

DEMAND CALIBRATION OF MULTIMODAL MICROSCOPIC TRAFFIC SIMULATION USING W-DSPSA

Simon Oh, Corresponding Author

Singapore-MIT Alliance for Research and Technology
1 CREATE Way, #09-02 CREATE Tower, Singapore 138602
Tel: +65-6601-1634 Fax: +65-6684-2118; Email: simon@smart.mit.edu

Ravi Seshadri

Singapore-MIT Alliance for Research and Technology
1 CREATE Way, #09-02 CREATE Tower, Singapore 138602
Tel: +65-6601-1634 Fax: +65-6684-2118; Email: ravi@smart.mit.edu

Carlos Lima Azevedo

Technical University of Denmark
Department of Management Engineering, 2800 Kgs. Lyngby, Denmark
Email: climaz@dtu.dk

Moshe E. Ben-Akiva

Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA
Tel: 617-253-5324; Fax: 617-253-1130; Email: mba@mit.edu

Word count: 5,245 words text + 2 tables / 7 figures x 250 words (each) = 7,495 words

Submission Date: 1st August 2018

ABSTRACT

This paper presents a stochastic approximation framework to solve a generalized problem of off-line calibration of demand for a multimodal microscopic (or mesoscopic) network simulation using aggregated sensor data. A key feature of this problem is that demand, although typically treated as a continuous variable is in fact discrete, particularly in the context of agent-based simulation. To address this, first, we use a discrete version of the weighted simultaneous perturbation stochastic approximation (termed W-DSPSA hereafter) algorithm for minimizing a Generalized Least Squares (GLS) objective (that measures the distance between simulated and observed measurements) defined over discrete sets. The algorithm computes the gradient at each iteration using a symmetric discrete perturbation of the calibration parameters and a multimodal weight matrix to improve accuracy of the gradient estimate. The W-DSPSA algorithm is then applied to the large-scale calibration of multimodal OD flows (including private vehicle (*PVT*) and public transit trips (*PT*)) in a microscopic network simulation model of Singapore. The results indicate that an acceptable margin of error on the vehicle loop count (*VLC*) and bus passenger count (*BPC*) are achieved at convergence with an improvement of 60~80% in root mean squared errors. Lastly, we validate the calibration results with observed travel-times on the network. Statistical comparison shows good agreements on both point-to-point travel-time (*PTT*) and public bus's stop-to-stop ride-time (*SRT*) with the field observations.

Keywords: Calibration and validation, Weighted Discrete Simultaneous Perturbation Stochastic Approximation, Multimodal demand estimation, SimMobility Short-term

1. INTRODUCTION

1.1 Off-line Calibration of Traffic Simulation Models

During the last few decades, extensive research has been devoted to the development of traffic simulation models that capture traffic dynamics at varying levels of resolution to accurately replicate the real-world. Simulation tools have important applications in understanding the implications of future scenarios in advance, enabling transportation planners and operators to estimate the impact of urban policy and operational decisions on the transportation system. Recognizing this, the Federal Highway Administration has funded the development of traffic simulation through the NGSIM program which has already played an important role in various policy assessments (1). Recently, emerging trends in mobility technologies and services (such as automated mobility-on-demand (2), ride-hailing services) have motivated the development of more accurate and high-fidelity traffic simulation models so as to evaluate and assess these mobility-sensitive scenarios (3).

A crucial prerequisite in carrying out policy and operational assessments is the ability to generate robust, accurate and reliable simulation results, for which, calibration is essential. Calibration is the process of estimating simulation parameters by matching simulated measurements against observed measurements, and has been traditionally formulated as an optimization problem where the objective function is based on goodness-of-fit measures (4).

An important input for traffic simulation is time-dependent OD (origin-destination) flows. This set of dynamic demand parameters has successfully been calibrated with traffic simulators using various methods: the traditional generalized least squares approach (5), bayesian inference (6), marginal computation with kinematic wave theory (7), principle component analysis (7). The most commonly used approaches for solving optimization formulations of the calibration problem are gradient descent methods such as the Newton–Raphson method (9) and steepest descent method (10). Another widely used approach is the finite difference stochastic approximation (FDSA) algorithm (11) which is known to be suitable given the non-linear nature of calibration problem and the presence of measurement errors. A key challenge with the FDSA algorithm lies in maintaining computational tractability given the expensive nature of gradient computations. To address this, the simultaneous perturbation stochastic approximation (SPSA) was proposed (12) which significantly improves computational performance by perturbing all parameters simultaneously during gradient computation. The SPSA algorithm has received significant attention because of efficiency (only two evaluations of the objective function are required in each iteration), applicability (simple to implement in high-dimensional real-world problems), accuracy (properly handles stochasticity and measurement errors), and has been applied to the off-line demand calibration problems by various researchers (13, 14, 15, 16).

