Improved Travel Demand Modeling with Synthetic Populations

Kaidi Wang[1], Wenwen Zhang[1], Henning Mortveit[2,3], and Samarth Swarup[3]

[1] Urban Affairs and Planning, Virginia Tech, Blacksburg, VA 24061, USA
{kaidi,wenwenz3}@vt.edu
[2] Department of Engineering Systems and Environment
[3] Biocomplexity Institute and Initiative,
University of Virginia, Charlottesville, VA 22911, USA
{henning.mortveit,swarup}@virginia.edu

Abstract. We compare synthetic population-based travel demand modeling with the state of the art travel demand models used by metropolitan planning offices in the United States. Our comparison of the models for three US cities shows that synthetic population-based models match the state of the art models closely for the temporal trip distributions and the spatial distribution of destinations. The advantages of the synthetic population-based method are that it provides greater spatial resolution, can be generalized to any region, and can be used for studying correlations with demographics and activity types, which are useful for modeling the effects of policy changes.

Keywords: Travel demand · Transportation · Synthetic population.

1 Introduction

Travel demand modeling refers to modeling population movements within a region, typically over the course of a fixed time period such as day or a week. Mobility depends on a number of factors, such as demographics, transportation infrastructure, the build environment, and more.

Transportation planning and demand modeling are required to receive federal transportation funds for larger urban areas in the U.S. [9]. Based the most recent regulation, the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU), transportation plans need to address many requirements, such as air quality issues, multimodal planning, better manage the existing system, expand public input, and financial requirements [15]. Transportation demand models play very important roles in forecasting and assessing whether the proposed transportation planning alternatives can help the region to meet the corresponding requirements. Therefore, all Metropolitan Planning Organizations (MPOs) for areas with population more than 50,000 have to develop, implement, and calibrate local travel demand models to evaluate a broad range of alternatives [9].

This has some limitations. First, MPOs don't do this planning for smaller regions. Thus, the coverage doesn't extend over the whole country. Second, there is a lack of spatial refinement in existing models, as all trips are attributed to Traffic Analysis Zones (TAZs), as we explain in the next section. Third, these models are not applicable in abnormal situations, such as mobility during disasters.

To address these limitations, we are exploring the use of synthetic populations [2, 20], which provide a disaggregated model of the population, their activity schedules, and activity locations. The synthetic population approach to generating travel demand is described in Section 3.

In the present work, we compare the synthetic population-derived travel demand with the travel demand generated by two models used by MPOs, for three US cities. The goals are to see how closely the models match, what the differences are, and where the synthetic population approach might be improved. Once the approach is validated, we can use it to do travel demand modeling for all regions in the US.

2 State of the Art in Mobility Modeling

Currently, a majority of MPOs in the United States adopt two genres of travel demand models, namely the conventional four-step travel demand model and the latest activity-based travel model. The four-step model is a widely adopted transportation demand forecast framework that can be dated back to the 1950s [19]. The model adopts four specific steps, including trip generation, trip distribution, mode choice, and trip assignment, to forecast future travel demand given changes in the spatial distribution in employment and population and performance of a transportation system within a region. The first trip generation step estimates the number of produced and attracted trips for each Traffic Analysis Zone (TAZ). The trip distribution step connects trip origins to destinations, which results in a person trip Origin-Destination (OD) matrix. The mode choice step divides the person trip OD matrix by travel mode, such as passenger vehicles, transit, etc., and generates mode-specific OD matrices for vehicle trips by the time of the day. The last trip assignment component forecasts the route for trips. The unit of analysis for the four-step model is zone-level trips. Thus, the model is not sensitive to demand and supply policies, as individual decision making is barely incorporated in the model [9].

The activity-based model advances the four-step model by forecasting travel demand at a more refined unit of analysis [5]. The activity-based model is typically developed at a disaggregated person level, enabling the model to evaluate possible changes in travel behavior and system performances across policy scenarios. However, the modeled geographic unit is similar to the four-step model, which is typically TAZs. In other words, all activities, trip origins, and destinations are assigned to TAZ centroids. Some MPOs tend to adopt more refined TAZ boundaries in the activity-based model compared with the four-step model [4]. The activity-based models, however, are more data and computational resource-consuming compared with the four-step model. Thus, only a limited number of

MPOs have adopted the activity-based model [10]. Though several MPOs have started to migrate from the four-step model to the activity-based model, the four-step model remains the most commonly used travel demand model in the U.S [10].

