

**Innovations Deserving
Exploratory Analysis Programs**

NCHRP IDEA Program

Synthetic Household Travel Data Using Consumer and Mobile Phone Data

Final Report for
NCHRP IDEA Project 184

Prepared by:
Josephine D. Kressner, Ph.D.
Transport Foundry, LLC

March 2017

 TRANSPORTATION RESEARCH BOARD

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Innovations Deserving Exploratory Analysis (IDEA) Programs Managed by the Transportation Research Board

This IDEA project was funded by the NCHRP IDEA Program.

The TRB currently manages the following three IDEA programs:

- The NCHRP IDEA Program, which focuses on advances in the design, construction, and maintenance of highway systems, is funded by American Association of State Highway and Transportation Officials (AASHTO) as part of the National Cooperative Highway Research Program (NCHRP).
- The Safety IDEA Program currently focuses on innovative approaches for improving railroad safety or performance. The program is currently funded by the Federal Railroad Administration (FRA). The program was previously jointly funded by the Federal Motor Carrier Safety Administration (FMCSA) and the FRA.
- The Transit IDEA Program, which supports development and testing of innovative concepts and methods for advancing transit practice, is funded by the Federal Transit Administration (FTA) as part of the Transit Cooperative Research Program (TCRP).

Management of the three IDEA programs is coordinated to promote the development and testing of innovative concepts, methods, and technologies.

For information on the IDEA programs, check the IDEA website (www.trb.org/idea). For questions, contact the IDEA programs office by telephone at (202) 334-3310.

IDEA Programs
Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

The project that is the subject of this contractor-authored report was a part of the Innovations Deserving Exploratory Analysis (IDEA) Programs, which are managed by the Transportation Research Board (TRB) with the approval of the National Academies of Sciences, Engineering, and Medicine. The members of the oversight committee that monitored the project and reviewed the report were chosen for their special competencies and with regard for appropriate balance. The views expressed in this report are those of the contractor who conducted the investigation documented in this report and do not necessarily reflect those of the Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; or the sponsors of the IDEA Programs.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; and the organizations that sponsor the IDEA Programs do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the investigation.

Synthetic Household Travel Data Using Consumer and Mobile Phone Data

IDEA Program Final Report

NCHRP-184

Prepared for the IDEA Program

Transportation Research Board

The National Academies

Josephine D. Kressner, Ph.D.

Transport Foundry, LLC

15 March 2017

ACKNOWLEDGEMENTS

This work was funded by the Transportation Research Board of the National Academies under the TRB-IDEA Program (NCHRP-184). A special thanks is extended to the TRB Project Director, Inam Jawed, and the Advisory Panel: Rebekah Straub Anderson, Ohio Department of Transportation; Richard Cunard, TRB; Clint Daniels, San Diego Association of Governments; Keith Killough, Arizona Department of Transportation; Becky Knudson, Oregon Department of Transportation; David Ory, MTC; Guy Rousseau, Atlanta Regional Commission; and Elizabeth Sall, UrbanLabs LLC.

Thank you also to Suzanne Childress and Billy Charlton at Puget Sound Regional Council (PSRC) for donating their time and brains to this effort.

Appreciation also goes to AirSage, Epsilon, HERE, and StreetLight Data for their enthusiasm and support of this project.

NCHRP IDEA PROGRAM COMMITTEE

CHAIR

DUANE BRAUTIGAM
Consultant

MEMBERS

ANNE ELLIS
Arizona DOT

ALLISON HARDT
Maryland State Highway Administration

JOE HORTON
California DOT

MAGDY MIKHAIL
Texas DOT

TOMMY NANTUNG
Indiana DOT

MARTIN PIETRUCHA
Pennsylvania State University

VALERIE SHUMAN
Shuman Consulting Group LLC

L.DAVID SUITS
North American Geosynthetics Society

FHWA LIAISON

DAVID KUEHN
Federal Highway Administration

AASHTO LIAISON

KEITH PLATTE
*American Association of State Highway
Transportation Officials*

TRB LIAISON

RICHARD CUNARD
Transportation Research Board

COOPERATIVE RESEARCH PROGRAM STAFF

STEPHEN PARKER
Senior Program Officer

IDEA PROGRAMS STAFF

STEPHEN R. GODWIN
Director for Studies and Special Programs

JON M. WILLIAMS
Program Director, IDEA and Synthesis Studies

INAM JAWED
Senior Program Officer

DEMISHA WILLIAMS
Senior Program Assistant

EXPERT REVIEW PANEL

REBEKAH STRAUB ANDERSON, *Ohio DOT*

CLINT DANIELS, *San Diego Assoc. of Government*

KEITH KILLOUGH, *Arizona DOT*

BECKY KNUDSON, *Oregon DOT*

DAVID ORY, *MTC*

GUY ROUSSEAU, *Atlanta Regional Commission*

ELIZABETH SALL, *UrbanLabs LLC*

TABLE OF CONTENTS

Executive Summary	1
Concept	1
Method	1
Findings and Conclusions	2
Report Organization	2
IDEA Product	3
Concept and Innovation	3
Investigation	4
Research Outline	4
Validation	4
Validation of Time Use	5
Validation of Tours Per Day	5
Validation of Trip Distribution.....	8
Conclusions	13
Recommendations	14
Small and Medium Regions	14
Large Regions.....	16
State Departments of Transportation.....	16
Autonomous Vehicles	17
Plans for Implementation	17
Near-Term Objectives	17
Long-Term Objectives	17
References	17
List of Acronyms	18
Appendix A: Asheville, N.C.	19
Overview	19
Validation	20
Conclusions	21
Planning Applications	21
Bridge Closure.....	21
Tolling.....	21
Conclusions	21

EXECUTIVE SUMMARY

CONCEPT

Metropolitan planning organizations (MPOs) and state departments of transportation (DOTs) forecast travel behavior for their region 20-30 years into the future to inform federally mandated long-range transportation plans, transportation improvement plans, and air quality conformity. Federal funding is often funneled through these plans. To create the forecasts, household travel surveys are typically conducted. However, they have low sample sizes (<1%), low response rates (<5%), significant sampling biases due to nonresponse, large gaps between surveys (~10 years), and most importantly a prohibitively high cost compared to most other planning functions. Even more, newer forecasting methods (e.g., activity-based modeling) attempt to capture a high level of detail, thereby requiring an even more onerous survey.

Meanwhile, industry has adopted the use of population synthesis, which creates detailed household-level data for a full-sized, representative population. The approach is useful for creating people in a simulation. Collectively, the synthetic people match the aggregate distributions of demographic and socioeconomic information in U.S. Census data. Considering the wide adoption of synthetic populations in forecasting and the concerns with household travel surveys, **this research aimed to develop a method that fuses consumer data with anonymous, passive location data to create synthetic populations with individual-level synthetic travel diaries.** The travel diaries detail each person's travel and activities at locations in a timeline format.

This low cost synthetic data could have broad impacts on policy and funding decisions at the federal, state, regional, and local levels because they could give entities up-to-date, detailed data that match the population's short-term and long-term movements. By making reliable, noninvasive, statistically representative household travel data available inexpensively as frequently as desired, performance of infrastructure projects or policies could be tracked in real-time and travel demand models and forecasts could be updated systematically on demand.

METHOD

To build the data-driven synthetic population with synthetic travel diaries, two different types of third-party "big" data were combined: consumer marketing data and passive location data. Targeted marketing firms compile consumer data about individuals aged 18 years and older from a variety of sources such as public records, credit report agencies, credit card transactions, email marketing lists, and internet behavior. These types of firms, who have been operating for decades, typically sell the inexpensive, up-to-date data to companies wanting to tailor marketing campaigns to potential customers. Coincidentally, these readily available data contain most household and individual demographic and socioeconomic fields that are used in travel demand applications plus an array of more.

Mobile phone data include both wireless signaling data and global positioning system (GPS) data, which are now available from many different providers in the U.S. In the case of wireless signaling data, mobile phones act as robust mobile sensors that are aggregated into summaries of population movements. In the case of GPS data, accurate point estimates for travel times and volumes on the transportation network can be obtained. Considering that the Pew Research Center estimated over 91% of all adults in the country in 2014 owned mobile phones with no difference in likelihood of ownership between ethnic or racial groups and small differences in likelihood between age and income groups, these data together provide a comprehensive picture of travel over most modes of transportation (Pew Research Center 2014).

These two types of data, consumer and mobile phone data, separately cannot be used to replace traditional household travel surveys. In consumer data, there are large amounts of data about people, but nothing about their trip-making behavior; and in mobile phone data, there is an excellent picture of travel in any given region at any given time, but we do not know who are making the trips due to warranted privacy regulations. By combining these two types of data using statistical, simulation, and data fusion techniques, their respective shortcomings can be overcome for use in travel demand modeling applications.

The research team previously built an initial implementation of the data fusion process in the City of Atlanta in Georgia. This IDEA project proposed building a synthetic household travel dataset for a different city that had a larger study area in land and population size. The four-county planning region of the Puget Sound Regional Council in metropolitan Seattle, Washington was selected. The project aimed to test the *transferability* of the concept, the *scalability* of the method to larger study areas, and the *suitability* of the specific implementation method chosen in the initial study. The research focused on developing a process that will be consistent nationally, rapidly deployable for any size city, and systematically updateable over regular time periods.

The first stage of the research effort refactored the methodology from the initial study in Atlanta so that the same code would build a synthetic population and travel diaries in the larger metro region of Atlanta and in metro Seattle. The required “big” data were obtained for metro Seattle and the synthetic travel diaries were built. The second phase of the research effort validated the resulting synthetic travel diaries for both Atlanta and Seattle against external sources. It also checked for internal consistencies with the passive data that were fed into the synthesizing method.

FINDINGS AND CONCLUSIONS

Technical research findings were as follows:

- The method produces locally sensitive synthetic populations with individual-level travel diaries using the same code in three different metropolitan regions: Seattle, Atlanta, and Asheville.
- The validations of time use, tours per day, and geographic distribution of trips were comparable between the Seattle and Atlanta synthetic travel diaries and the household travel surveys, the activity-based models in use by the MPOs in each region, the Census Transportation Planning Package (CTPP) flow data, and the Longitudinal Employer-Household Dynamics (LEHD) data.
- The differences discovered so far in these measures appear to be reasonable considering the variability in the regional travel estimates.
- To quantify the conclusions further, the synthesizing process was treated as a travel demand model and fed into both a static assignment and MATSim model in Asheville, N.C. The resulting volumes compared well against traffic counts. See Appendix A.

