

Integrated simulation of activity-based demand and multi-modal dynamic supply for energy assessment

Siyu Chen
Isabel Hemerly Viegas de Lima
Zach Needell
Andrea Araldo
Arun Prakash Akkinipally
Jessica Trancik
Moshe Ben-Akiva

Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

Nicholas Fournier
Aikaterini Deliali
Eleni Christofa
University of Massachusetts, Amherst
Amherst, Massachusetts, USA

Carlos Lima Azevedo
Technical University of Denmark
Lyngby Kgs., Denmark

Abstract—The development of a large scale agent-based simulation model for the Greater Boston Area is presented, closing the gap between state-of-the art integrated demand-supply modeling techniques (SimMobility) with advanced energy estimation models (TripEnergy) and shedding light on its practical application to large urban areas. This paper describes the technical details of its three key components (activity-based demand, multi-modal dynamic supply, and trajectory-based energy models), the used data, the model estimation, integration and calibration processes. The proposed model can simulate any day with and without congestion in order to capture changes in energy use across all dimensions of a mobility system, namely temporal, spatial, modal or functional. For an average 24h in the Greater Boston Area the simulated travel of 4.5-million people resulted in 15-million trips and a total vehicle energy consumption of 548 thousand equivalent gallons of gasoline. Our proposed platform allows for the comprehensive and consistent assessment of energy related policies, technologies and services affecting traveler behavior, the transportation system's and vehicle energy performances.

Index Terms—Mobility, energy, simulation, microscopic, activity-based, multi-modal.

I. INTRODUCTION

The objective of this research is to develop a large scale transportation simulation model that can be used to estimate energy consumption and optimize incentives for network-wide energy consumption reduction. The System Model (SM) modeling and simulation platform is developed to act as a proxy for the real world and to replicate as closely as possible travelers' reaction to information and incentives, as well as the multi-modal transportation system performance for the Greater Boston area (GBA). SM consists of three main components: supply, demand, and energy models. Supply and demand are simulated using SimMobility [1] and the energy consumption is simulated using TripEnergy [2].

SimMobility is an integrated agent-based simulation platform used to evaluate a wide range of future mobility related scenarios. It is comprised of three primary modules differentiated by the time-frame in which we consider the

behavior and operation of an urban system: short-term (a microscopic mobility simulation, few hours stimulated at 0.1-second time resolution), mid-term (an activity-based model integrated with a dynamic multi-modal assignment simulator, daily simulation at 5-second resolution), and long-term (a land use and long-term behavioral model, at 6-months to one year resolution) [1]. For this work we are utilizing the mid-term (MT) module where SimMobility agents behavior is modeled in terms of activity, travel plans and actions [3]. It is categorized as a mesoscopic simulator since it combines the activity-based micro-simulator on the demand side with macroscopic simulation on the supply side.

The energy model has two components: a trip matching algorithm and a vehicle energy model. The matching algorithm merges high-resolution 1-Hz velocity histories with more representative but low-detail trip data (i.e., the trajectories produced from the mesoscopic simulator). The matched velocity histories are then fed into the vehicle energy model to produce energy estimates that account for driving style and trip type [4, 5].

II. LITERATURE REVIEW

With advances in communications and computational power, large-scale integrated models are becoming increasingly common and complex. These models are typically composed of several loosely interconnected models, each with a specialized purpose. For example a land-use, demographic, and economic simulator (e.g. Urbansim [6]) feeds into a demand modeler (e.g. ALBATROSS [7] or FEATHERS [8]), which then feeds into a traffic assignment simulator (e.g. DynaMIT [9]). These components are often integrated within platforms, such as TRANSIMS [10], CEMDAP [11], or MATSim [12], which simulate both demand and network assignment.