Both of the aforementioned demand forecasting frameworks were developed for regional planning purposes. In those scenarios, a TAZ-level model is considered sufficient for planning-related decision making. However, the model fails to support policy making at refined spatial scales to address emerging transportation problems (e.g., chaotic curb uses) introduced by disruptive transportation modes, especially ride-hailing services and the envisioned Shared and Private Autonomous Vehicles (AVs). Even after incorporating these emerging travel modes into the four-step [28] and activity models [12, 30], the model outputs are constrained at TAZ level, which are not very useful to support refined decision making, such as block level curb spaces allocation. Meanwhile, the ride-hailing companies, such as Uber and Lyft, are reluctant to release detailed trip data, due to competition and privacy concerns. Finally, different MPOs tend to model mobility demand using various data sources (e.g., National Household Travel Survey [NHTS] vs. local household travel survey) and are calibrated using different base year data, rendering it difficulty to conduct research for cross-city and region comparisons [19].

Therefore, in this study, we proposed a disaggregated travel demand modeling approach that is built upon synthetic populations (developed using multiple datasets, as described in the next section) and nationally available transportation network and Point of Interest (PoI) data to fill the current demand model and data gaps. We validate our modeling outputs by comparing spatiotemporal distributions of synthesized trips with Origin-Destination (OD) matrices (i.e., the product of mode choice). The OD matrices contain the number of estimated trips for each pair of origin and destination. Given that in most regions, vehicle travel is the dominant travel mode, our comparison will only focus on vehicle trips. We obtained OD matrices from three different regions with various urban forms, travel patterns and current transportation infrastructures, namely Richmond, VA, Seattle, WA, and Atlanta, GA. The adopted travel demand models differ across these cities, in terms of travel demand data sources, modeling framework, and modelled time periods, as displayed in Table 1.

Table 1: Model Settings for Validation OD Metrics

Model Settings	**Atlanta, GA**	**Richmond, VA**	**Seattle, WA**
Model Framework	Activity-based	Four-step	Four-step
Model Data Source	2011 local survey	2009 NHTS	2014-2015 local survey
Caliberated Base Year	2015	2012	2014
Model Time Periods	5	4	12

3 The Synthetic Population Approach

A "synthetic population" [2, 11, 16] is a very detailed model of a region, including the resident population, their daily or weekly activity patterns, their networks of interaction, and the built environment. The last includes buildings, and also infrastructures for transportation, power, communication, etc.

Synthetic populations have been used as the basis for multi-agent simulations in a variety of domains, including computational epidemiology [14], disaster response [6], transportation planning [3, 30], and more [24, 27]. They provide high resolution, high fidelity representations, enabling realistic simulations which can be used for meaningful policy recommendations [7]. A synthetic population is generated through a series of steps. We describe the initial steps briefly below, and present the mobility modeling step (assigning locations to activities) in more details. Further information is given in a technical report [20].

Generating agents with demographics: We use data from the American Community Survey [25], which provides demographic distributions for each block-group and a 5% sample of complete records for a slightly larger area, known as the Public Use Microdata Sample (PUMS). These are combined using the statistical technique called Iterative Proportional Fitting (IPF) [13, 8] to generate a joint distribution over selected demographic variables. We chose *age of householder*, *household income*, and *household size* as the variables for the IPF step. From this, we sample the resulting joint distribution and select matching households from the PUMS data to create the population of synthetic agents.