Based on the implementations in Seattle, Atlanta, and Asheville and on the comments by the project's expert panel, it was concluded that:

- This method would be most useful, as it is functioning right now, in small- and medium-sized regions for planning.
- For large regions that have invested in large household travel survey collection programs and sophisticated activity-based models (ABMs), there is a desire to merge a data-driven approach into their existing activity-based models, specifically for modeling areas that ABMs currently struggle with (e.g. external trips). This is an area of future research. Otherwise, this approach could be used alongside ABMs to inform model validation through model alignment, which is also known as A/B modeling in the simulation world.
- For state departments of transportation, especially those that support modeling efforts for all the small- and medium-sized communities in their state, this data-driven approach has many benefits. Some of these benefits include lessening (or removing) the need for a statewide household travel survey program, standardization of the travel models in use without time-intensive local calibration, and standardization of analysis of projects for transportation improvement programs within the state.
- For regions interested in analyzing the impact of autonomous vehicles (AVs), this method can be combined with open-source MATSim to rapidly analyze short-term responses to AVs assuming shared fleets (i.e. "mobility as a service"), privately owned fleets, or a mix of the two. Research will be presented at the 2017 TRB Applications Conference.

REPORT ORGANIZATION

The body of this report is divided into the following sections:

- The first section, IDEA Product, describes the product that will result from the investigation in terms of new capabilities and potential impact on transportation practice.
- The second section, Concept and Innovation, highlights the principles of the innovation and describes the technical basis and uniqueness of the IDEA product for application to practice.
- The third section, Investigation, outlines the research effort conducted during this study in both Seattle, Washington and Atlanta, Georgia.
- The fourth section, Plans for Implementation, provides a description of the continuing efforts being undertaken to develop the technology.
- The main body of the report is followed by an Appendix that contains results from a related study conducted in parallel to this research in Asheville, North Carolina.

IDEA PRODUCT

The product described in this report is a data fusion system that synthesizes individual travel diaries from multiple passive “big” data sources for use in long-range and shorter-term transportation planning. A travel diary explains a timeline and the locations of a person’s travel and activities. See Table 1 for an example. An example of the raw output from the IDEA product is shown in Table 1. A random individual's travel diary is shown. The `per_id` is a unique identifier for each synthetic person. It is relatable to a person and household table with his/her demographic and socio-economic information. Each row in this table represents an activity, either Home, Work, Other, or Travel with a `start` and `end` time, shown in military time. The `place_zone` depicts the location of each rows’ activity. The other columns give information about how the activities are chained together.

This low cost synthetic data could have broad impacts on policy and funding decisions at the federal, state, regional, and local levels because the data give entities up-to-date, detailed data that match the population’s short-term and long-term movements. By making reliable, noninvasive, statistically representative household travel data available inexpensively, performance of infrastructure projects or policies can be tracked in real-time and travel demand models and forecasts can be updated systematically on demand to test future infrastructure or policy scenarios. The resources that are available for planning studies can focus on analysis entirely rather than on data collection, model building, and then condensed analysis.

TABLE 1 Example Synthetic Travel Diary from Seattle

<code>per_id</code>	<code>event</code>	<code>tour_number</code>	<code>tour_type</code>	<code>event_type</code>	<code>place_zone</code>	<code>place_number</code>	<code>start</code>	<code>end</code>
P1339021	1	0	No tour	Home	53061041704	0	0400	0658
P1339021	2	1	Composite from work	Travel		0	0658	0718
P1339021	3	1	Composite from work	Work	53061040800	1	0718	1459
P1339021	4	1	Composite from work	Travel		0	1459	1523
P1339021	5	1	Composite from work	Other	53033021903	2	1523	1527
P1339021	6	1	Composite from work	Travel		0	1527	1547
P1339021	7	0	No tour	Home	53061041704	0	1547	

Showing 1 to 7 of 7 entries

Previous 1 Next

CONCEPT AND INNOVATION

There has been increasing interest in using non-traditional data sources for travel demand modeling and planning, and this interest is motivated in part by the explosion of large, third-party data sources. Third-party data are inexpensive, prolific, and information-rich. To build the data-driven synthetic population with synthetic travel diaries, two different types of third-party “big” data are combined: consumer marketing data and passive location data. Targeted marketing firms compile consumer data about individuals aged 18 years and older from a variety of sources such as public records, credit report agencies, credit card transactions, email marketing lists, and internet behavior. These types of firms, who have been operating for decades, typically sell the inexpensive, up-to-date data to companies wanting to tailor marketing campaigns to potential customers. Coincidentally, these readily available data contain most household and individual demographic and socioeconomic fields that are used in travel demand applications plus an array of more.

Mobile phone data include both wireless signaling data and global positioning system (GPS) data, which are now available from many different providers in the U.S. In the case of wireless signaling data, mobile phones act as robust mobile sensors that are aggregated into summaries of population movements. In the case of GPS data, accurate point

estimates for travel times and volumes on the transportation network can be obtained. Considering that the Pew Research Center estimated over 91% of all adults in the country in 2014 owned mobile phones with no difference in likelihood of ownership between ethnic or racial groups and small differences in likelihood between age and income groups, these data together provide a comprehensive picture of travel over most modes of transportation (Pew Research Center 2014).

These two types of data, consumer and mobile phone data, separately cannot be used to replace traditional household travel surveys. In consumer data, there are large amounts of data about people, but nothing about their trip-making behavior; and in mobile phone data, there is an excellent picture of travel in any given region at any given time, but we do not know who are making the trips due to warranted privacy regulations. By combining these two types of data using statistical, simulation, and data fusion techniques, their respective shortcomings can be overcome for use in travel demand modeling applications.

The synthetic household travel data will be generated using a process like the conventional method for synthesizing populations in activity-based modeling, where disaggregate data, in this case from multiple consumer data firms, are used as “seeds” to create individual records. Collectively, the individual records are consistent with cross-tabulations provided by aggregate data like the U.S. Census Bureau’s Decennial Census and, in this case, aggregate mobile phone data from multiple data providers. The process will be more complex than this simplification, but the key is that the synthetic data can be produced every month, quarter, or year systematically from up-to-date consumer and mobile phone data.

In its final form, the process of creating the synthetic household travel data will operate in two steps. The first will use real household-level data from multiple consumer firms as seeds to create household-level records that are collectively consistent with tabulations and distributions provided by U.S. Census and consumer data. This synthetic population generation will provide a high level of privacy protection. The disclosure avoidance technique will operate similarly to the processes used by the U.S. Census for their American Community Survey (ACS) Public Use Microdata Sample (PUMS). The second step in the process will append synthetic travel data to each individual with realistic time schedules, origins, and destinations. Aggregate mobile phone data summarizing travel behavior at small levels of geography will be used to generate trips for each synthesized household and individual. Speed and travel time data from mobile phones will be used to estimate the length of each trip generated according to the origin and destinations. Eventually, these trip patterns from mobile phone data will be enhanced and further informed by additional data from sources like transit smart card data, General Transit Feed Specification (GTFS) data, bike share or other bike data, and pedestrian data.

INVESTIGATION

RESEARCH OUTLINE

The research team previously built an initial implementation of the data fusion process in the City of Atlanta in Georgia. This IDEA project proposed building a synthetic household travel dataset for a different city that had a larger study area in land and population size. The project aimed to test the *transferability* of the concept, the *scalability* of the method to larger study areas, and the *suitability* of the specific implementation method chosen in the initial study. The research focused on developing a process that will be consistent nationally, rapidly deployable for any size city, and systematically updateable over regular time periods.

The research effort was split into two stages. During the first stage, a study area was defined in the Seattle metropolitan region that covered the entire four-county Puget Sound region, as defined by the Puget Sound Region Council (PSRC). Also during the first stage, all input data was prepared and presented to the expert panel interactively online. During the second stage, metro Seattle synthetic travel diaries were created and then compared to other datasets for validation. Alongside the Seattle validation, an Atlanta implementation was synthesized to show that the same synthesizing process transfers between regions without local calibration. The validation effort is the focus of this report because validation is the most important to the success of the methodology.

VALIDATION

The synthetic travel diaries are compared to the input data (quality control, or internal validation) and to external data (external validation) along three dimensions: time use, trip distribution, and tours per day. The internal validations include comparisons of the synthetic travel diaries to the input passive data, namely the mobile phone location data. These validations check that the synthesizing process is behaving as expected. External validations include comparisons to data not used in the synthesizing process at all. The most recent Puget Sound Regional Council (PSRC, Seattle MPO)

and Atlanta Regional Commission (ARC, Atlanta MPO) household travel surveys, trip tables from the PSRC and ARC activity-based models, Census Transportation Planning Package (CTPP) worker flows, and Longitudinal Employer-Household Dynamics (LEHD) Longitudinal Origin-Destination Employment Statistics (LODES) data are used.

Other Validation and Application

Concurrent to the Seattle and Atlanta implementations, the same passive data synthesis process was implemented in the less complex region of Asheville, North Carolina. This was done so that the synthesis process could be externally validated against link-level volumes, which is traditionally done during calibration/validation. Additionally, two planning applications, a bridge closure and a tolling scenario, were implemented using the passive data demand model combined with MATSim (an open-source, agent-based assignment model). This comparison could be done in a more timely manner in Asheville than it could be done in Seattle or Atlanta. The next application will include modeling several automated vehicle scenarios with future population forecasts. See Appendix A for more details on Asheville.

Throughout the following comparisons, remember that no part of the synthesizing process is specific to or calibrated for one metropolitan or micropolitan region. The same code runs the Seattle, Atlanta, and Asheville applications.

Validation of Time Use

Time use area plots (Figure 1) show what activity all individuals are engaged in throughout the day. They are useful for validation because this dimension of schedules are not directly controlled for in the synthesizing process (nor in activity-based models, ABMs). Time use area plots are shown below for a collection of household travel surveys and synthetic travel diaries in Seattle and Atlanta. The comparisons against National Household Travel Survey (NHTS) data can be considered internal quality control, as the NHTS data are currently used to determine tour characteristics in the process. (Note that for both Seattle and Atlanta, only NHTS trip records are used for people who live in core-based statistical areas, CBSAs, with populations greater than 3 million. In Asheville, the NHTS data are filtered for CBSAs with populations less than 500k.) The comparisons against the other datasets are all external; none of them are used in the synthesizing process. The PSRC ABM data was not available for comparison due to restrictions on their economic data.

The synthesizing process, sometimes referred to as a simulation, captures the swing between home and work/other activities throughout the day, although it is still more rounded out than the household travel surveys (HHTS) and the ARC ABM by leaving about 10% too many people at home. The simulation does keep the right magnitude of people at home all day making no trips (shown with a white line in the area plots). The percentages are as follows: ARC ABM 18.8%, NHTS 11.8%, PSRC HHTS 11.9%, Synthetic (Atlanta) 17.2%, and Synthetic (Seattle) 16.9%. The synthesis process misses the increase of people at Home, at Other, or doing Travel around lunchtime. These time use issues are related to the way the discrete event simulation is currently structured. Several other structural approaches are being tested to address the higher number of people at Home and the missing lunchtime "buzz."

The ARC ABM plot does overestimate the number of people traveling due to the fact that ARC's flavor of ABM produces trip lists with start times that are binned into 30 minute segments and that sometimes specify a person is in more than one place at a time. Accordingly, true travel diaries cannot be assembled. For example, an individual can have two or more trips that start at 8am, some of which have travel times that overlap trips starting at 8:30am in the next time bin. In cases where multiple trips start in the same time bin, which make up a significant portion of the population, only one trip (the first occurring in the table) is kept for the time use sampling. Also and more importantly, every single trip shows up in the time use plot because they all start at the top of the hour when the time use sampling occurs.