Nearly all transport models provide speed and density of traffic on links in the network. This information can then be used to estimate energy or emissions using a vehicle

model. Coarse estimates can be achieved using empirically measured energy and emission rates based on link speed with models such as COPERT [13]. However, these models fail to capture the effect of acceleration/deceleration on energy and emissions. More accurate estimates can be made using high resolution trajectories, such as with ADVISOR [14], MOVES [15], VT-Micro [16], or CMEM [17]; but are computationally intensive and are not practical for large-scale networks. More recently, simulations have interpolated higher detail trajectories from moderate-resolution trajectories produced by a mesoscopic network simulation [18, 19]. These models rely on traffic flow theory to model acceleration and deceleration behavior of drivers based on traffic density, and infer vehicle specific power from the interpolated trajectories.

The challenge in integrating long-term macroscopic models (e.g. housing and economics) and short term microscopic models (e.g. traffic), is that it can impose certain limitations on the simulation platform, such as restricting the ability of agents to change their daily activity patterns. This poses a problem when attempting to integrate auxiliary models, such as emissions or energy models where outcomes may be dependent upon intra-day or intra-trip changes.

III. DEMAND MODELS

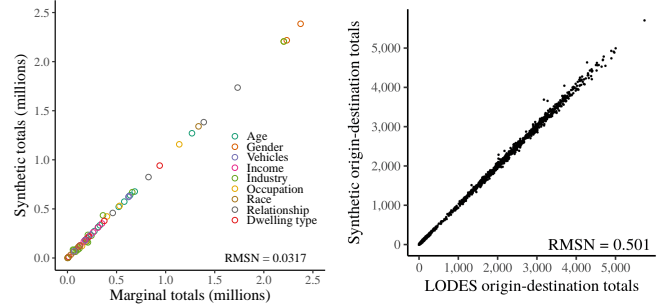
Travel demand is modeled using three components: synthetic population, pre-day model and within-day model. The synthetic population provides a disaggregated set of individuals as demand generators. The pre-day model estimates the planned activities and trips of individuals and the within-day model accounts for its modifications that occur during a day.

A. Population Synthesis

A fully disaggregated population of individuals, households, and vehicles with attributes is needed for agent-based simulation. Such a population can be synthesized using detailed samples and marginal total data available for the simulated area. The synthetic population is generated in three phases, (1) baseline generation with fixed work-trip distribution, (2) auxiliary models and attribute assignment, and (3) synthetic vehicle population and ownership model. The baseline population is generated using Iterative Proportional Fitting (IPF) [20]. IPF is a method to fit a joint distribution of cells to match known marginal totals by proportionally scaling cells across dimensions until error converges. Since IPF is also used for fixed work-trip distribution, it has been integrated directly into the population synthesis IPF for improved accuracy [21].

The synthetic population was generated using US Census data with aggregated marginals from the American Community Survey (ACS) [22], disaggregated samples from Public Use Microdata Samples (PUMS) [23], and trip distribution marginals from the Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics [24].

In order to estimate energy and emissions, a detailed vehicle fleet is necessary. A synthetic population of vehicles allocated to individuals is also generated and assigned to owners in the person population. This is achieved using a



(a) Marginal fit of population (b) Fixed work trip distribution fit

Fig. 1: Fixed work trip distribution fit

three-part process: (1) estimate a vehicle ownership choice model as a multinomial logit model; (2) generate a pool of assigned vehicle classes directly from individual persons choices, and (3) assign a specific vehicle type to each vehicle, based on its assigned class.