Assigning activity patterns: Each person p created in the previous step is assigned an *activity sequence* $\alpha(p) = (a_{i,p})_i$ where each *activity* $a_{i,p}$ has a start time, a duration, and an activity type. For the synthetic population used in this work, the activity types are from the set

$$\mathcal{A} = \{\texttt{Home}, \texttt{Work}, \texttt{School}, \texttt{Shopping}, \texttt{Religion}, \texttt{Other}\} . \tag{1}$$

The activity sequence survey data was taken from the National Household Travel Survey (NHTS) 2017 [23]. From this, consistent week-long activity sequences were constructed and assigned using CART and the Fitted Values Means method [18].

Assigning locations to activities: This modeling step connects people and their activities to the set $\mathcal{L}$ of *residence*- and *activity locations* of the given region.

The first part of this modeling step constructs the *locations*. This is done based on the MS Building Footprint data [21] which we have augmented with a residential/non-residential classification based on the HERE Premium StreetMap landuse classifications and extended POI listings [17]. Each non-residence location, which we refer to as an *activity location*, is additionally augmented with a weight for each non-`Home` activity reflecting the likelihood of people conducting that particular activity at the given location. Each household is mapped to a residence location. The assignment of residence locations is done for each block-group. First each possible residence location is assigned one household, to ensure that there are no residence locations without at least one household. The remaining households are assigned residence locations with probability proportional to the area of the building footprint.

The second step assigns people's activities to locations. Abstractly, for each person p this step constructs a map $\lambda_p\colon \alpha(p) \longrightarrow \mathcal{L}$ that assigns to each activity $a_{i,p}$ of p a location $\ell \in \mathcal{L}$. For the various activity types, this algorithm has the following sequence of steps:

- Using NCES data [22], assign to each residence, the vector-valued ID containing the nearest public school for each grade level;
- Construct the normalized county/county work commute flow matrix M adjusted with county self-references using ACS commute flow data [1];
- For each person p:
  - assign each activity a_i of type `Home` to the residence location of p;
  - assign each activity a_i of type `School` to the age-appropriate school location assigned to their residence location;
  - select a work location using this 2-step process: (a) randomly select a target county c' using the probability distribution M_c where c is the county of p. (b) For county c', randomly select a work location ℓ from the set of activity locations $\mathcal{L}_A|_{c'}$ of c' using the probability distribution induced by the locations' `Work` weights. Assign all `Work` activities of p to ℓ. Thus, each working person has a consistent work location for the entire period.
  - if c' supports `Shopping` (resp. `Other`), assign `Shopping` (resp. `Other`) activities independently at random to the set of activity locations $\mathcal{L}_A|_{c'}$ of c' using the distribution induced by their `Shopping` (resp. `Other`) weights. If c' has no activity locations supporting `Shopping` (resp. `Other`), repeat this process using the home county c. If c does not support `Shopping` (resp. `Other`), select a county using the probability distribution M_c and repeat.
  - if c supports `Religion`, randomly select a location ℓ from the set of activity locations $\mathcal{L}_A|_c$ of c using the distribution induced by their `Religion` weights. If c has no activity location supporting `Religion`, select a county c'' using the probability distribution M_c and repeat for c''.
- Additionally, one may construct a person-person contact network using some form of location co-occupancy model; we do not need that for this work.

In the present work, we extract the collection of activities that take place on a Tuesday. Travel demand is constructed from the activity schedules by extracting the locations for successive activities. The start time for the travel is taken to be the end time of the first activity. If two successive activities take place at the same location, there is no travel, and this pair is not included in the travel demand file. Next we describe the comparison between travel demand constructed from synthetic populations and travel demand data obtained from three US MPOs, who use traditional models.

4 Comparison of Results

We validated the synthetic approach by comparing the distributions of synthesized trips with that of OD matrices generated by local travel demand models, using data from three cities, namely Richmond, VA, Atlanta, GA, and Seattle, WA. These three cities are selected because they tend to have significantly different urban forms and transit infrastructures. Cities like Richmond and Atlanta have more urban sprawl and have limited transit systems, while Seattle is more densely developed and maintains an extensive transit system. Specifically, we compared the number of generated trips, the distributions of trip departure times, as well as the spatial distributions of the trip origins and destinations, to determine if the synthetic trips are representative and can replicate the distributions generated by travel demand models, including both the activity-based model and the conventional four-step model.

