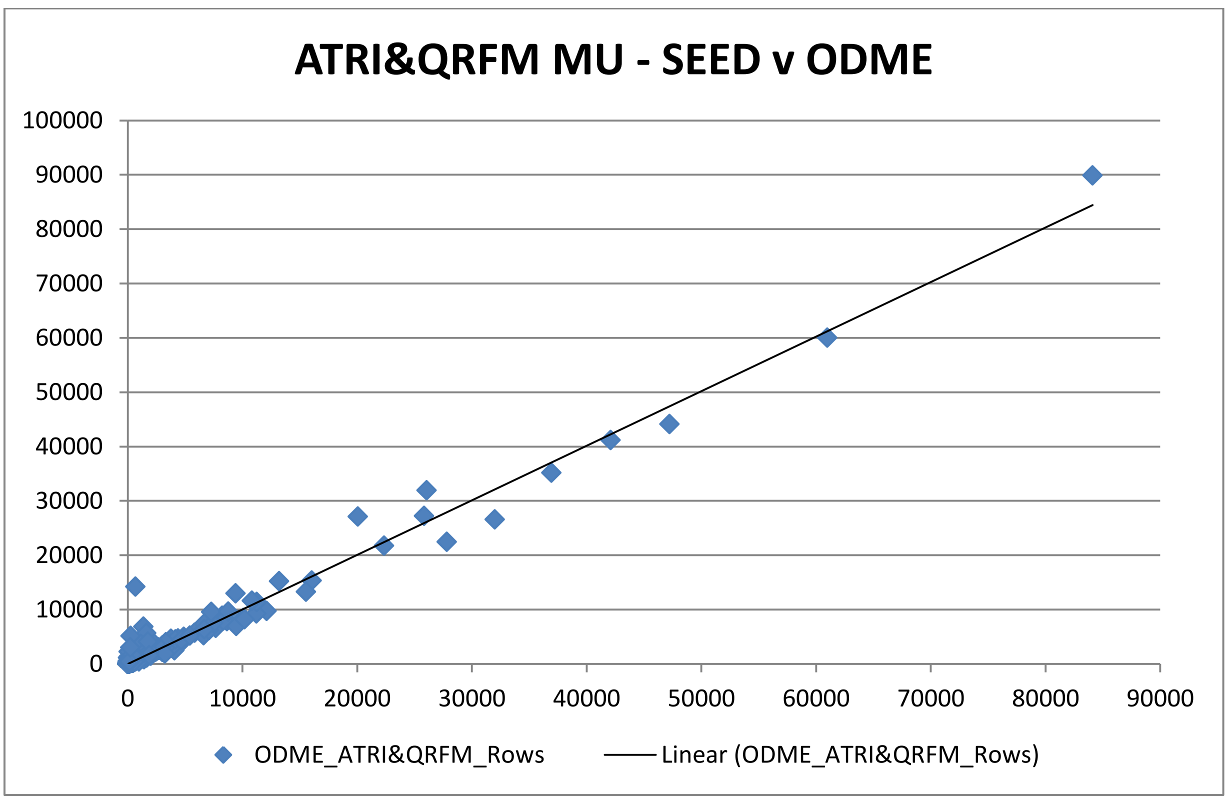
Preliminary Results

The resulting trip records from the eight samples were adjusted based on INDOT monthly adjustment factors. The ATRI trips were further adjusted by a sample expansion factor of 0.125 to generate approximately 300,000 daily trips. The adjustment was based on comparing the weekly ATRI trip total to the daily truck commodity trips from the statewide model.

An initial ODME run was made based on over 8,400 truck counts throughout the study area and using a seed trip table created by adding the ATRI trip table to QRFM multi-unit trucks, under the preliminary assumption that ATRI trip table corresponds to and could potentially replace the multi-unit commodity flow-based truck trips. Alternative hypotheses, such as that the ATRI data includes a reasonable sample of local multi-unit or even single unit trucks, will also be tested.