However, recent work has identified an issue of unstable convergence of the SPSA algorithm in large-scale applications arising from the gradient approximation (17, 18, 19). To overcome this, the c-SPSA (cluster-wise SPSA) and W-SPSA (Weighted SPSA) algorithms were proposed. The c-SPSA (17) approximates the gradient within a small number of homogeneous clusters one at a time. The W-SPSA (18, 19) incorporates spatio-temporal correlations between parameters and measurements in approximating the gradient thereby minimizing the impact of uncorrelated measurements. The enhanced algorithms outperform the original SPSA showing improved robustness and convergence.

In summary, the literature has focused largely on solving the calibration problem using stochastic approximation (SA) algorithms in the continuous domain. However, certain calibration parameters are discrete valued (Eg: OD demands) and the discrete nature of these

variables is not explicitly handled. This paper contributes to the existing literature by presenting a solution approach using the discrete version of W-SPSA (W-DSPSA) which is appropriate when the domain of the objective function is defined on a discrete set of points. A case study on a large-scale multimodal network using a traffic microsimulator examines performance of the W-DSPSA algorithm. The motivation is to view the demand calibration problem from the perspective of a discrete resource allocation problem (20, 21). Some researchers have presented mathematical formulations for the discrete version of SPSA (22, 23, 24), however these have not been applied to calibration problems within the transportation domain to the best of our knowledge.

1.2 Multimodal Microscopic Traffic Simulation

SimMobility is a multi-level, activity-based, multi-modal simulation platform (25) which consists of three integrated simulators operating at different time scales. The short-term simulator (ST) functions at the operational level and simulates movement of agents at a microscopic granularity (26). The mid-term (day-to-day) simulator (MT) handles transportation demand; it simulates agents' behaviour which includes their activity and travel patterns and the movement of vehicles at a mesoscopic level (27). The long-term (year-to-year) model (LT) captures land use and economic activity, and predicts the evolution of land use and property development and use, the associated life cycle decisions of agents, and accounts for interactions among individuals and firms (28). This modelling approach allows a full integration of all individual choices and the reliable simulation of its impacts in the transportation system under future mobility scenarios. Although various simulators have been developed for each of the different time-scales (such as UrbanSim (29), DynaMIT (30), MITSIMLab (31)), these have not been integrated in a coherent manner. SimMobility is unique in that the same pool of agents is used across all timeframes: agents' long-term behavior is already established when their behavior is modelled in the mid-term/short-term simulation. SimMobility has been applied to explore how future scenarios induce shifts in the distribution of people, activities, land use, and transportation network performance in several contexts: autonomous mobility-on-demand (32, 33), freight (34), public transit (35, 36), pricing (37).

This paper focuses on the demand calibration of SimMobility Short-term (ST) which simulates the high-resolution movement of agents (traffic, transit, pedestrians and goods) and the operation of different mobility services and control systems. This simulator comprises three main components: microscopic traffic simulator, control and operation systems, and communication network simulator (details on modeling framework and system architecture may be found in (26)). The microscopic traffic simulator includes modeling of driving behavior (car-following, lane-changing, intersection and merging) and travel behavior (route-choice for private vehicle and public transit trips). SimMobility ST also includes the detailed modeling of the public transit system through a bus controller agent which is responsible for scheduling and dispatching of buses, and bus driver agents that manage bus movement near and between bus stops, dwelling at bus stops, and the monitoring of real-time occupancy. Bus stop agents on the other hand are responsible for passenger boarding and alighting at bus stops and bus occupancy at stops. SimMobility ST uses a mode-specific activity-based demand representation generated by SimMobility Mid-term (MT) instead of the traditional OD flow based demand representation. The trip-chain represents a sequence of multimodal trips (sub-trips) to be simulated and the agent's role (pedestrian, passenger, and driver) determines role-specific behavioral characteristics such as aggressiveness, yielding behavior, look-ahead distance, reaction-time, walking speed. For the multimodal OD estimation in this paper, we aggregate this trip-activity

schedule (trip-chain) into the form of an OD vector and then perform a disaggregation into the trip-chain representation after the calibration process is completed.