Validation of Tours Per Day

The number of tours per day (Figure 2) is not controlled in the simulation directly either; unlike in an ABM. This plot works well for validation of the synthetic travel diaries. As was done in the time use comparisons, the synthetic travel diaries in Seattle and Atlanta are compared with a collection of household travel surveys and activity-based models.

The synthetic travel diaries in Atlanta and Seattle might have too many people who stay home all day by about 5% when compared with the household travel surveys, which was also observed in the time use area plots. Otherwise, the number of tours in a day match up well to both the travel surveys and activity-based models.

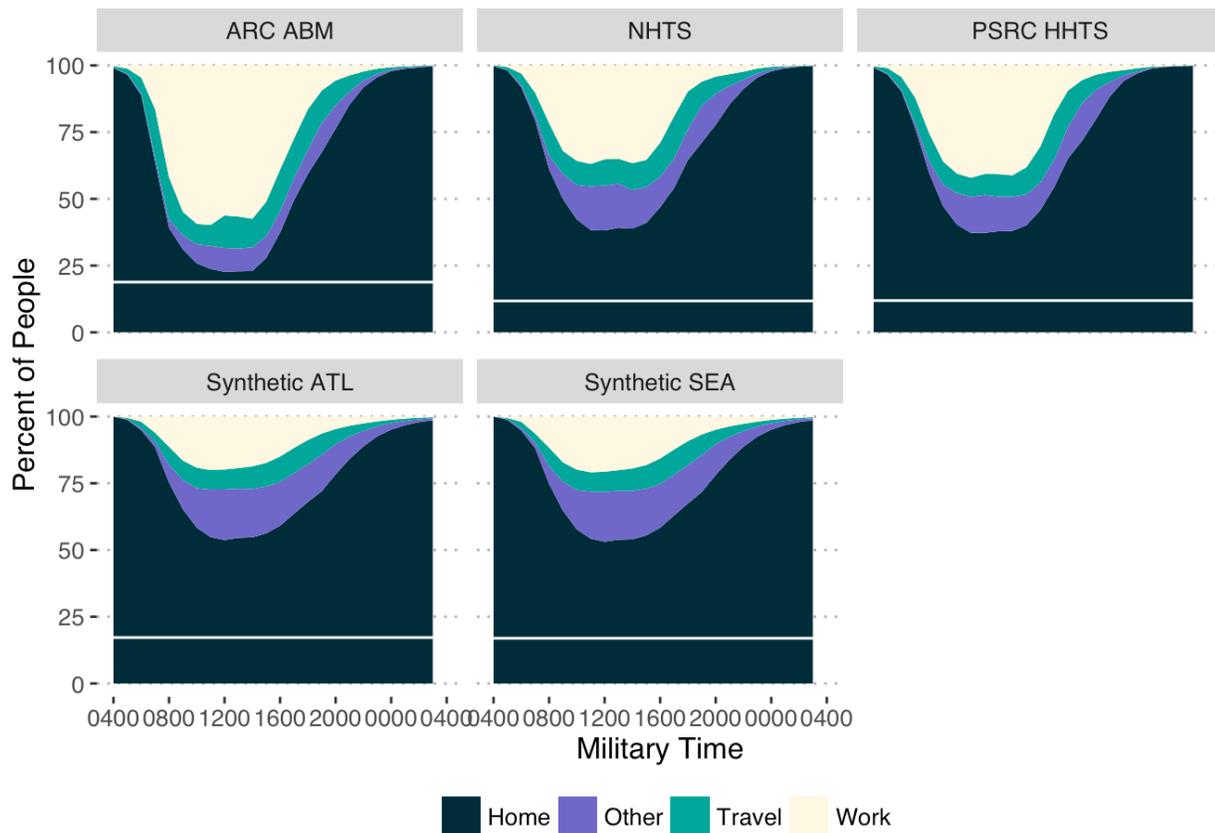


FIGURE 1 Activity by time of day as a percentage of the population for the synthetic travel diaries in Atlanta and Seattle and the comparison datasets (Atlanta Regional Commission's activity-based model, National Household Travel Survey, and Puget Sound Regional Council's household travel survey).

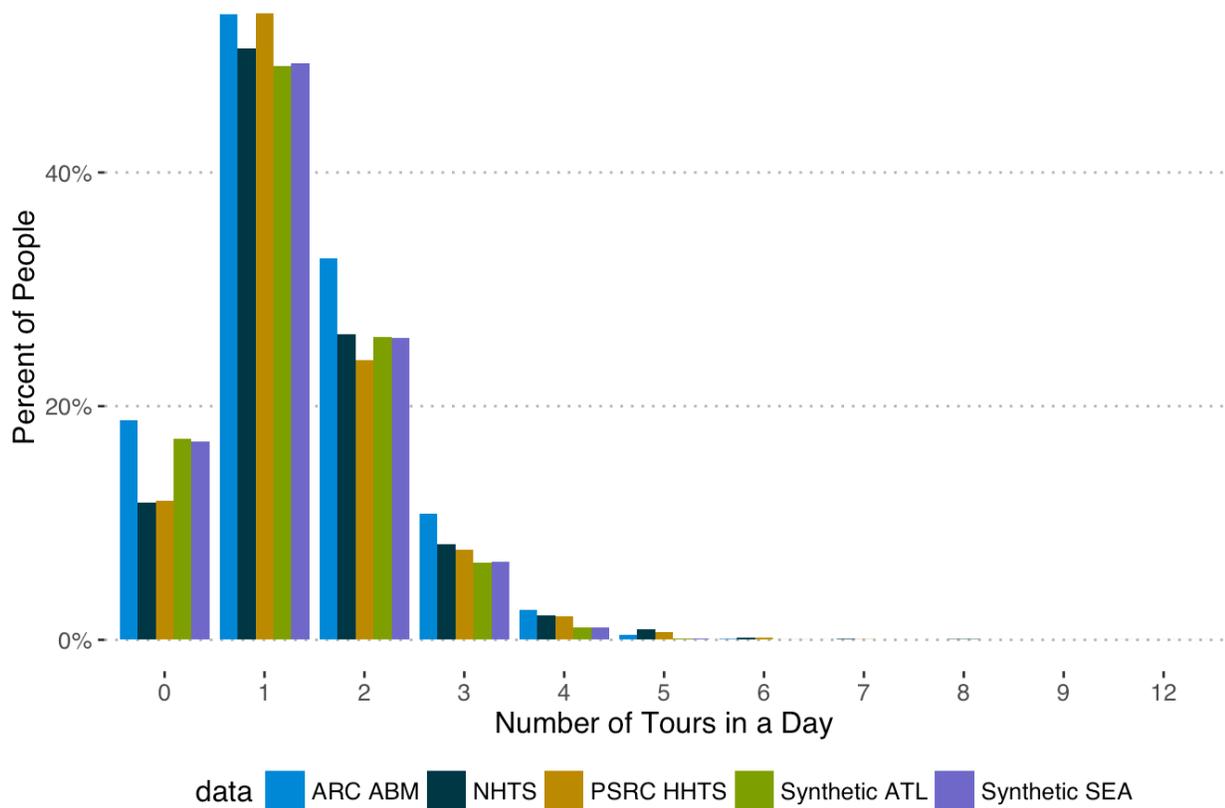


FIGURE 2 Percentage of the population by number of tours in a day for the synthetic travel diaries in Seattle and Atlanta and the comparison datasets (Atlanta Regional Commission’s activity-based model, National Household Travel Survey, and Puget Sound Regional Council’s household travel survey).

Validation of Trip Distribution

Introduction to "Flow Analyzer"

The geographic distribution of trips is modeled directly in the simulation using an origin-destination (OD) matrix from passive location data, which comes from mobile phones. An OD matrix summarizes, for a given time of day and sometimes for a specific trip purpose, the portion of trips going to and from an exhaustive, mutually exclusive set of polygons that cover a geographic area. The distribution of trips in a region is vital to the usefulness of a travel model that transfers to different regions. It therefore was included in the validation efforts. To compare OD matrices at different geographic levels, an interactive "chord diagram" was plotted beside an interactive map. This type of visualization will be referred to as a flow analyzer. For an interactive introduction to interpreting the flow analyzer, visit <http://forge.transportfoundry.com/tour>.

In general, a chord diagram can be thought of as a pie chart that is specific to matrix data. In the case of an origin-destination matrix, it allows one to visually identify the origin and destination of each pie piece in the pie chart without labeling each piece. The data are arranged radially around a circle with the relationships between the origin and destination of each element of the matrix drawn as chords that connect reciprocal pie pieces. For example, if 20% of all trips go from Place A to Place B, and 10% go from Place B to Place A (the reciprocal), a chord would connect the two "pie pieces" together through the middle of the circle. On one side of the chord, the outer edge would proportionally take up 20% of the pie, and on the other it would take up 10%. The outer arcs, where the ends of the chords are proportionally 20% or 10% like a pie chart, are labeled with text as "A" or "B". The chords (or connected pie pieces) are grouped by origin. As such, the end of chord that takes up 20% of the outer arc would have an "A" label, and the end that makes up 10% would have a "B" label. When one sees a chord with a 20%-sized chord edge by "A" and a 10%-sized chord edge by "B", it is obvious that 20% of the total trips originate in Place A and go to Place B and 10% go in the other direction. In total, this origin and destination pair have 30% of all trips.

The flow analyzer also has a map in the right panel that responds to mouse movements and is tied to the chord diagram on the left. The map highlights an origin using a wider outline. When an origin is outlined, a proportional bubble appears over each of the destinations according to the matrix. This shows the geographic distribution of the flow of people out of the origin.

The flow analyzer is used to compare trip tables and home-work matrices at four geographic levels: county level (4 counties in Seattle, 14 counties in Atlanta); a larger aggregate zone level (~25 zones each), a smaller aggregate zone level (~100 zones each), and the census tract level (776 tracts in Seattle, 858 tracts in Atlanta). The two aggregate zone levels are census tracts clustered by driving distance between population center points using a k-medoids algorithm. The clustering was restricted within county boundaries for the 25-zone clustering but was not for the 100-zone clustering. At each of these levels, comparisons for both internal validations and external validations were made. In the next few sections, a few themes and conclusions are highlighted from these comparisons.

All of the comparisons are available at http://projects.transportfoundry.com/trbidea/report2.html#all_comparisons. In all the flow analyzers, the estimates were converted to percentages to provide consistency between geography levels and were rounded to two decimal places to speed up loading time (except in the case of the tract level where three decimal places were used to capture the smaller flow estimates). The tract level comparisons can be slow depending on your browser, internet speed, and browser version. As the zones become smaller, the chord diagrams become less meaningful, but in these cases the map interactions are still compelling.

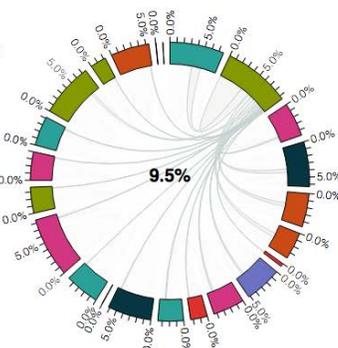
The validation of trip distribution is split into internal validations that check the internal consistency of the resulting synthetic travel diaries with the OD matrix that was input into the data fusion from mobile phones. The external validations compare the synthetic travel diaries with data sets that were not used at all in the synthesis process.