B. Pre-day

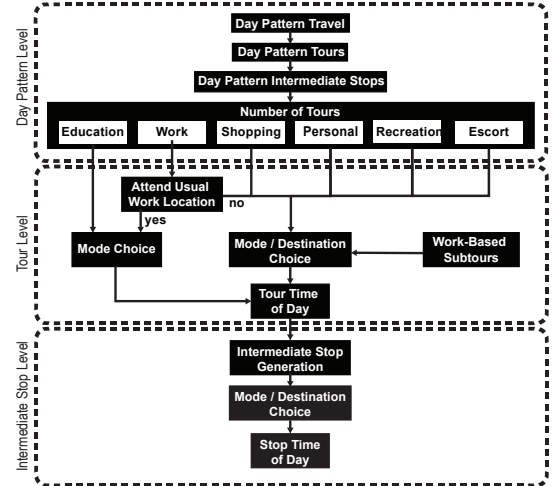


Fig. 2: Pre-day Model Structure

Pre-day models follow the Daily Activity Schedule (DAS) approach [25] to decide an initial overall daily activity schedule of an agent. This includes tours, sub-tours, preferred modes, departure times by half-hour slots, and destinations. This is based on the sequential application of hierarchical discrete choice models using a Monte-Carlo simulation. This allows for a direct modeling of individual trip purposes and is capable of capturing the dependencies between within-day decision making, fundamental for the proper evaluation of behavioral change and impacts. Figure 2 presents how the pre-day models from lower levels are conditioned on decisions made with models from higher levels. There are three different hierarchies in the system: day pattern level, tour level, and intermediate stop level.

1. Day pattern models: Agents decide on the types of activities to be performed during the day. These activities

can either be the primary activities of tours, or activities performed at intermediate stops. After selecting their activities, agents decide on the exact number of tours performed for each activity type (e.g., work, education, shopping, recreation, personal, or escort).

2. **Tour level models:** These include the tour mode/destination models and tour time of day models. Mode/destination choice models vary by primary activity: for education tours the destination is known and therefore only mode choice is simulated; for work tours, an additional model is used depending on whether an individual has a usual work location or not. If the usual work location is known, only the tour mode is simulated. Otherwise if not, then mode and destination are simulated simultaneously. In addition, a work-based sub-tour generation model is applied to work tours in order to simulate the number of sub-tours performed by an individual during a work activity. Finally, a tour time of day model is applied in order to simulate the arrival and departure time for the primary activity of each tour simultaneously.

3. **Intermediate stop level models:** An intermediate stop generation model is applied for each tour in order to simulate the number of stops within the tour. The mode/destination choice is then simulated for each stop. Finally, the stop time of the day model is applied in order to simulate the arrival or departure time for each intermediate stop.

The pre-day models for GBA travelers have been estimated using Massachusetts Travel Survey (MTS) data [26], which include activity diaries for 33,000 individuals belonging to 15,000 households. Days are divided into four different sections: AM peak, PM peak, mid-day, and nighttime. Activities in MTS were categorized into six main categories, which are work, education, shopping, personal activities, recreation, and escort [27]. Individual travel behavior is validated using the data from future mobility survey in GBA.

C. Within-day

The within-day models take the generated DAS from pre-day and modify it throughout the day to account for the changes in schedule due to event triggers, such as incidents, information, controls, or incentives. As the simulation proceeds, the schedules of the individuals are monitored. If an individual is on schedule, their activities and travel are simulated as per the schedule generated in pre-day. Otherwise if the individual is off-schedule, their schedule is recomputed for the rest of the day based on a framework similar to pre-day, but adjusted for information on the current state [27].

Figure 3 shows the times at which people make their trips classified by tour types. With the exception of under-representing the personal tour peak during the day which is probably due to the fact that the time of day model decides the time in a hierarchical order (personal is decided after work and education), the model successfully captured the time of day choices for other tour purposes.

IV. SUPPLY MODELS

The supply model consists of the road network (i.e., geometry and traffic performance characteristics), the transit