The rest of the paper is organized as follows. Section 2 formulates the off-line multimodal demand calibration problem and Section 3 applies a discrete version of the w-SPSA solution algorithm to solve the problem. Section 4 applies the framework and solution algorithm to the calibration of SimMobility ST for a sub-network in Singapore and finally, Section 5 concludes this paper.

2. PROBLEM FORMULATION

The general off-line demand calibration problem is formulated as an optimization problem to determine the simulation parameter vector (θ) that minimizes the objective function (Z):

$$\min Z(\theta) \quad (1)$$

$$\text{s.t. } \theta \in \Theta \quad (2)$$

where Θ is the domain of the parameter vector θ .

In this study, we define Z as an aggregate error between simulated and observed measurements and θ is the time-dependent multimodal Origin-Destination (OD) demand matrix. Let $\mathbb{N} = \{1, 2, \dots, N\}$ denote the set of the transportation modes, $\mathcal{H} = \{1, 2, \dots, H\}$ the collection of time intervals in the simulation period, and $\mathcal{P}_n = \{1, 2, \dots, P_n\}$, the set of unique OD pairs for mode n , where $n \in \mathbb{N}$. Thus, the time-dependent OD vector (θ) is composed of elements $\theta_{n, p_n, h}$ and has a dimension of $\sum_{n=1}^N P_n * H$. Further, let $\mathcal{M}_n = \{1, 2, \dots, M_n\}$ denote the set of sensors associated with mode $n, n \in \mathbb{N}$. Sensors could, for example, refer to loop detectors measuring vehicle counts or the recording of transit traveler tap-ins and tap-outs at stops/stations using smart cards. The measurement vector which contains measurements obtained from the specified sensors is denoted by $M = (M_{n,m,h}); n \in \mathbb{N}, m \in \mathcal{M}_n, h \in \mathcal{H}$ and is composed of time-dependent entries for each individual sensor over the simulation time period \mathcal{H} ; hence, we have $M_n * H$ elements for each mode in \mathbb{N} . The superscripts ‘Obs’ and ‘Sim’ are used to denote observed and simulated measurements respectively. Thus, we have the generalized formulation:

$$\min_{\theta} Z(\theta) = \sum_{n=1}^N \sum_{m=1}^{M_n} \sum_{h=1}^H \bar{Z} \left(M_{(n,m,h)}^{\text{Obs}}, M_{(n,m,h)}^{\text{Sim}} \right) \quad (3)$$

$$\text{s. t. } \mathbf{lb} \leq \theta \leq \mathbf{ub}; M^{\text{Sim}} = f(\theta, \beta)$$

where $\bar{Z} \left(M_{(n,m,h)}^{\text{Obs}}, M_{(n,m,h)}^{\text{Sim}} \right)$ is a measure of the error between $M_{(n,m,h)}^{\text{Obs}}$ and $M_{(n,m,h)}^{\text{Sim}}$. The constraints involve upper and lower bounds on the OD demands and the mapping of measurements to the simulation parameters through the simulator denoted by $f()$. In addition to the OD demands, the simulator takes as an input other parameters such as driver behavior and route choice which are denoted by β . The function $Z(\theta)$ is defined using a generalized least squares objective function expressed as:

$$Z(\theta) = \epsilon_M^T \Omega_M^{-1} \epsilon_M \quad (4)$$

where $\epsilon_M = \left(M_{(n,m,h)}^{\text{Obs}} - M_{(n,m,h)}^{\text{Sim}} \right)$, $\forall n \in \mathbb{N}, m \in \mathcal{M}_n, h \in \mathcal{H}$ is a vector containing the differences between the observed and simulated measurements, Ω_M^{-1} is the inverse of the variance-covariance matrix associated with the measurement vector M . The matrix Ω_M^{-1} in effect provides suitable weights to the different measurements in the objective function.

In the case study presented in Section 4, the set of transport modes include private vehicles (pvt) and public transport buses (pt), i.e. $\mathbb{N} = \{pvt, pt\}$. The OD vector is defined as: $\theta_{OD} = [\theta_{pvt}; \theta_{pt}]$, having $P_{pvt} * H_{pvt}$ and $P_{pt} * H_{pt}$ elements respectively. Note that the time interval of the demand vector may differ from the time interval of the measurements. Furthermore, each type of measurement may be affected by only part of the OD vector: for example, θ_{pvt} (PVT) and θ_{pt} (PT) will affect the vehicle loop count (VLC) and bus passenger count (BPC) respectively.