Internal Validations

Synthetic Travel Diaries vs Input Origin-Destination Matrix

The flow analyzer screenshot in Figure 3 shows the trip table from the synthetic travel diaries in the top row and the input origin-destination (OD) matrix in the bottom row using the 25-zone system. Overall, the chord diagrams show very similar flow patterns. The biggest differences in Seattle relate to the polygons closest to downtown Seattle. These zones have a larger share of trip origins in the synthetic travel diaries than they do in the input OD matrix. For example, mousing over the downtown Seattle polygon, one sees that 9.5% of all trips in the synthetic travel diaries start in this zone whereas 9.0% do in the input OD matrix. Due to the fact that the passive data OD matrix is used in the synthesizing

Synthetic



Input OD Matrix

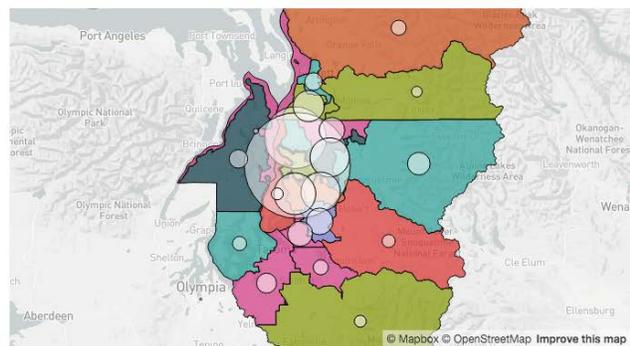
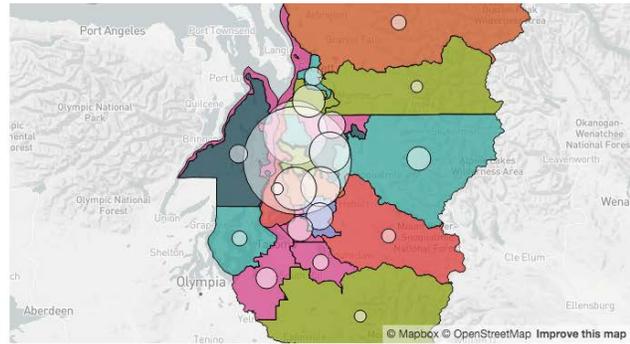
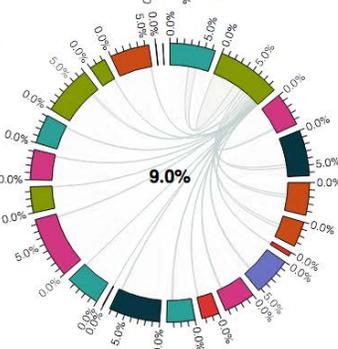


FIGURE 3 Screenshot of the flow analysis of Seattle between the synthetic travel diaries and the OD matrix that was input into the synthesizing process. The geographies are counties, and the map scale is the same as Figure 4.

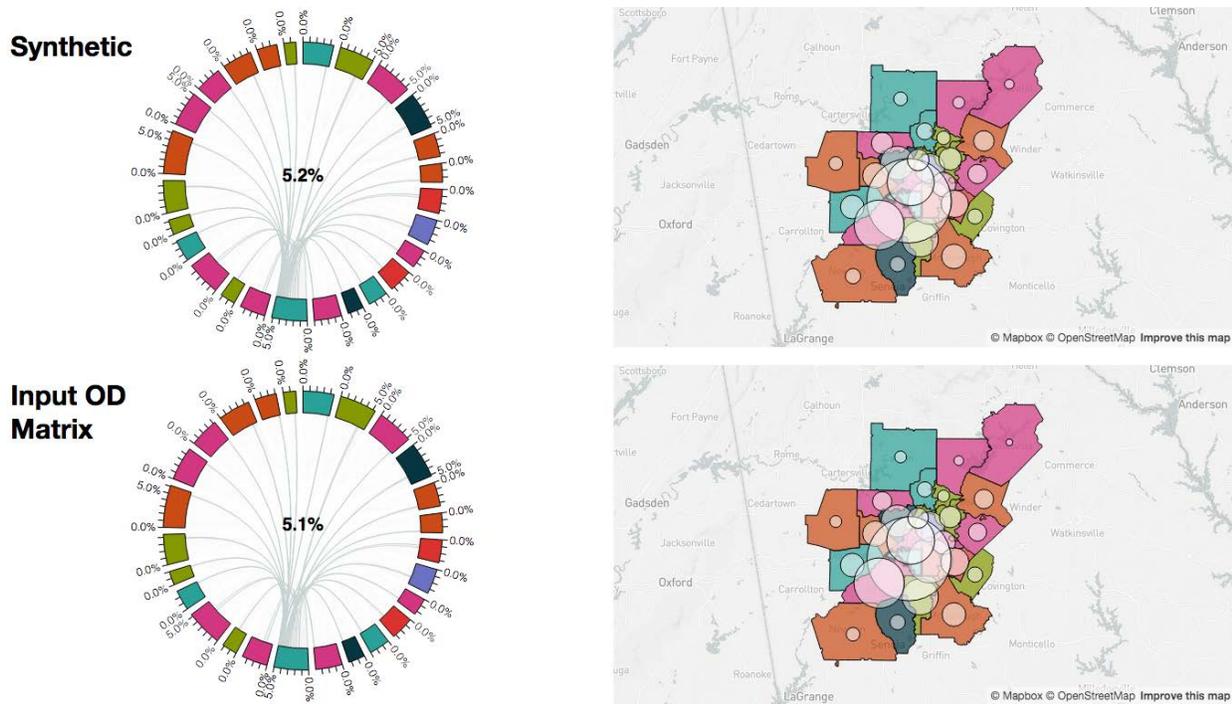


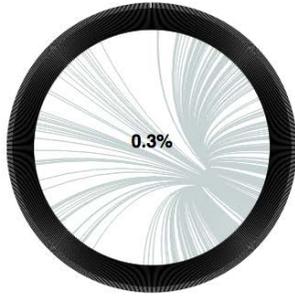
FIGURE 4 Screenshot of the flow analysis of Atlanta between the synthetic travel diaries and the OD matrix that was input into the synthesizing process. The geographies are counties, and the map scale is the same as Figure 3.

to *distribute* trips like a typical destination choice model would and not to *generate* trips, the process adjusts the scale of each origin vector according to the synthetic people living within each zone. Said another way, the synthetic people generate trips, not the OD matrix. Figure 4 shows a screenshot of the same comparison for metro Atlanta. In downtown Atlanta, the difference is less prevalent (5.2% versus 5.1%). The following image shows the flow analyzer at the same zoom level as Seattle.

Looking closer in Figure 5 in Seattle, one can see that the overall differences are subtler at the census tract level. Note that the input OD matrix was purchased from the data provider at the tract level in Seattle and at the block group level in Atlanta.) These subtle differences are due to the above reasoning as well as the additive smoothing implemented in the synthesizing process. Stated simply, additive smoothing is a technique that assigns very small probabilities to unobserved categorical data due to the fact that 100% of all trips were not observed in the mobile phone data. The same flow analyzer for tract-level in Atlanta shows similar results but is not included here.

In general, the internal comparisons show that the synthesizing process is distributing trips as intended. In the future, these comparisons need to be done for each trip purpose and time of day category.

Synthetic



Input OD Matrix

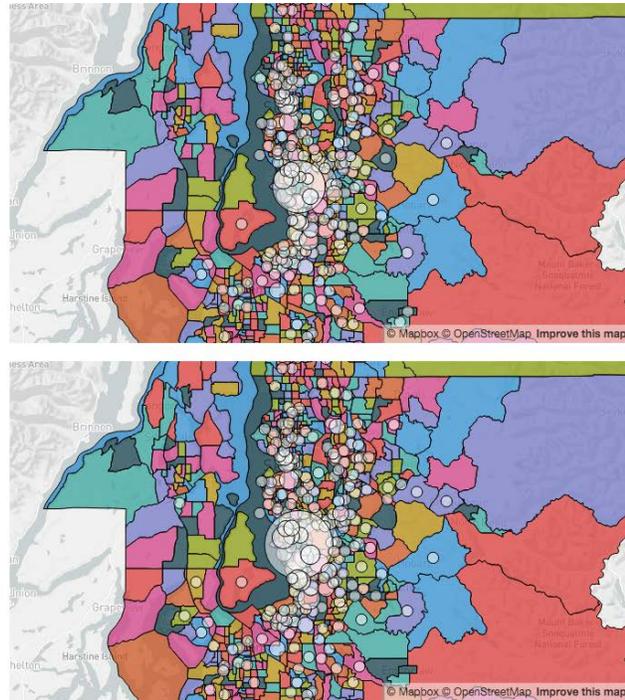
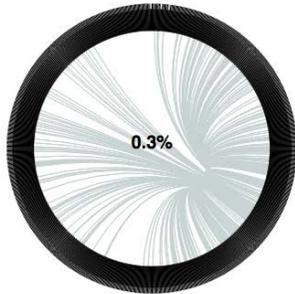


FIGURE 5 Screenshot of the flow analysis of Seattle between the synthetic travel diaries and the OD matrix that was input into the synthesizing process. The geographies are census tracts.

External Validation

Synthetic Travel Diaries vs PSRC/ARC Activity Based Model

Looking at the comparison of the synthetic travel diaries to PSRC'S ABM which is not used in the synthesizing process, greater differences are present than are in the internal validation, as expected. See Figure 6. In the PSRC ABM, 10.9% of all trips start in the downtown Seattle polygon rather than 9.5%. Despite this difference, the overall characteristic or distribution of trips are very similar at this zone level.

In both cities, a greater number of trips occur internally within zones in the PSRC and ARC ABM models than in the synthetic travel diaries. This is most easily seen in the 25-zone system: Seattle and Atlanta. This could be due to an over correction on the part of the data providers internally. Historically, intrazonal trips in these matrices were too high. Especially in the case of triangulated signal location data, it can be difficult to identify true movement when the geographic accuracy has higher tolerances than the trip distance itself. GPS-based location data could be used to smooth out these intrazonal trip estimates, but this has not yet been fully tested.

Other Observations

In Atlanta, the input OD matrix was obtained about 2 years ago without a "buffer" zone around the study area (due to budget and contract constraints). As a result, one can see in Figure 7 that the zones bordering the study area with interstate highways passing through them have external-external, external-internal, and internal-external trips mixed into the probabilities. This can be easily corrected with a refresh of the OD matrix using a different zone structure. In fact, these types of external trips can be captured and modeled explicitly within the synthesizing process.

In Seattle in Figure 8, the zone covering the air force base has no originating trips in the ABM. With the passive data, there are trips. The ABM mostly likely relies on ACS PUMS data to synthesize its population, and the ACS PUMS data have separate estimates for group quarters.

