

USING PASSIVE DATA TO BUILD AN AGILE TOUR-BASED MODEL: A CASE STUDY IN ASHEVILLE

Josephine D. Kressner, PhD (Corresponding Author)

Transport Foundry
3423 Piedmont Rd NE Suite 481
Atlanta, GA 30305

Gregory S. Macfarlane, PhD

WSP | Parsons Brinckerhoff
434 Fayetteville St, Suite 1500
Raleigh, NC 27601

Leta Huntsinger, PhD, PE

WSP | Parsons Brinckerhoff
434 Fayetteville St, Suite 1500
Raleigh, NC 27601

Rick Donnelly, PhD, AICP

WSP | Parsons Brinckerhoff
6100 Uptown Blvd NE Suite 700
Albuquerque, NM 87110

Word Count: 1751 words + 1 figure(s) + 0 table(s) = 2001 words

Submission Date: 2 October, 2015

ABSTRACT

To date, the primary sources of data for travel modeling have been comprehensive household travel surveys, which are collected at high cost and yet may suffer from measurement error and nonresponse bias. There has been interest in using passive data to develop travel models, but modern trip- and tour-based models require a link between demographics and trip-making behavior that is uncommon in passive data. This project will demonstrate an attempt to overcome these limitations by building an agile tour-based model from passive data, relying on an innovative person-based discrete event simulation framework. This model will be validated against a modern trip-based model recently developed for the North Carolina Department of Transportation (NCDOT) and the French Broad River Metropolitan Planning Organization (FBRMPO) covering the Asheville region in North Carolina. The main result from this project will be a detailed comparison of costs, development time, complexity, usability, and accuracy between NCDOT's recent aggregate trip-based model and this new tour-based model constructed from passive data. We believe the agile modeling approach developed through this research will be a good fit for small and medium-sized communities that will help them overcome the challenges of costly data collection and aggregate models that may not fully understand travel behavior.

OBJECTIVES AND MOTIVATIONS

To date, the primary sources of data for travel modeling have been comprehensive household travel surveys, which are collected at high cost and yet may suffer from measurement error and nonresponse bias (e.g., 1, 2). Many have investigated the use of passive location data including GPS, WiFi positioning, Bluetooth, and triangulated signal data as an alternative or supplement to surveys (e.g., 3, 4). However, these types of passive location data typically lack information about the people that make each trip when collected without a coordinating traditional survey. This makes these data ill-suited for building current trip- or tour-based models; instead, the data have primarily been used for survey supplementation (e.g., 5), model validation (e.g., 6, 7), or for specific project studies (e.g., 8). For a comprehensive review of emerging data usage in research and practice to date, see Lee, Sener, and Mullins (9).

Meanwhile, consumer data firms have been compiling information about individuals and households for decades, typically selling the inexpensive, up-to-date data to commercial marketers. These readily available data contain the majority of household and individual demographic and socioeconomic fields that are used in travel demand modeling, but these data lack trip-making behavior. This weakness, along with some concerns over data integrity, has led modelers to rely instead on relatively limited and outdated U.S. Census data.

This project will demonstrate an attempt to overcome the independent respective limitations of passive location and consumer data by building an agile tour-based model that relies on an innovative person-based simulation framework. This model will be validated against a modern trip-based model recently developed for the North Carolina Department of Transportation (NCDOT) and the French Broad River Metropolitan Planning Organization (FBRMPO) covering the Asheville region in North Carolina. The model will be built without local household travel survey data, and will test whether an advanced microsimulation model might be built without the long development times typically experienced with modern activity-based models.

METHODOLOGY

The difference between modern travel modeling approaches and the approach we will use in this project can be explained by examining a simple queue at a bank. Using modern approaches for modeling, one customer out of every 100-200 customers who wait for service would be surveyed. The surveyor would follow that single customer through the queue asking questions and measuring experienced wait times along the way by that customer. To create a model of the system, the survey data would be expanded to represent all of the customers using sophisticated data expansion techniques. In travel modeling specifically, the expansion is done by comparing the demographics of the survey respondents to the demographics of the estimated full population.