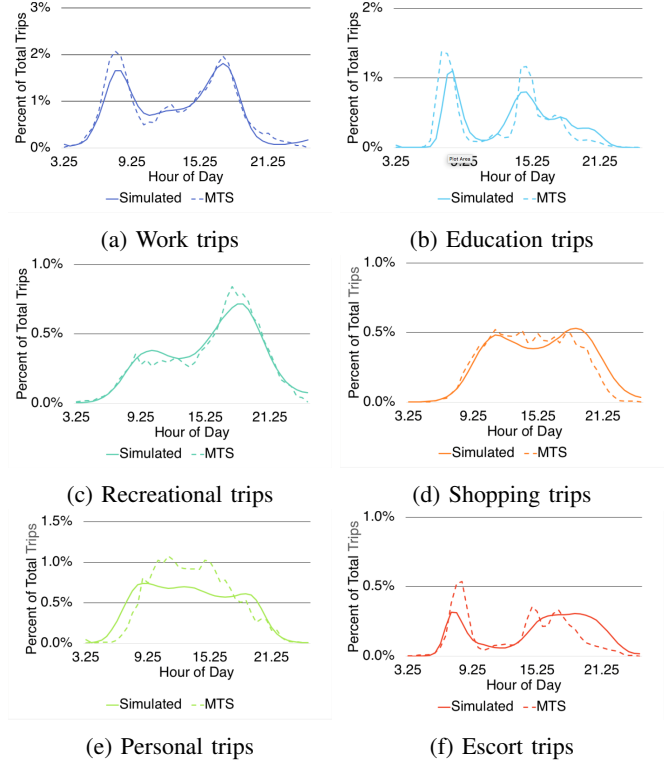


Fig. 3: Tours by time-of-day

network (i.e., bus and train routes, stops/stations and schedules) and the mobility service controllers (for the operation and decision making of bus, rail and on-demand mobility services).

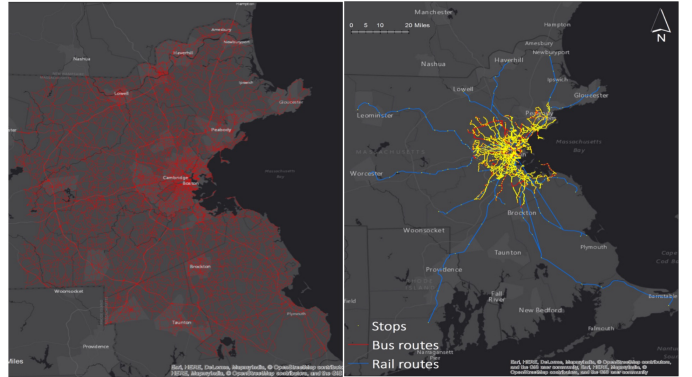


Fig. 4: Supply network

The road network in SM is represented by a hierarchical structure composed of links, segments, lane groups, lanes, connectors, and turnings. The vehicle simulation in the SM supply simulator is mesoscopic where the segment is the basic processing unit. Each segment represents a section of homogeneous roadway which is further divided into two traffic flow regions: the moving area and the queuing area [9]. Vehicles in the moving area travel at some uniform positive speed determined by a segment-specific, predefined macroscopic speed-density function. Vehicles in the queuing area form a horizontal queue whenever the arrival rate of the

traffic flow exceeds the capacity of the segment. In terms of resolution, the network is detailed up to the point of small local streets. The final SM network for GBA has 17,817 nodes, 44,672 links, and 164,802 segments (see Figure 4).

Buses are also simulated to include bus stop related movements and is composed of four steps: lane selection, entering bus stop, boarding and alighting, and re-joining traffic flow. Since our mesoscopic simulator is able to simulate vehicle movement, the interaction between buses and the main flow of traffic is also considered under this framework (such as queue spillback). As a result, the impact of crowdedness, waiting times, and denied boardings on traffic can be captured, which is important for assessing the effectiveness of incentives and controls.

The rail controller was developed for SM as a mesoscopic simulator with three different components: (1) train vehicles agents with predefined capacities that constantly accelerate or decelerate towards the defined speed limit; (2) a controller that dynamically defines all movement properties (dwell time, speed limits, dispatching) based on internal predefined control functions and the current state of the network; and (3) agents assigned with the rail mode in their DAS trips; these agents are considered as rail users and are allowed to interact with the rail controller and the train vehicle.

The rail network is constructed in the simulator as set of station platforms as nodes connected by lines and poly-lines as tracks. The attributes of a platform include the geographical location on the map and its associated walk times. Train routes and schedules are defined in the SimMobility database as an input to the controller (as a sequence of track blocks, train stops and train dispatching frequency). The simulated transit network includes 8,387 stops/stations, 12 commuter rail routes, 13 rapid transit routes, and 183 bus routes.