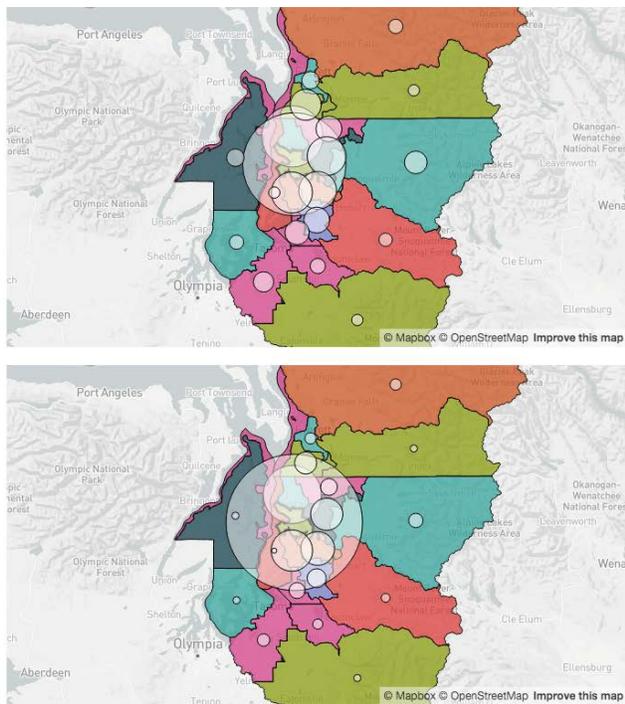
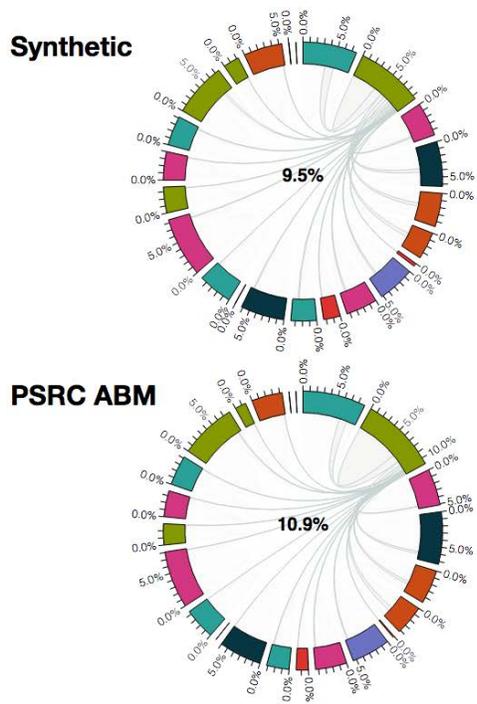


FIGURE 6 Screenshot of the flow analysis of Seattle between the synthetic travel diaries and the regional MPO’s (PSRC) activity based model (ABM). The geographies are clustered census tracts.

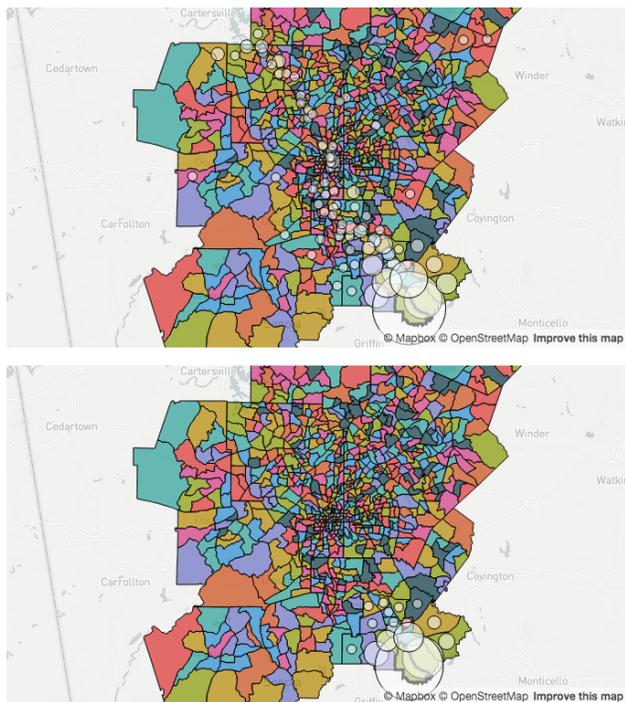
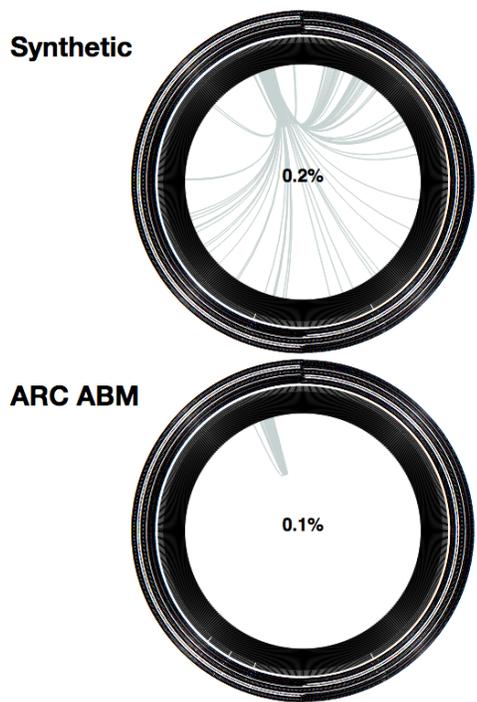


FIGURE 7 Screenshot of the flow analysis of Atlanta between the synthetic travel diaries and the regional MPO’s (ARC) activity based model (ABM). Highlights the external trips traveling along the interstate.

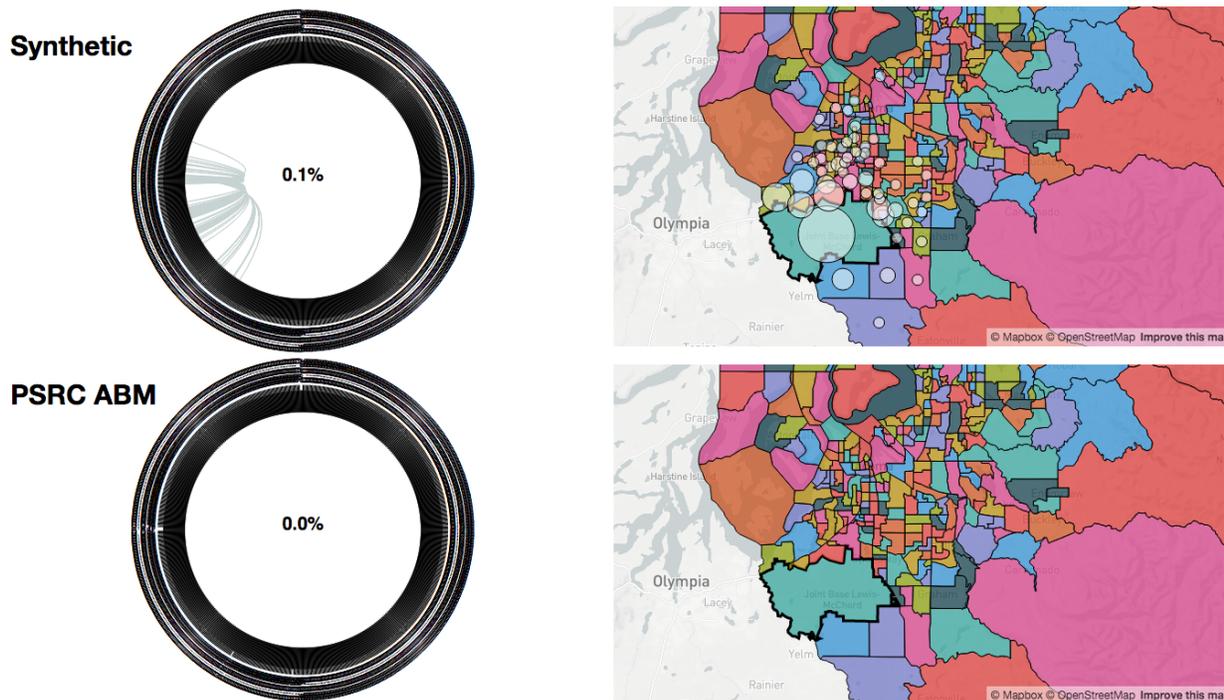


FIGURE 8 Screenshot of the flow analysis of Seattle between the synthetic travel diaries and the regional MPO’s (PSRC) activity based model (ABM). Highlights the air force base, a group quarter not present in many US Census datasets.

Synthetic Home-Work Trips vs Census Transportation Planning Package

Census Transportation Planning Package (CTPP) flows were not used within the synthesizing process. The flow patterns of the synthetic homes and work locations match very well to CTPP flows. See Figure 9. This comparison could indicate that the input OD matrix were scaled by the data providers to CTPP data at some geographic level. For example, hovering over downtown Seattle where the highest percentage of jobs are in the region, one can see that the distribution of commuters are very similar between the synthetic travel diaries and the CTPP flows (17.3% vs 19.0%, respectively). Note that the flow analyzer is organized with work locations as the origins and home locations as the destination. This approach works well for investigating distributions of home locations when hovering over a business district. At the lower level of 100 zones, the patterns are still very similar. At the lowest level of census tracts, greater variability is present between the two datasets.

Synthetic HW vs Longitudinal Employer-Household Dynamics

Looking at the same comparison as the one above with Longitudinal Employer-Household Dynamics (LEHD) flows in Figure 10, there are greater differences. The large commute patterns still are identifiable though. The fact that the CTPP consistently matches the synthetic travel diaries better than the LEHD also indicates that the CTPP may have been used internally by the data provider(s).

CONCLUSIONS

The validations of time use, tours per day, and geographic distribution of trips in this report were comparable between the Seattle and Atlanta synthetic travel diaries and the household travel surveys, the activity-based models, the CTPP flow data, and the LEHD data. The differences discovered so far in these measures appear to be reasonable considering the variability in the regional travel estimates. To quantify the conclusions further, the synthesizing process was treated as a

demand model and fed into both a static assignment and MATSim model in Asheville, N.C. The resulting volumes were compared against traffic counts. See Appendix A.

RECOMMENDATIONS

Small and Medium Regions

Based on the implementations in Seattle, Atlanta, and Asheville and on the comments by the project's expert panel, the method explored during this TRB IDEA project for producing synthetic travel diaries from passive "big" data would be most useful in small- and medium-sized regions. Often in these regions, local household travel surveys are out of reach due to the high cost of conducting them. For those regions that can afford a local household travel survey, cost constraints often limit them to a small sample size that results in data that are behaviorally rich, but pose significant challenges when disaggregating the data for understanding travel markets. In these cases, the regions are often left with more aggregate travel models that may not fully represent the diverse travel behavior within the region. This can impact the understanding of tradeoffs between various land use and transportation alternatives that a region may wish to evaluate.