In this project's model, different measurements will be taken. Rather than individual customer experiences being surveyed, *events* will be passively measured. This framework is based upon discrete event simulation from the industrial systems engineering field. In the simple queue example, the time between customer arrivals, service times at the teller window, and overall teller details like number of tellers by time of day would be collected by teller computers and other passive monitors. Using the passive data and the discrete event simulation framework, the queue can be modeled in a statistically robust way, where the model is not unduly affected by outliers. At the same time, an analyst studying the simulated queue can still examine individual customer experiences in detail.

For building a person-level tour-based travel model, events can be measured in the same way as they were in the simple queue example. Passive data collected by many different third-party data providers measure events in a transportation system. For example, a GPS data provider can provide statistical characterizations of the time of day that people travel to work from home, summarized by small geographic home zones. Using the same data, an analyst can obtain statistical characterizations of the length of time spent at a workplace before making another trip, summarized by small geographic workplace zones. Using discrete event simulation, data fusion, and other statistical techniques with these statistical characterizations, one can systematically synthesize person-level travel diaries using the same mathematical techniques as in modeling a simple queue.

Using passive data shared by the North Carolina Department of Transportation (NCDOT) and a prototype of this approach, we will first build base-year demand trip tables for the FBRMPO region and feed them into a static assignment model. We will compare the results with NCDOT's demand trip tables produced by their modern trip-based model and fed into the same static assignment model. We will compare the two model results using standard measures of model output as a means of validation. If time permits, we expect to repeat the comparison between this new model and the trip-based one using a dynamic traffic assignment framework such as MatSIM.

EXPECTED RESULTS

The main result from this project will be a detailed comparison of costs, development time, complexity, usability, and accuracy between NCDOT's recent aggregate trip-based model and this new tour-based model constructed from passive data. We expect to find that the passive data model will produce a model with satisfactory accuracy compared with an aggregate trip-based model using standard validation measures. We also expect that the costs and timeline to produce such a model, when compared with a traditional household travel survey and model development, will be lower.

A past National Science Foundation Small Business Innovation Research (NSF SBIR) project and ongoing Transportation Research Board Innovations Deserving Exploratory Analysis (TRB IDEA) project, in cooperation with Puget Sound Regional Council and Atlanta Regional Commission, has and will implement this approach in the city of Atlanta (NSF SBIR) and metropolitan Seattle (TRB IDEA), respectively. The Atlanta and Seattle implementations aim to produce validated synthetic populations with 48 hour travel diaries. The diaries contain person, trip, and other attributes that typically come from travel surveys. The synthetic population diaries currently lack mode and route information.

This project will build on this work and experiment with linking the synthetic travel diaries with an assignment model. We have selected FBRMPO as the case study MPO to implement the approach due to its size, data availability, and its small share of transit and walk/bike trips. The FBRMPO region has an estimated mode split of auto 94.5 percent, walk/bike 5.2 percent, and transit 0.3 percent. With a share of just 0.3 percent, we will ignore transit in this experimental research. We do expect to implement simple heuristics to handle the non-motorized travel. If this project is successful, we expect to explore passive data sources that can inform mode splits and the associated differences in origin and destination patterns so that the model can be used in regions where mode splits are more pertinent.

Figure 1 shows a Gantt chart that outlines the schedule.