4.1 Trip Counts Comparison

The number of trips generated by each approach for each city is illustrated in Table 2. For the three study areas, we only compare the trips that both start and end within the city boundaries. The number of daily trips generated by four-step travel demand model in Richmond is 370,998. Synthetic approach generates 401,042 daily trips, which is 8.1% more than that in four-step travel demand model. Seattle also sees slightly more synthetic trips. Notice that the Richmond travel demand model is calibrated using 2012 ACS data, while the Seattle model is calibrated using 2014 ACS data. The accuracy of the projected 2017 OD matrices from these models may vary depending on the quality of local population and employment forecasting model. The synthetic trips are generated using 2017 ACS data, which should be considered as more accurate compared with local forecasts. It is interesting that Atlanta's activity model generates significantly more trips than synthetic trips. ARC calibrated and validated the activity-based model (ABM) using 2011 regional household travel survey and then forecast travel demand in 2015, while the synthetic trip profiles are generated using 2017 NHTS data. This is largely because the synthetic population method assigns some destinations to locations outside the Census blockgroups that are within the city of Atlanta. These get eliminated when we restrict our analysis to travel demand in Atlanta. There may be two additional reasons for the discrepancy in daily trip generation: (1) the trip generation rate for Atlanta may decrease from 2011 to 2017, and (2) Atlanta NHTS data are not representative in the 2017 survey. In the concluding section, we address the possibility of reducing this discrepancy by using other data sets. It is important to note that the synthetic population model is generated for the entire state, so the total number of trips taken by the residents of Atlanta are not fully represented here. In other words, the discrepancy is not due to a systematic bias in the travel demand, but due to the fact that we are restricting the analysis to a subregion.

Table 2: Trip Counts Comparison

City	# of Trips Obtained from MPO	# of Synthetic Trips	Percentage Difference
Richmond	370,998	401,042	8.1%
Seattle	1,087,814	1,152,136	5.6%
Atlanta	1,087,418	796,688	-26.7%

4.2 Departure Time Comparison

The number and percentage of trips by the periods of departure time is shown in Table 3 for each city. Overall, the temporal patterns of trips generated by the travel demand model and synthetic approach are similar. The share of trips in each of the time periods by the two methods is close. Thus, we conclude that the distribution of departure time of synthetic trips matches that in demand models. Notice that Seattle and Atlanta have more time periods in their travel demand model. We aggregated the trips into 4 time periods to make the comparison more intuitive. The MPO for Richmond defined AM period as 6:30 am - 8:30 am and PM period as 4:30 pm - 6:30 pm. ARC has five periods in the ABM. i.e. early morning, morning, midday, afternoon, evening. We collapsed early morning and evening into a night period. Finally, the AM period is 6 am - 10 am and PM period is 3 pm - 7 pm. Puget Sound Regional Council (PSRC), the MPO for Seattle, used 12 periods in the demand model. We aligned the periods with the other two cities as much as possible and end up with AM period between 7 am - 10 am and PM period between 4 pm - 8 pm. The synthetic population model gives actual trip start times, so it can be aggregated for any definition of the bins. In any case, the definition of time periods does not influence the comparison of models as we compare for each period rather than across periods.

Table 3: Trip Count by Time Period

Period	Model	Richmond	Seattle	Atlanta
AM	synthetic	54,996 (14%)	251,371 (22%)	190,226 (24%)
	demand model	34,563 (9%)	191,111 (18%)	216,115 (20%)
MD	synthetic	201,063 (50%)	453,975 (39%)	258,049 (32%)
	demand model	198,413 (53%)	402,391 (37%)	399,421 (37%)
PM	synthetic	68,554 (17%)	316,430 (27%)	258,057 (32%)
	demand model	42,648 (11%)	371,374 (34%)	305,209 (28%)
NT	synthetic	76,429 (19%)	130,360 (11%)	90,356 (11%)
	demand model	95,374 (26%)	122,938 (11%)	166,673 (15%)