The passive data simulation explored in this project can be used in combination with an existing regional assignment model (or an open-source implementation of MATSim) as a data-driven travel model for planning applications, as was done in Asheville. This data-driven model permits a tour-based, microsimulation model to be developed that can capture diverse travel behaviors with substantially less investment in time, money, and data. In other words, it will offer the benefit of person- and tour-based analysis without the costs and development requirements of an activity-based model. The universal applicability of the model as well as the reduced needs for manual data collection can also save these regions in development costs.

Planning

Most importantly, the data-driven model can be used as an effective scenario analysis and forecasting tool to examine common policy scenarios for these sized communities, including road closures or expansions, tolling, and effects of population growth or decline. These types of scenarios can be applied now in the existing form of the model. For straightforward scenario analysis involving changes to the network, the assignment model and its network can be altered

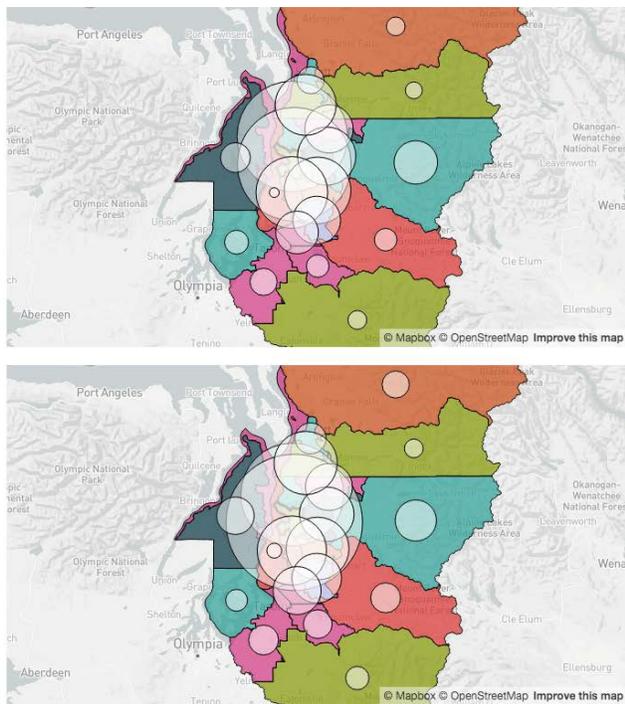
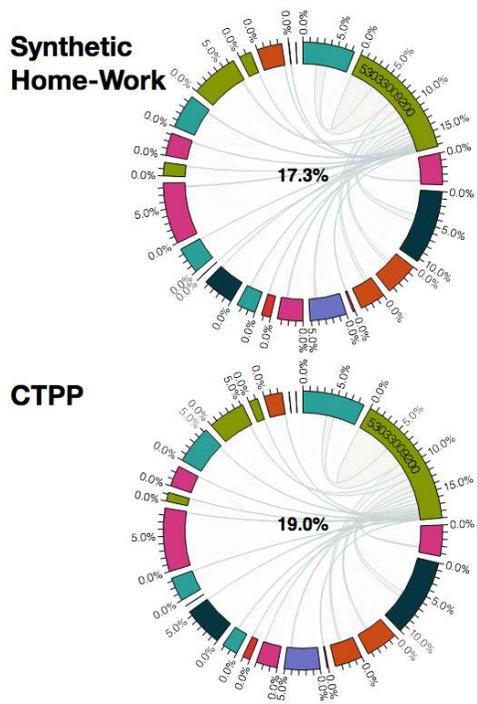


FIGURE 9 Screenshot of the flow analysis of Seattle between the home-work locations in the synthetic travel diaries and the Census Transportation Planning Package (CTPP) data in the same area. Organized with work locations as the origins and home locations as the destination. This approach works well for investigating distributions of home locations when hovering over a business district.

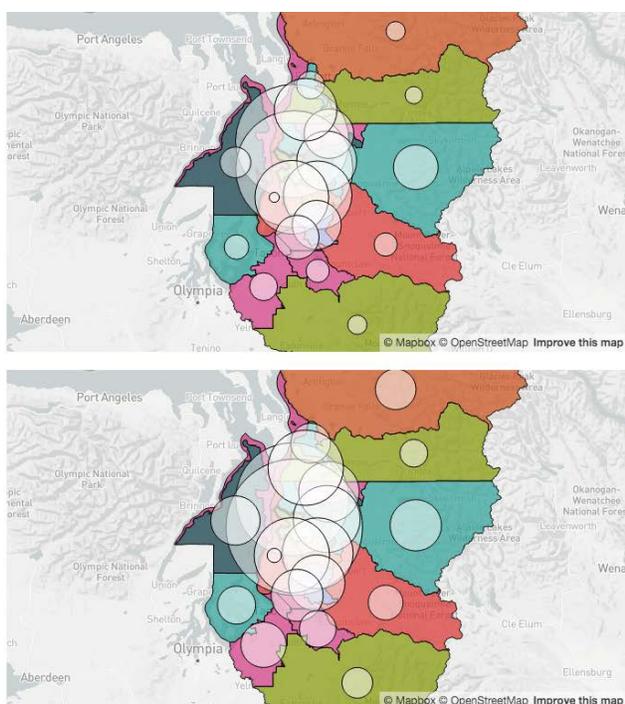
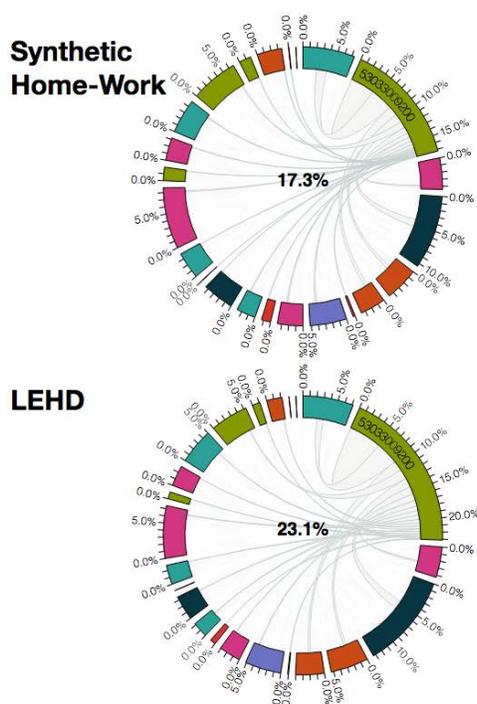


FIGURE 10 Screenshot of the flow analysis of Seattle between the home-work locations in the synthetic travel diaries and the Longitudinal Employer-Household Dynamics (LEHD) data in the same area. Organized with work locations as the origins and home locations as the destination. This approach works well for investigating distributions of home locations when hovering over a business district.

and the model rerun to examine differences. This operates the same way a four-step or activity-based model does. For forecasting work, the base-year population can be altered to match a future forecast in the same way it is done with a four-step or activity-based model, and the base-year model can be rerun with the updated population. As an added benefit though, with the lower costs and time requirements of updating this data-driven model, communities would also be able to update the model much more frequently than current practice allows. This could better inform population forecasts and also provide opportunities for forecasting behavioral changes more directly.

Because the data-driven model is not calibrated with local adjustment factors, the input behavioral data can be tweaked in a straightforward manner. For instance, analysts working with state-of-the-practice models cannot adjust parameters that would create more trip chaining behavior without upsetting the model calibration results in other interrelated but non-obvious steps of the model. In the data-driven microsimulation model by contrast, the user can directly increase the number of activities on a tour and the simulated individuals' daily patterns will adjust, leaving activity durations and paths between them relatively unchanged. Similar conclusions can be drawn with changes in the rate of work trips versus non-work trips or changes in commute departure times. Due to the tight relationship between model input and model output, it may also be easier to capture *uncertainty* in all of these.

Large Regions

The project's expert panel consisted of several panelists from metropolitan planning organizations that have invested in large household travel survey collection programs and sophisticated activity-based models. For these types of regions, there is a desire to merge a data-driven approach into their existing activity-based models, specifically for modeling areas that ABMs currently struggle with (e.g. external trips). This may be possible in the future.

A/B Modeling

With a desire to focus on the now though, these regions can consider using this type of data-driven model as a method for comprehensive scenario analysis validation. The availability of passive "big" data, from AirSage origin-destination patterns to INRIX travel time data to billions of crowd-sourced route histories in Waze, has made it possible to more comprehensively validate an ABM than was possible in the past. However, adding one or more of these data to an existing ABM can be tricky and cumbersome, often introducing as much noise as it eliminates, as specification errors or effects of constants in one part of the model propagate to other parts.

Rather, these data can be used in a different way to inform model validation through model alignment, which is also known as A/B modeling in the simulation world. In A/B modeling, two teams start from the same data and problem statement, but create separate models. The two teams then come back together to compare approaches and outcomes. The results are often different, of course. But the most important insights are gained by understanding why the different approaches reached different outcomes. The insights can also build confidence in the models, build confidence in the policy analyses carried out using the models, provide insight into additional data and methods required for robust analyses within the problem space, and provide guidance on which way to best "fix" an existing modeling system. The ability to mine the complete patterns and interactions from the data-driven model can add significant information to the ABM, far in excess of simply adding additional data to it. The ability to look at differences between the synthetic and actual travel diaries, for example, can provide important insights that are simply not possible to obtain when looking at either in isolation. Asking why the differences occur can provide the insight needed to improve planning efforts and each individual model. Because the cost of the data-driven model is a small fraction of the cost of an ABM, this A/B modeling approach is now feasible.

State Departments of Transportation

For the state departments of transportation that support modeling efforts for all the small- and medium-sized communities in their state, this data-driven approach has additional benefits. Adopting this approach statewide would lessen (or remove) the need for a statewide household travel survey program, it would standardize the travel models in use without time-intensive local calibration, it could standardize analysis of projects and prioritization for transportation improvement programs within the state, it could provide performance measures while resolving issues around data and target consistency, and it would lower overall passive data costs through the bulk purchase of passive data. It would also enable an examination of the potential differences in autonomous vehicle scenarios between different sized regions within the state.

Autonomous Vehicles

According to U.S. DOT Secretary Foxx, autonomous vehicles (AVs) have “enormous potential to save lives, reduce greenhouse gas emissions, and transform mobility for the American people.” The current government and large industry leaders are investing in policies and technologies to make automated vehicles a reality. Even though AVs seem to be right around the corner, few four-step or activity-based travel models are able to study the effects of AVs. The Asheville data-driven model will be used to analyze short-term responses to AVs assuming shared fleets (i.e. "mobility as a service"), privately owned fleets, or a mix of the two. Although AVs present an array of questions-- notably the interaction of AVs with transit and the changes to trip-making behavior-- the first study, which will be available in mid-2017, will focus on measuring the sensitivity of total vehicle miles traveled (VMT) and average commute time to differing assumptions of AV adoption and ownership. The ability to run a lot of different scenarios is made possible by the data-driven model's fast runtimes and tight relationship between model input and model output. This will allow for a *pattern-based* analysis, where conclusions can be drawn about the likelihood of different outcomes based on the stability of an outcome through different scenarios.

PLANS FOR IMPLEMENTATION

The research team is proceeding with the development of this technology. Future implementation plans are divided into near- and longer-term objectives. Near-term objectives represent plans extending approximately 12–18 months into the future. Longer-term objectives represent activities beyond this near-term period.