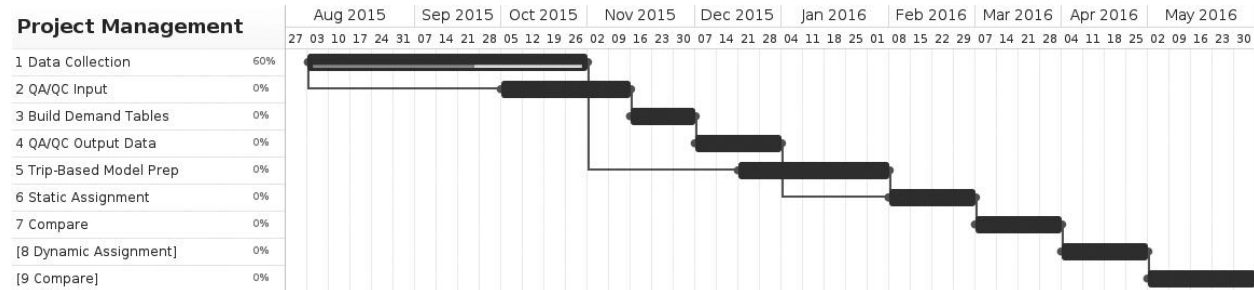


FIGURE 1 Schedule.

Gantt chart that outlines the schedule of this research.

IMPLICATIONS

For many small and medium-sized communities household travel surveys are out of reach due to the high cost of conducting them. For those communities that can afford a household travel survey, cost constraints often limit them to a small sample size resulting in data that are behaviorally rich, but pose significant challenges in the disaggregation of the data for better understanding travel markets. As a result, these communities are often left with more aggregate travel models that may not fully represent the diverse travel choices within the region. This may further impact the application of these aggregate models with respect to understanding the tradeoffs between various land use and transportation alternatives that a community may wish to evaluate. We believe the agile modeling approach developed through this research will be a good fit for small and medium-sized communities that will help them overcome the challenges of costly data collection and aggregate models that may not fully understand travel behavior. It will offer the benefit of person- and tour-based analysis without the costs and development requirements of an activity-based model. The reduced needs for data collection and universal portability of the methodology may also save these regions development costs.

STATEMENT OF INNOVATION

This study contributes to the modeling community by testing out a novel approach for modeling travel with passive data rather than survey data.

STATEMENT OF FINANCIAL INTEREST

The authors do not have any direct financial interest with regard to the publication of this research. The research was self-funded in an effort to advance practice.

REFERENCES

- [1] Roux, Sophie, and Jimmy Armoogum. Calibration Strategies to Correct Nonresponse in a National Travel Survey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2246, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 1-7.
- [2] Son, Sanghoon, Asad Khattak, Xin Wang, Paul Agnello, and Ju-Yin Chen. Quantifying Key Errors in Household Travel Surveys: Comparison of Random-Digit-Dial Survey and Address-Based Survey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2354, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 9–18.

- [3] Bricka, Stacey, Johanna Zmud, Jean Wolf, and Joel Freedman. Household Travel Surveys with GPS: An Experiment. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2105, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 51–56.
- [4] Wolf, J. Applications of New Technologies in Travel Surveys. Presented at the International Conference on Transport Survey Quality and Innovation, Costa Rica, 2004.
- [5] Wolf, J., W. Bachman, M. S. Oliveira, J. Auld, A. Mohammadian, and P. Vovsha. Applying GPS Data to Understand Travel Behavior, Volume I: Background, Methods, and Tests. Publication NCHRP Report 775. National Cooperative Highway Research Program, Washington, D.C., 2014
- [6] Gur, Y. J., S. Bekhor, C. Solomon, and L. Kheifits. Intercity Person Trip Tables for Nationwide Transportation Planning in Israel Obtained from Massive Cell Phone Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2121, 2009, pp. 145–151.
- [7] Huntsinger, L. F., and R. Donnelly. Reconciliation of Regional Travel Model and Passive Device Tracking Data. Presented at the 93rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2014.
- [8] Bachmann, C., M. J. Roorda, B. Abdulhai, and B. Moshiri. Fusing a Bluetooth Traffic Monitoring System with Loop Detector Data for Improved Freeway Traffic Speed Estimation. *Journal of Intelligent Transportation Systems*, Vol. 17, No. 2, 2013, pp. 152– 164.
- [9] Lee, Richard J., Ipek N. Sener, and James A. Mullins III. Emerging Data Collection Techniques for Travel Demand Modeling: A Literature Review. State of Texas. Strategic Research Program. <http://d2dtl5nnlpfr0r.cloudfront.net/tti.tamu.edu/documents/161407-2.pdf>