4.3 Spatial Distribution Comparison

We validated the spatial distributions of the trips by examining the spatial correlation of buffered trip origins and destinations. Each origin or destination is a traffic analysis zone (TAZ). There are 219 TAZs in Richmond, 856 TAZs in Seattle and 829 TAZs in Atlanta. Table 4 shows the computed correlations using data from two sources by time periods for each city. A buffered TAZ includes a TAZ and its neighbor TAZs based on queen contiguity criteria, which means if two TAZs share a vertex or edge, they are neighbors. The reason we do not directly impute spatial correlation of origins and destinations is that the correlation cannot reflect the actual spatial pattern. For example, the spatial correlation of destinations could be extremely low even though the synthetic trips end in areas close to the destinations of trips in demand model. Including neighbor TAZs when comparing the distribution will mitigate this issue.

Table 4: Spatial Distribution Pearson Correlation of Origins and Destinations

City	AM	MD	PM	NT
Correlations of Origins				
Richmond	0.85	0.79	0.73	0.81
Seattle	0.86	0.53	0.40	0.54
Atlanta	0.84	0.71	0.65	0.82
Correlations of Destinations				
Richmond	0.62	0.77	0.85	0.83
Seattle	0.35	0.58	0.70	0.39
Atlanta	0.58	0.73	0.84	0.86

It can be observed that the spatial correlation of origins peaks in AM and that of destinations peaks in PM or NT for all of the three cities. In contrast, they all experience least correlated origins in PM and destinations in AM. The low correlation is because of different methods of estimating employment in the two approaches. MPOs estimate employment based on ACS block-level data while the synthetic approach uses county-level commute flows. The destinations in AM and origins in PM are mostly the locations of jobs. Therefore, the different estimates lead to low correlations. This indicates one aspect in which the synthetic population model can be refined.

As shown in Table 4, the destinations in AM in Seattle is the least spatially correlated in all the cities and periods. Their spatial distribution is illustrated in Figure 1. Both maps demonstrate that the trips tend to end in commercial zones such as downtown Seattle in the middle, industrial district in the southeast and Northgate in the north. Only a small portion of trips travel to Northeast Seattle and Ballard, where most land use type is residential zones. Notice that the very west and very east areas in the map are mostly water areas. Therefore, few trips end there. In general, the spatial distribution of synthetic trips matches that of trips generated by the MPOs well.

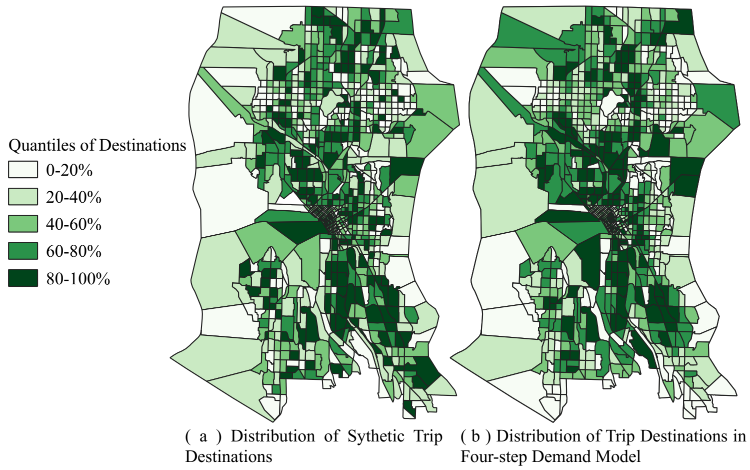

Fig. 1: Distribution of Destinations in AM in Seattle

5 Conclusion

This study evaluated the synthetic population-based model by comparing it with the state of the art travel demand models used by Metropolitan Planning Offices in the United States. This was done by comparing the spatial and temporal distribution of the trips generated by the two approaches. The results indicate that the trip count estimated by synthetic approach is close to that estimated by travel demand model. The synthetic approach matches demand models for the distribution of departure times. The spatial correlations of origins and destinations are mostly high except some specific periods. The low spatial correlation of origins in PM and destinations in AM reflects the difference of the two models. The synthetic population model uses ACS county-county commuter flows, but within counties chooses destinations randomly (though weighted by an estimate of building capacity). Switching to a higher resolution, such as the Census LODES data product [26], might help with this.