On-demand mobility services is simulated to allow multiple and different services to operate at the same time, e.g., regular taxi services and on-call services (Uber, Lyft). The agents involved in these services are controllers (one per each service), drivers, and passengers. Passengers can request a trip by either hailing a taxi on the street or by sending a request to one of the available controllers. Drivers can both operate as taxi drivers or subscribe at the same time to different on-call services. While operating taxis, drivers can pick up passengers hailing on the street, without waiting for controller instructions. If subscribed to an on-call service, they can also accept instructions from the controller.

Figure 5 shows zone to zone vehicle flows validation between results from simulator and Massachusetts Department of Transportation (MassDOT) count data, and zone to zone vehicle travel time validation between results from simulator and Uber travel time data in three time intervals. With the exception of some data points, which maybe because of different route choices, the simulator successfully reflected real world traffic situations.

V. ENERGY MODELS

The energy estimator is based on the TripEnergy model [2, 28], which matches lower-resolution data on vehicle trips

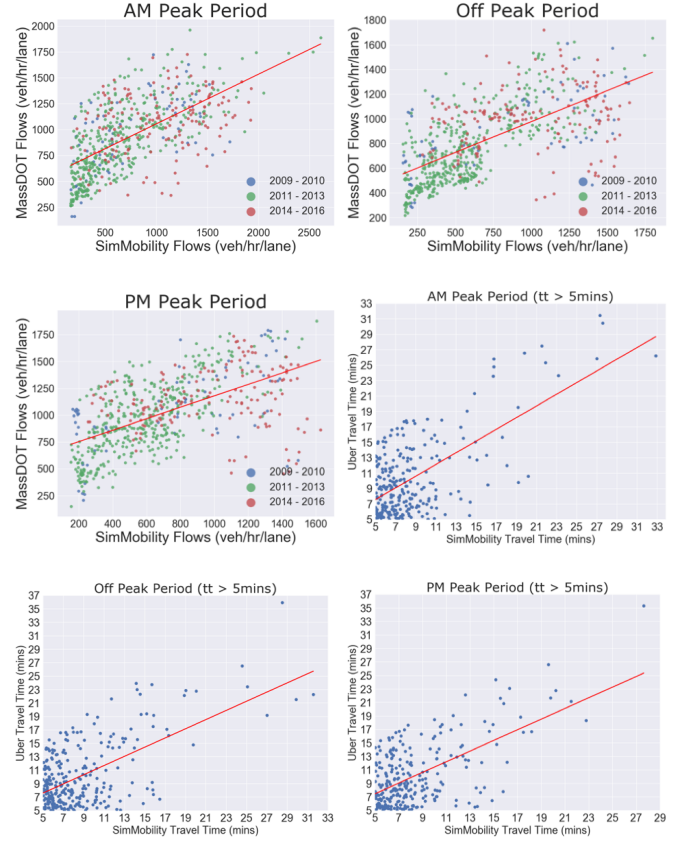


Fig. 5: Zone to zone flows and travel time validation

with a less representative set of high-resolution GPS data from real-world driving. Here, the low resolution inputs are the timestep average speeds for each vehicle simulated, and the high resolution data are portions of 1-Hz velocity histories from over 100,000 trips, accounting for over 28,000 hours of real-world driving. Before simulation, these high resolution trajectories are divided into 5 second partial trajectories and stored in a compressed format such that they can be easily accessed by the simulation on the fly [28].

These high-resolution partial trajectories are matched with the simulated vehicles movement over a timestep based on three inputs: (1) the average velocity of a given vehicle for the 5-second time period being estimated, (2) the average velocity over the preceding time periods and (3) the average velocity over the following time periods. These three quantities are each binned, together pointing to a bin containing a set of aggregate driving behavior parameters that have been calculated from all matched partial trajectories taken under similar conditions. This procedure is described in detail in [28]. These high-resolution trajectory characteristics are used to estimate the vehicle's tractive and braking energy use (E_{accel} and E_{brake} , respectively) over the timestep, which are combined with the vehicle-specific efficiency parameters for the total vehicle net energy consumption estimation:

$$E_{tot} = E_{accel}/\eta_{drive} - E_{brake}\eta_{brake} + P_{idle}T \quad (1)$$

where P_{idle} is the idling power consumption, η_{drive} is the

peak efficiency between the energy storage device (gas tank or battery) and the wheels, η_{brake} is the average efficiency between the brakes and the energy storage device for vehicles with regenerative braking—parameters that are calibrated off-line in advance based on fuel economy test result data—and where T is the trip duration. The current vehicle attributes database, including vehicle mass and coastdown coefficients needed for dynamometer tests and the unadjusted results of the CAFE tests, considers approximately 10,000 vehicle types tested by the EPA with model years 2010-2018.

VI. RESULTS

The complete integrated simulation takes about 22 hours to run with 20 cores which can be improved when additional resources are available given the parallel implementation in SimMobility. A total daily personal vehicle energy consumption of 531 GJ (which is equivalent to 0.95 gallons of gasoline per capita) was estimated from a 24-hour simulation of the GBA network. This equates to an average fuel economy of 25.8-MPG, which is slightly higher than the average fleet wide fuel economy in the U.S. for 2015 of 22.0-MPG. This discrepancy may be due to localized fleet differences in GBA compared to the average U.S. fleet as well as increased penetration of alternative fuel and higher fuel economy vehicles since 2015.

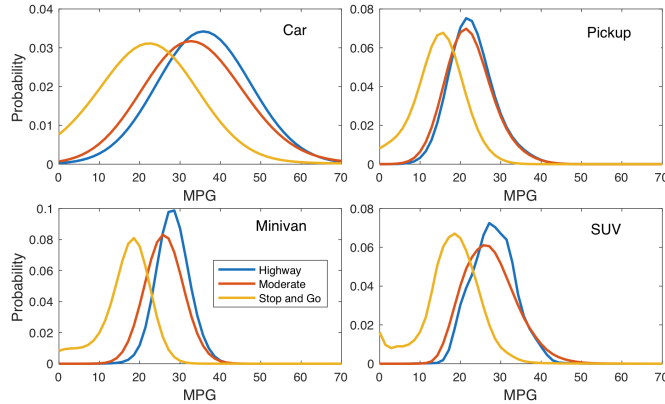


Fig. 6: Fuel economy for different vehicles, by vehicle class.

The energy estimates for vehicle class matched the expectations of vehicles performing better on highways compared to stop-and-go travel, shown in Figure 6. Compared to trucks, minivans, and SUVs; cars result in the greatest variability in their energy estimates due to the wide range of vehicle power trains, such as high-performance sport cars versus hybrid and electric vehicles. Age groups also yielded varying levels of fuel economy due to driving style (e.g. fast driving younger drivers) and off-peak travel (e.g. older drivers). In Figure 7 it would appear that suburban areas provide higher fuel economy; however, this is likely due to less congestion and higher traffic speeds experienced in these low density areas.

VII. CONCLUSIONS

The integration of the state-of-the-art models of TripEnergy and SimMobility offers a new avenue of research