NEAR-TERM OBJECTIVES

In the near-term, the objective is to build a user-interface that enables transportation planners and modelers to interact with and test the data fusion methodology. It is the team's intention to build a user-interface that allows for rapid testing of a variety of infrastructure projects and policy scenarios that small- and medium-sized regions deal with on a regular basis such as road closures/expansions, tolling, and accessibility questions specific to certain population segments. The research team anticipates providing this user interface as a web-based application that alleviates small- and medium-sized planning organizations from the computing and data storage requirements of this methodology. Our aim is to make it a plug and play application.

The other near-term objective is to engage a small group of MPO and DOT users that find value in this innovative, forward-thinking technology. Through this small initial group, the research team hopes to refine the functionality and its value-add to the planning process.

LONG-TERM OBJECTIVES

In the long-term, the research team hopes to make lasting change to the current state of practice in transportation planning. By building a methodology that can truly be carried between different planning agencies, a community of innovation can develop among planners and travel modelers. It is our aim to cultivate this community so that public agencies can continue to be relevant in today's changing transportation industry.

REFERENCES

- Greaves, Stephen P., and Peter R. Stopher. 2000. “Creating a Synthetic Household Travel and Activity Survey: Rationale and Feasibility Analysis.” *Transportation Research Record: Journal of the Transportation Research Board* 1706: 82–91.
- Pew Research Center. 2014. “Cell Phone Ownership Hits 91% of Adults.” <http://www.pewresearch.org/fact-tank/2013/06/06/cell-phone-ownership-hits-91-of-adults/>.
- PTV NuStats. 2011. “Regional Travel Survey: Final Report.” Atlanta Regional Commission.

LIST OF ACRONYMS

ABM	activity-based model
ACS	American Community Survey, a US Census data product
ARC	Atlanta Regional Commission, the metropolitan planning organization for Atlanta, Georgia
AV	autonomous vehicle
CTPP	Census Transportation Planning Package, a tabulation of the US Census' American Community Survey
DOT	department of transportation
GPS	global positioning system
GTFS	General Transit Feed Specification
HHTS	household travel survey
LEHD	Longitudinal Employer-Household Dynamics, a US Census data product
MPO	metropolitan planning organization
NHTS	National Household Travel Survey
OD	origin-destination, typical used as OD matrix
PSRC	Puget Sound Regional Council, the metropolitan planning organization for Seattle, Washington
PUMS	Public Use Microdata Sample, a tabulation of the US Census' American Community Survey
VMT	vehicle miles traveled

APPENDIX A: ASHEVILLE, N.C.

OVERVIEW

Concurrent to the TRB IDEA implementations in Seattle and Atlanta, the same passive data synthesis process was implemented in the smaller region of Asheville, North Carolina. This was done so that the synthesis process could be externally validated against volumes as is done traditionally during calibration/validation. Additionally, two planning applications were implemented using the passive data model and MATSim combination: a bridge closure and a tolling scenario. This comparison could be done in a more timely manner in Asheville than it could be done in Seattle. The next application will include modeling several automated vehicle scenarios with future population forecasts.

The flowchart in Figure 11 shows diagrammatically how the synthesis process, which is referred to as a discrete-event simulation (DES) here, compares to the four-step model that Asheville metropolitan region currently uses. For context, an activity-based model (ABM) process is shown too. Both the four-step and activity-based models essentially use the same type of inputs. In the case of ABMs, the inputs are more detailed. Although ABMs estimate more sophisticated behavioral dynamics than trip-based models, they both flow through sequential steps, where matrices of information are passed between the steps. In either case, the demand model outputs get summed up and feed into an assignment model, which can be either a static-assignment or a dynamic one.¹ The assignment model loops back to the demand model until equilibrium is met (not shown in the flowchart).

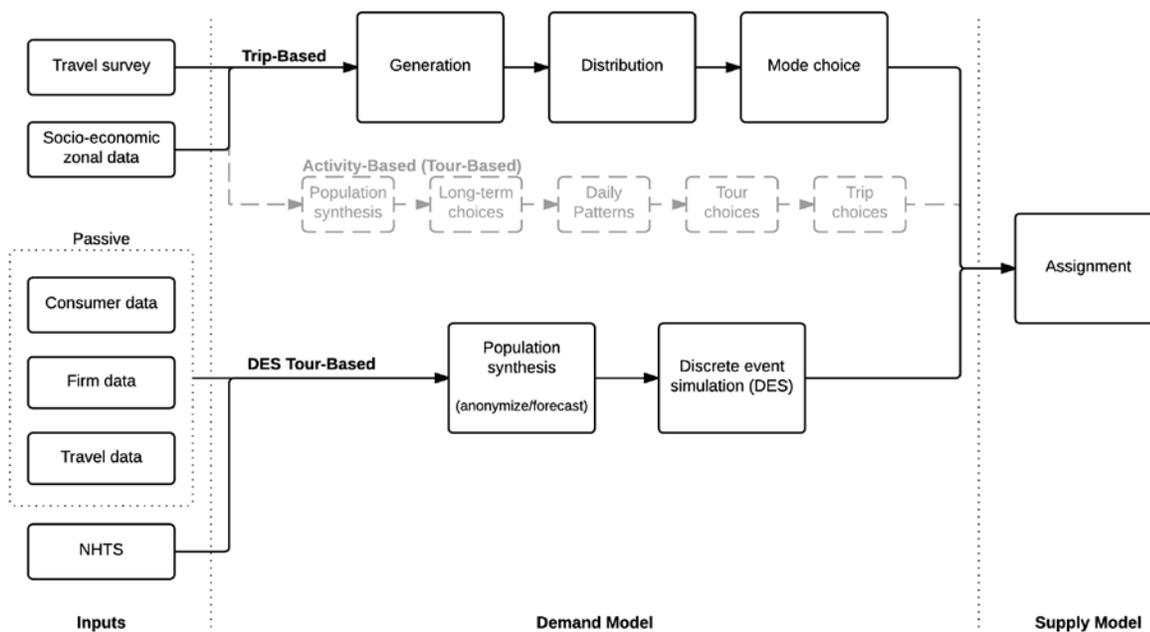


FIGURE 11 Comparison of trip-based (four-step), activity-based, and discrete event simulation tour-based (data-driven) models.

¹ ARC is working on integrating their ABM with a dynamic traffic assignment model.

TABLE 2 Example Data from Data-Driven Demand Model

per_id	event	tour_number	tour_type	event_type	place_id	place_number	start	end
H00001R06P02	1	0	No tour	Home	H00001R06	0	0400	0944
H00001R06P02	2	1	Composite from work	Travel		0	0944	1018
H00001R06P02	3	1	Composite from work	Work	F18373	1	1018	1228
H00001R06P02	4	1	Composite from work	Travel		0	1228	1257
H00001R06P02	5	1	Composite from work	Other	F17134	2	1257	1345
H00001R06P02	6	1	Composite from work	Travel		0	1345	1430
H00001R06P02	7	0	No tour	Home	H00001R06	0	1430	1809
H00001R06P02	8	2	Simple non-work	Travel		0	1809	1809
H00001R06P02	9	2	Simple non-work	Other	F28315	1	1809	1815
H00001R06P02	10	2	Simple non-work	Travel		0	1815	1816

Showing 1 to 10 of 11 entries

Previous 1 2 Next

In the case of this DES approach, a different set of inputs are used. The inputs rely on passively-collected "big" data. In the demand model, person and firm populations are synthesized with consumer and firm data as the seeds and, in a similar way to population synthesis for ABMs, is controlled with U.S. Census data totals. Then rather than sequentially estimating large matrices for each step, a DES is used. DESs are a way to model the behavior of a complex system as ordered events along a timeline, where each event occurs at a particular instant in time and marks a change of state in the system. For example, a change in the state of the system might be leaving one's house or arriving at one's work. The DES is trained using tour patterns from NHTS, and in the future will be trained with other passive data to reflect up-to-date and local changes in tour behavior. In other words, the DES models the same aspects of behavior you find in four-step or ABMs, but with a unified simulation rather than sequential steps.² It then passes the demand trip tables into an assignment model.

An example of the trip table that passes from the DES to the assignment model is shown in Table 3 with one random individual's travel diary from Asheville. This table is relatable to person and household sociodemographic data via `per_id` and to firm data via `place_id`. The `start` and `end` columns are shown in military time.

In Asheville, the DES was fed first into the existing static-assignment model used in the region, which was built in TransCAD. Following that, the DES was fed into MATSim, an agent-based assignment model (similar to a discrete event simulation). No feedback loops were used in either case, although MATSim does have a replanning phase within the model itself that maximizes utility for each individuals' schedule. In the future, one could imagine a single agent-based model that would do both demand and assignment in "planning" and "real-time" layers.

VALIDATION

The passive data demand model with both the static-assignment and agent-based assignment models were compared against traffic counts. Alongside, the existing regional four-step model with the static-assignment model are shown. Note that the passive data model currently only handles internal trips to date; demand tables for commercial vehicles, visitors, and external trips were carried over directly from the four-step model. Note that the existing four-step/static-assignment

² Mode choice is being tested in both the DES and in MATSim.

model is calibrated specifically to the local traffic counts whereas the two passive data implementations are not calibrated to Asheville.

Conclusions

The passive data model produced effectively equivalent levels of accuracy when compared with the current four-step/static model. Figure 12 and Table 3 summarize the results. In the left column of the scatterplot, the reference four-step model validation is depicted. In the right two columns, the DES models are depicted. The top row illustrates the count scatterplots, and the bottom row illustrates the maximum desirable deviation tolerances. The points are color coded by facility type. Note that 41 and 47 links in the static-assignment models, respectively had percent error greater than 200%. However, these were all on Other Major or Minor Thoroughfares with AWDT volumes under 3,000 (shown with a vertical dashed line). For visual clarity, they are not shown.

The largest deviation from the maximum desirable deviation tolerances are with expressways in the DES+MATSim model. In this case, the resident data only are fed into MATSim. When agents figure out their plans, they do not see any other traffic from commercial, external, or visitor trips. They will fill up the expressways first. Post MATSim, the commercial, external, and visitor trips from the TransCAD assignment are added to the MATSim network results. Because of this simplification, the expressway assignment is expected to be high.

PLANNING APPLICATIONS

In the following planning applications in Asheville, the synthetic travel diaries (also referred to here as DES) and MATSim are used.

Bridge Closure

In this application, a bridge link on Amboy Road was deleted in both directions from the network. Figure 13 shows the location of the bridge in relation to downtown Asheville. The network difference plot in Figure 14 shows the effect of the bridge closure. Red/yellow mean fewer trips (where full red means a loss of almost all the links' volume), blue means more, and grey means unchanged. The difference is relative to the base scenario with the bridge open (i.e., $(V_{noblidge} - V_{base})/V_{base}$).