Synthetic population-based models are more flexible compared to four-step travel demand model. They are less computationally expensive compared to activity-based model while providing more details on the trips and the travelers. Besides, this approach is generalizable that most areas can use it to estimate travel demand. Thus it can be used to do a multiple city study as it enables direct comparison between cities. Future work can also be to combine it with research on automated vehicles. For example, the social interaction potential could be explored with a shared automated vehicle (SAV) system using synthetic trips as input, e.g., by integrating with simulations of potential shared autonomous vehicle use [29, 28].

Acknowledgments

We thank our external collaborators and members of the Network Systems Science and Advanced Computing (NSSAC) division for their suggestions and comments. This work was partially supported by DTRA Grant HDTRA1-17-F-0118 and NASA grant 80NSSC18K1594. We also thank Atlanta Regional Commission (ARC), Puget Sound Regional Council (PSRC) and Richmond Regional Planning District Commission for providing us with travel demand model outputs.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views, official policies or endorsements, either expressed or implied, of NASA, DTRA, or the U.S. Government.

References

1. 2011-2015 5-year ACS commuting flows, `https://www.census.gov/data/tables/2015/demo/metro-micro/commuting-flows-2015.html`, last accessed: 24 Feb 2020
2. Adiga, A., Agashe, A., Arifuzzaman, S., Barrett, C.L., Beckman, R.J., Bisset, K.R., Chen, J., Chungbaek, Y., Eubank, S.G., Gupta, S., Khan, M., Kuhlman, C.J., Lofgren, E., Lewis, B.L., Marathe, A., Marathe, M.V., Mortveit, H.S., Nordberg, E., Rivers, C., Stretz, P., Swarup, S., Wilson, A., Xie, D.: Generating a synthetic population of the United States. Tech. Rep. NDSSL 15-009, Network Dynamics and Simulation Science Laboratory, Virginia Bioinformatics Institute, Virginia Tech (2015)
3. Adiga, A., Marathe, M., Mortveit, H., Wu, S., Swarup, S.: Modeling urban transportation in the aftermath of a nuclear disaster: The role of human behavioral responses. In: The Conference on Agent-Based Modeling in Transportation Planning and Operations. Blacksburg, VA (Sep 30 - Oct 2 2013)
4. Atlanta Regional Commission, et al.: Activity-based travel model specifications: Coordinated travel–regional activity based modeling platform (CT-RAMP) for the Atlanta region (2012), `http://documents.atlantaregional.com/The-Atlanta-Region-s-Plan/RTP/abm-specification-report.pdf`
5. Axhausen, K.W., Gärling, T.: Activity-based approaches to travel analysis: Conceptual frameworks, models, and research problems. Transport Reviews **12**(4), 323–341 (1992)
6. Barrett, C., Bisset, K., Chandan, S., Chen, J., Chungbaek, Y., Eubank, S., Evrenosoğlu, Y., Lewis, B., Lum, K., Marathe, A., Marathe, M., Mortveit, H., Parikh, N., Phadke, A., Reed, J., Rivers, C., Saha, S., Stretz, P., Swarup, S., Thorp, J., Vullikanti, A., Xie, D.: Planning and response in the aftermath of a large crisis: An agent-based informatics framework. In: Pasupathy, R., Kim, S.H., Tolk, A., Hill, R., Kuhl, M.E. (eds.) Proceedings of the 2013 Winter Simulation Conference. pp. 1515–1526. IEEE Press, Piscataway, NJ, USA (2013)
7. Barrett, C., Eubank, S., Marathe, A., Marathe, M., Pan, Z., Swarup, S.: Information integration to support model-based policy informatics. The Innovation Journal **16**(1), article 2 (2011)
8. Beckman, R.J., Baggerly, K.A., McKay, M.D.: Creating synthetic base-line populations. Transportation Research A – Policy and Practice **30**, 415–429 (1996)
9. Beimborn, E., et al.: A transportation modeling primer (2006)