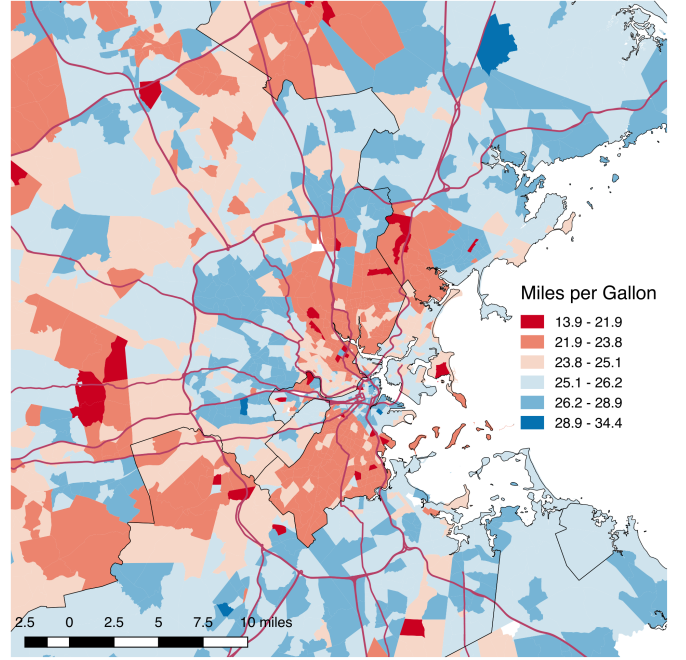


Fig. 7: Average fuel economy of overall Traffic Analysis Zones (TAZs) in the GBA

potential to evaluate real-time system response. Although the tests performed need further improvements in terms of calibration, runtime performance, and overall scalability, the paper present a framework to evaluate systemic effects—in both mobility and energy—of policies and controls. Further simulations are being carried out to test the performance of the proposed model when incidents and disruptions are present in the network. Next steps will include integrating of the simulation platform into a larger real-time optimization framework that minimizes energy and emissions through incentivized behavioral change.

ACKNOWLEDGMENT

This work was funded by the U.S. Department of Energy, Advanced Research Projects Agency-Energy (ARPA-E) under the Tripod Project, Grant Number: DE-AR0000611. The authors would like to thank K. Basak, R. Seshadri and T. Chu Yaw at SMART for their important support, and M. Danaf, and R. Basu for their contributions to the initial GBA model.

REFERENCES

- [1] M. Adnan, F. C. Pereira, C. M. L. Azevedo, K. Basak, M. Lovric, S. Raveau, Y. Zhu, J. Ferreira, C. Zegras, and M. Ben-Akiva, "Simmobility: A multi-scale integrated agent-based simulation platform," in *95th Annual Meeting of the Transportation Research Board*, 2016.
- [2] J. McNerney, Z. Needell, M. Chang, M. Miotti, and J. Trancik, "Tripenergy: Estimating personal vehicle energy consumption given limited travel survey data," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2628, pp. 58–66, 2017.
- [3] Y. Lu, M. Adnan, K. Basak, F. C. Pereira, C. Carrion, V. H. Saber, H. Loganathan, and M. E. Ben-Akiva, "Simmobility mid-term: A state of the art integrated agent based demand and supply model," in *94th Annual Meeting of the Transportation Research Board*, 2015.