The figure below shows the comparison of the trip ends from the seed trip table versus the ODME output. The R-squared value was 0.9751, suggesting that very little adjustments to the seed trip table were necessary to reproduce the observed truck counts on the network.

Figure 3   Comparison of Trip Ends between ATRI + QFRM Seed and ODME Output



Application

Based on the favorable outcome of the ODME that implied the seed was a good fit to the traffic counts, attention has focused on how to apply the ODME to the Indiana Statewide Model. As discussed previously, the intent was to use the ODME generated from the ATRI data as a replacement for the previous ODME trip table, but retain the same structure of the model.

To validate this methodology, the resulting orientation or distribution pattern of the ATRI trips versus the freight trips was considered as well as the average trip length. What was found was that the ATRI trips and the output from the commodity truck model were not consistent geographically or in magnitude. The average trip length from the freight trips from the ISTDM is 180 miles versus the average of 67 miles for the ATRI data. Further, the commodity trips are more focused on external movements with nearly ½ of the total trips being external – external.

Table 2 Preliminary GPS-Based District to District Flows

ATRI	Externals	Ohio	Michigan	Illinois	Kentucky	Central Indianna	Southwest Indianna	Northwest Indianna	Northeast Indianna	Southeast Indianna	Total
Externals	2,913	5,016	2,131	8,250	5,041	820	1,238	2,967	1,541	1,310	31,226
Ohio	4,702	29,604	3,173	1,016	2,690	610	332	1,398	2,433	1,167	47,124
Michigan	2,011	3,259	23,760	1,076	303	143	84	1,232	1,272	79	33,219
Illinois	8,415	1,066	1,025	53,952	1,017	738	1,261	5,980	1,074	586	75,114
Kentucky	5,227	2,521	282	1,027	15,242	352	1,671	542	231	1,264	28,360
Central Indianna	828	584	162	742	355	5,506	814	1,962	1,130	1,176	13,259
Southwest Indianna	1,332	353	95	1,147	1,818	789	7,080	852	302	908	14,676
Northwest Indianna	2,996	1,376	1,192	6,243	563	1,846	980	13,551	2,068	943	31,759
Northeast Indianna	1,461	2,260	1,242	1,143	253	1,129	302	2,313	12,723	545	23,370
Southeast Indianna	1,131	1,214	114	697	1,105	1,268	889	1,049	603	3,853	11,924
Total	31,017	47,253	33,176	75,294	28,386	13,201	14,651	31,846	23,377	11,831	310,031

Several possibilities are being explored at the time of this writing. One possibility is that the ATRI data includes significantly more short-haul, local delivery movements which are not reflected in the commodity flow data. If this is the case, the ATRI data may be better compared to the ISTDM's total truck trip table rather than only the commodity-flow-based portion. There is preliminary evidence for this assumption in the fact that average truck trip length for all ISTDM trucks is 60.2 miles, which compares favorably with the ATRI average of 67 miles. Another possibility with some support from the geographic distribution of the ATRI data, which shows many stops just off of major truck routes, is that despite screening out major known truck stops, the ATRI data may still include many stops on truck trips for re-fueling, meals, restroom breaks or rest breaks related to Hours of Service regulations which are not predicted by the commodity flow models.

Conclusion

A methodology has been presented for processing raw GPS truck data to produce a truck trip table and successfully applied in Indiana. Analysis is on-going at the time of this submission and further research should help to refine the approach and determine how the GPS data can best be used to improve the ISTDM. However, the initial results that relatively small adjustments were needed in order to produce a very high degree of agreement with a large set of observed truck counts suggests substantial agreement between these two large sample datasets. ATRI's GPS data could be used in many other applications. Future work on the ISTDM truck model may use the dataset to revise and improve either or both the commodity flow and/or local truck models in addition to the base truck trip table. For instance, ATRI data could be used for updating local truck trip generation equations or even estimating destination choice models for local trucks. On the freight side, truck trip ends could be used for disaggregating FAF or other commodity flow databases. Far from offering the final word on the use of GPS data in truck modeling, this short paper serves rather as an initial demonstration of the feasibility of using large sample GPS databases to improve truck models.