10. Castiglione, J., Bradley, M., Gliebe, J.: Activity-based travel demand models: A primer. No. SHRP 2 Report S2-C46-RR-1 (2015)
11. Chapuis, K., Taillandier, P., Renaud, M., Drogoul, A.: Gen*: A generic toolkit to generate spatially explicit synthetic populations. International Journal of Geographical Information Science **32**(6), 1194–1210 (2018)
12. Childress, S., Nichols, B., Charlton, B., Coe, S.: Using an activity-based model to explore the potential impacts of automated vehicles. Transportation Research Record **2493**(1), 99–106 (2015)
13. Deming, W.E., Stephan, F.F.: On a least squares adjustment of a sampled frequency table when the expected marginal tables are known. Annals Math. Stats **11**(4), 427–444 (1940)
14. Eubank, S., Guclu, H., Kumar, V.S.A., Marathe, M., Srinivasan, A., Toroczkai, Z., Wang, N.: Modelling disease outbreaks in realistic urban social networks. Nature **429**, 180–184 (May 2004)
15. Fischer, J.W., Resources, S., Division, I.: Safe, accountable, flexible, efficient transportation equity act–a legacy for users (safetea-lu or safetea): Selected major provisions. Congressional Research Service, The Library of Congress (2005)
16. Gallagher, S., Richardson, L.F., Ventura, S.L., Eddy, W.F.: SPEW: Synthetic Populations and Ecosystems of the World. Journal of Computational and Graphical Statistics **27**(4), 773–784 (2018)
17. HERE premium streetmap for the U.S., `https://www.here.com/`, last accessed: 24 Feb 2020
18. Lum, K., Chungbaek, Y., Eubank, S., Marathe, M.: A two-stage, fitted values approach to activity matching. International Journal of Transportation **4**, 41–56 (2016). https://doi.org/10.14257/ijt.2016.4.1.03
19. McNally, M.G.: The four step model. Handbook of transport modelling **1**, 35–41 (2000)
20. Mortveit, H.S., Adiga, A., Barrett, C.L., Chen, J., Chungbaek, Y., Eubank, S., Kuhlman, C.J., Lewis, B., Swarup, S., Venkatramanan, S., Wilson, A., Xie, D., Marathe, M.V.: Synthetic populations and interaction networks for the U.S. Tech. rep. (2020), nSSAC Technical Report: #2019-025
21. Microsoft building footprint data U.S., `https://github.com/microsoft/USBuildingFootprints`, last accessed: 24 Feb 2020
22. National Center for Education Statistics, `https://nces.ed.gov/datatools/`, last accessed: 24 Feb 2020
23. National Household Travel Survey, `https://nhts.ornl.gov/`, last accessed: 24 Feb 2020
24. Swarup, S., Gohlke, J.M., Bohland, J.R.: A microsimulation model of population heat exposure. In: Proceedings of the 2nd International Workshop on Agent-based Modeling of Urban Systems (ABMUS). Sãu Paulo, Brazil (2017)
25. United States Census Bureau: American Community Survey 2013-2017 5-year estimates, `AmericanFactFinder:https://factfinder.census.gov/`
26. U.S. Census Bureau: LEHD origin-destination employment statistics (LODES), `https://lehd.ces.census.gov/data/`
27. Waddell, P.: UrbanSim: Modeling urban development for land use, transportation, and environmental planning. Journal of the American Planning Association **68**(3), 297–314 (2002)
28. Zhang, W., Guhathakurta, S.: Parking spaces in the age of shared autonomous vehicles: How much parking will we need and where? Transportation Research Record **2651**(1), 80–91 (2017)

29. Zhang, W., Guhathakurta, S., Fang, J., Zhang, G.: Exploring the impact of shared autonomous vehicles on urban parking demand: An agent-based simulation approach. Sustainable Cities and Society **19**, 34–45 (2015)
30. Zhang, W., Mortveit, H.S., Swarup, S.: Estimating shared autonomous vehicle fleet size to meet urban daily travel demand. In: The 3rd International Workshop on Agent-based Modeling of Urban Systems (ABMUS). Stockholm, Sweden (2018)