- [4] I. M. Berry, "The effects of driving style and vehicle performance on the real-world fuel consumption of us light-duty vehicles," Ph.D. dissertation, Massachusetts Institute of Technology, 2010.
- [5] T. Yuksel and J. J. Michalek, "Effects of regional temperature on electric vehicle efficiency, range, and emissions in the united states," *Environmental science & technology*, vol. 49, no. 6, pp. 3974–3980, 2015.
- [6] P. Waddell, "Urbansim: Modeling urban development for land use, transportation, and environmental planning," *Journal of the American planning association*, vol. 68, no. 3, pp. 297–314, 2002.
- [7] T. Arentze, F. Hofman, H. van Mourik, and H. Timmermans, "Albatross: multiagent, rule-based model of activity pattern decisions," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1706, pp. 136–144, 2000.
- [8] T. Bellemans, B. Kochan, D. Janssens, G. Wets, T. Arentze, and H. Timmermans, "Implementation framework and development trajectory of feathers activity-based simulation platform," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2175, pp. 111–119, 2010.
- [9] Y. Lu, R. Seshadri, F. Pereira, A. O'Sullivan, C. Antoniou, and M. Ben-Akiva, "Dynamit 2.0: Architecture design and preliminary results on real-time data fusion for traffic prediction and crisis management," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*. IEEE, 2015, pp. 2250–2255.
- [10] L. Smith, R. Beckman, and K. Baggerly, "TRANSIMS: Transportation analysis and simulation system," Los Alamos National Laboratory (LANL), Los Alamos, NM, Tech. Rep., 1995.
- [11] C. Bhat, J. Guo, S. Srinivasan, A. Pinjari, N. Eluru, R. Coperman, and I. Sener, "The comprehensive econometric microsimulator for daily activity-travel patterns (cemdap), report 4080-s, prepared for the texas department of transportation," *Center for Transportation Research, The University of Texas at Austin*, 2006.
- [12] M. Balmer, M. Rieser, K. Meister, D. Charypar, N. Lefebvre, and K. Nagel, "Matsim-t: Architecture and simulation times," in *Multi-agent systems for traffic and transportation engineering*. IGI Gl., 2009, pp. 57–78.
- [13] C. Samaras, L. Ntziachristos, and Z. Samaras, "Copert micro: a tool to calculate the vehicle emissions in urban areas," in *5th Transport Research Arena: Transport Solutions from Research to Deployment*, 2014.
- [14] T. Markel, A. Brooker, T. Hendricks, V. Johnson, K. Kelly, B. Kramer, M. O'Keefe, S. Sprik, and K. Wipke, "Advisor: a systems analysis tool for advanced vehicle modeling," *Journal of power sources*, vol. 110, no. 2, pp. 255–266, 2002.
- [15] H. C. Frey and B. Liu, "Development and evaluation of simplified version of moves for coupling with traffic simulation model," in *Transportation Research Board 92nd Annual Meeting, Washington, D.C.*, 2013.
- [16] H. Rakha, K. Ahn, and A. Trani, "Development of vt-micro model for estimating hot stabilized light duty vehicle and truck emissions," *Transportation Research Part D: Transport and Environment*, vol. 9, no. 1, 2004.
- [17] G. Scora and M. Barth, "Comprehensive modal emissions model (cmem), version 3.01," *User guide. Centre for Environmental Research and Technology. University of California, Riverside*, vol. 1070, 2006.
- [18] X. Zhou, J. Liu, and Y. Qu, "Evaluating and calibrating emission impacts of traffic management strategies through simplified emission estimation model and mesoscopic dynamic traffic simulators," 2016.
- [19] S. Zegeye, B. De Schutter, J. Hellendoorn, E. Breunese, and A. Hegyi, "Integrated macroscopic traffic flow, emission, and fuel consumption model for control purposes," *Transportation Research Part C: Emerging Technologies*, vol. 31, pp. 158–171, 2013.
- [20] W. E. Deming and F. F. Stephan, "On a least squares adjustment of a sampled frequency table when the expected marginal totals are known," *Annals of Mathematical Statistics*, vol. 11, no. 4, pp. 427–444, 1940.
- [21] N. Fournier, E. Christofa, A. Akkinipally, and C. L. Azevedo, "An integration of population synthesis methods for agent-based microsimulation," in *Transportation Research Board 97th Ann. Meeting*, 2018.
- [22] U.S. Census Bureau, "5-year American Community Survey Tables," 2015. [Online]. Available: <https://www.census.gov/data.html>
- [23] U.S. Census Bureau American Community Survey, "2011-2015 ACS 5-year PUMS," 2015. [Online]. Available: <https://www.census.gov/data.html>
- [24] U.S. Census Bureau Center for Economic Studies, "Longitudinal Employer-Household Dynamics," 2015. [Online]. Available: <https://lehd.ces.census.gov/data.html>
- [25] J. L. Bowman and M. E. Ben-Akiva, "Activity-based disaggregate travel demand model system with activity schedules," *Transportation research part a: policy and practice*, vol. 35, no. 1, pp. 1–28, 2001.
- [26] MassDOT, "Massachusetts Travel Survey," 2012. [Online]. Available: <http://www.massdot.state.ma.us/>
- [27] I. Viegas de Lima, M. Danaf, A. Akkinipally, C. L. Azevedo, and M. Ben-Akiva, "Modelling framework and implementation of activity- and agent-based simulation: Application to the greater boston area," in *Transportation Research Board 97th Ann. Meeting*, 2018.
- [28] Z. Needell and J. Trancik, "Efficiently simulating personal vehicle energy consumption in mesoscopic transport models," *Transportation Research Record: Journal of the Transportation Research Board*, 2018.