Tolling

In this tolling scenario, a toll of 10 cents per mile from 7am to 9am and 4pm to 6pm was implemented along the I-40/I-240 loop surrounding central Asheville. The relative network difference plots in Figures 15 and 16 show the network just before the morning toll begins (around 6:40am) and just after (around 7:00am), respectively. Again, yellow indicates a decrease in volume whereas red would indicate a change to zero volume.

Conclusions

In both applications, the model responded as the team expected in terms of volume changes and diversion. In addition to being able to query changes in traffic volumes on specific links as one can do with the existing four-step model, the synthetic travel diaries and MATSim combination can also break down the people whose behavior changed. For example, one can explore the income distribution of those who choose to pay the toll and those who did not in the tolling scenario. This kind of analysis is not typically built into a four-step model with a static assignment. Other similar advantages can be found in a variety of road closure/expansion scenarios such as the bridge closure.

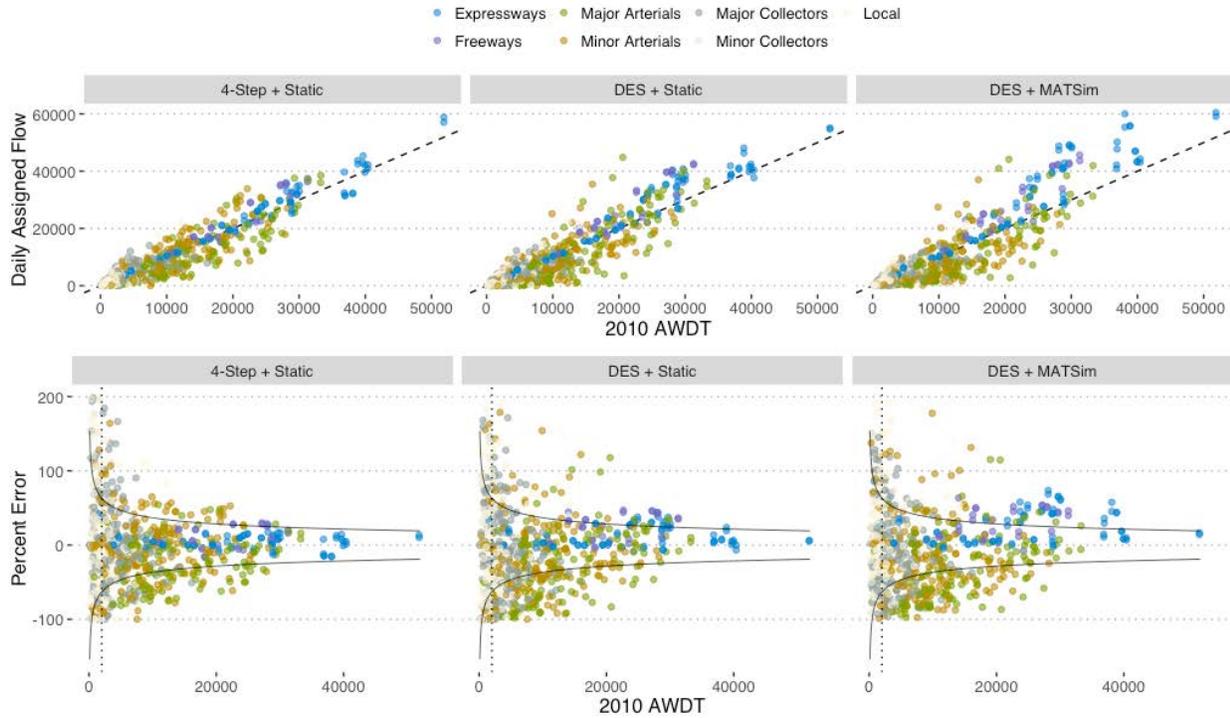


FIGURE 12 Validation of link volumes in Asheville, N.C. for the existing model (4-Step + Static), the data-driven demand model with the existing assignment model (DES + Static), and the data-driven model with MATSim (DES + MATSim).

TABLE 3 Validation of Link Volumes

Facility Type	n	Avg Count	Four-step + Static			DES + Static		
			VHT	RMSE	PRMSE	VHT	RMSE	PRMSE
Freeways	103	22,935.6	62,051.2	3,321.6	14.5%	78,042.0	5475.1	23.9%
Expressways	6	13,732.2	2,022.9	2,168.2	15.8%	1,965.0	1984.5	14.5%
Boulevards	15	15,965.1	2,081.2	5,695.2	35.7%	2,302.2	5920.4	37.1%
Other Major Thoroughfares	337	10,322.2	31,646.6	4,181.5	40.5%	32,980.0	5113.7	49.5%
Minor Thoroughfares	414	2,938.9	18,360.0	2,026.8	69.0%	18,856.2	2359.8	80.3%
Total	875	8,433.7	116,162.0	3,250.3	38.5%	134,145.4	4106.4	48.7%

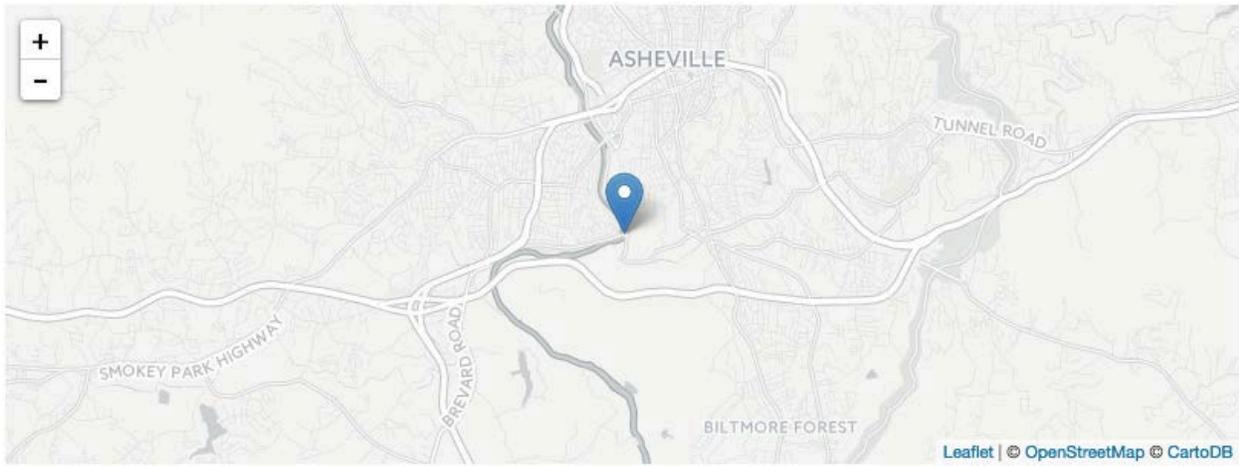


FIGURE 13 Map of bridge closure scenario in Asheville, N.C.

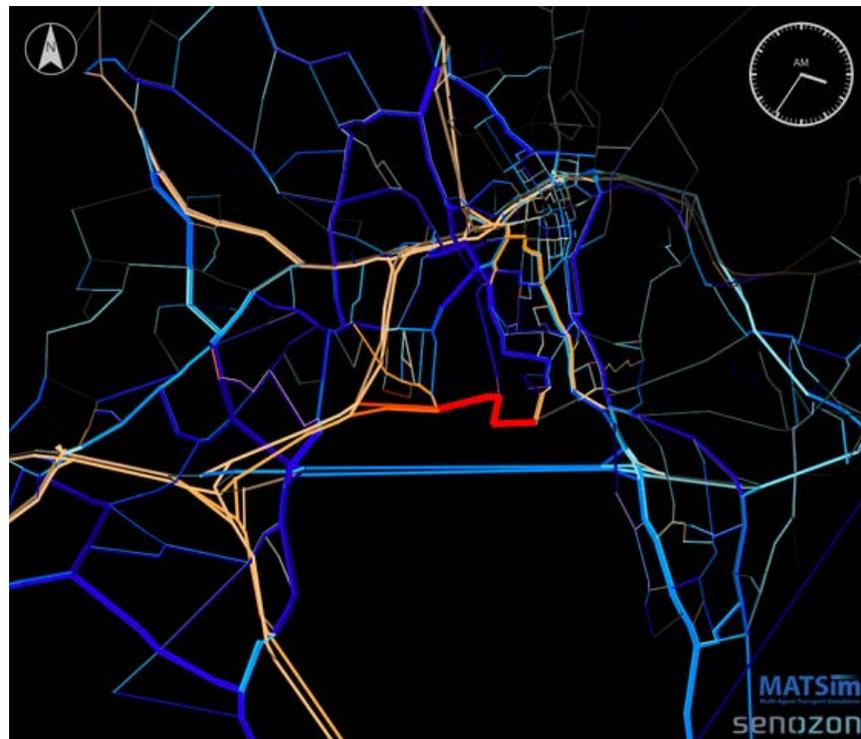


FIGURE 14 Network difference plot for the bridge closure scenario in Asheville, N.C.

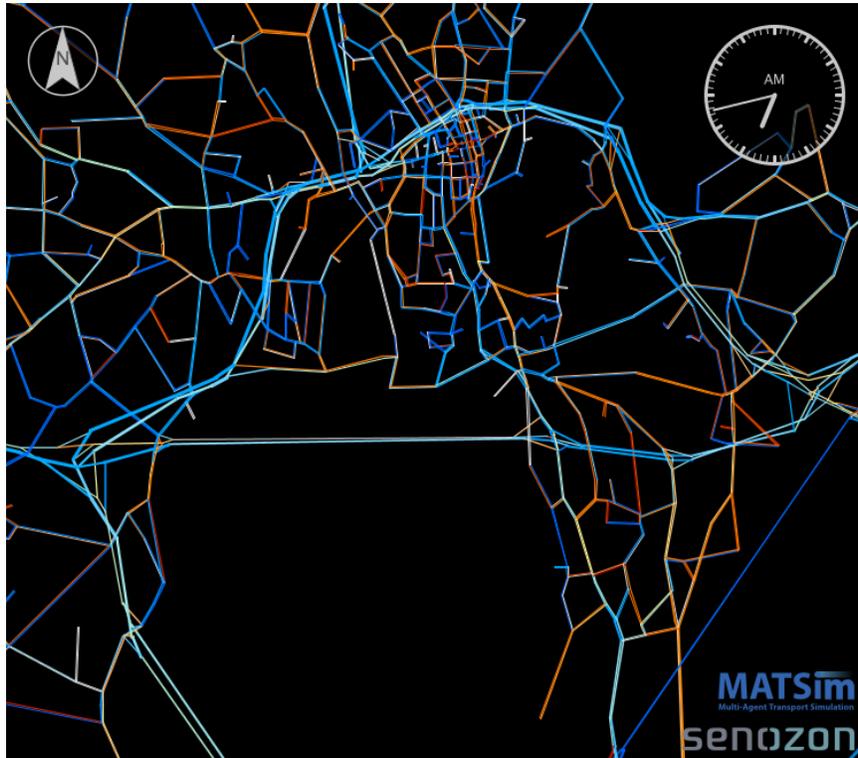


FIGURE 15 Relative network difference plot for the tolling scenario just before the morning toll in Asheville.



FIGURE 16 Relative network difference plot for the tolling scenario just after the morning toll in Asheville.