3. PROPOSED SOLUTION

3.1 DSPSA: Discrete version of SPSA

SPSA has been widely applied in the context of offline calibration involving stochastic simulators with a large number of parameters. SPSA is a type of stochastic approximation algorithm that determines the optimal solution through the iterative update rule:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k) \quad (6)$$

where $\hat{g}_k(\hat{\theta}_k)$ is an unbiased estimate of the gradient at the estimated point of the decision vector in the k^{th} iteration ($\hat{\theta}_k$). a_k is computed based on algorithm parameters A, a, α as,

$$a_k = \frac{a}{(A+k+1)^\alpha} \quad (7)$$

The gradient is computed using two evaluations of the objective function in the k^{th} iteration as:

$$\hat{g}_k(\hat{\theta}_k) = \frac{z(\hat{\theta}_k^+) - z(\hat{\theta}_k^-)}{2c_k} = \frac{z(\hat{\theta}_k + c_k \Delta_k) - z(\hat{\theta}_k - c_k \Delta_k)}{2c_k} \quad (8)$$

where, Δ_k is a random perturbation vector following a Bernoulli process taking either -1 or 1 with probability of 0.5. c_k is computed based on algorithm parameters (c, γ) as,

$$c_k = \frac{c}{(k+1)^\gamma} \quad (9)$$

The standard SPSA algorithm assumes a continuous problem domain. However, demand calibration involves the allocation of discrete number of trips to a finite set of OD pairs and hence, involves a discrete problem domain. Recently, a discrete version of SPSA (DSPSA) has been proposed by researchers to deal with discrete optimization problems such as resource allocation. This DSPSA algorithm has been tested and compared with the other algorithms including the stochastic ruler and stochastic comparison (24). The main idea is to approximate the gradient for a discrete function L by taking the middle point between the ceiling ($\lceil \theta \rceil$) and floor ($\lfloor \theta \rfloor$) integer points (24).

$$\theta_{original}^\pm = \pi(\theta) \pm \frac{1}{2} \quad (10)$$

where, $\pi(\theta) = \lfloor \theta \rfloor + \frac{1}{2}$. However, the perturbed parameters ($\theta_{original}^+, \theta_{original}^-$) always take the value between θ and $\theta + 1$ if θ is an integer value which is asymmetric resulting a biased

gradient to the plus side. Hence, we modify this fomula to derive a symmetric perturbation while maintaining the discrete domain as:

$$\theta_{revised}^{\pm} = \left(\lfloor \theta \rfloor \pm \frac{1}{2} \right) \pm \frac{1}{2} \quad (11)$$

With the perturbation parameter (Δ_k), the gradient at k^{th} iteration is approximated as:

$$\hat{g}_k(\hat{\theta}_k) = \frac{Z\left(\left(\lfloor \hat{\theta}_k \rfloor + \frac{1}{2} \text{sgn}(\Delta_k)\right) + \frac{1}{2} \Delta_k\right) - Z\left(\left(\lfloor \hat{\theta}_k \rfloor - \frac{1}{2} \text{sgn}(\Delta_k)\right) - \frac{1}{2} \Delta_k\right)}{\Delta_k} \quad (12)$$

where, $\text{sgn}(\Delta)$ yields the sign of Δ . Then, we update the parameter vector for next iteration ($k+1$) according to Eq.6.

3.2 W-DSPSA: Incorporating Multimodal Weight Matrix

Although the SPSA algorithm has been successfully applied to the calibration of DTA models, it has been observed that its accuracy deteriorates as the problem scale increases in terms of network size and time intervals due to a gradient ‘approximation error’ (18). To decrease this approximation error, researchers (18, 19) have introduced the concept of a weight matrix (W) to incorporate the spatial and temporal correlations between OD parameters and measurements which are determined by network topology and route choice behavior. Accordingly, each component ($w_{p+P \times (h_1-1), m+M \times (h_2-1)}$) in the weight matrix represents the relative correlation between the p^{th} OD parameter at the interval h_1 and the m^{th} individual sensor measurement in the interval h_2 . We extend this weight matrix W to the multimodal dimension ($(P_{pvt} * H_{p_{pvt}} + P_{pt} * H_{p_{pt}})$ by $(M_{pvt} * H_{M_{pvt}} + M_{pt} * H_{M_{pt}})$) to capture the correlations determined by the traveler’s route choice, departure time, and travel-time for both private vehicle and public transit trips. Note that the notation in Section 2 applies, i.e. P_{pvt}, P_{pt} denote the number of OD pairs associated with private vehicle and public transit trips respectively; $H_{p_{pvt}}, H_{p_{pt}}$ denote the number of time intervals of private vehicle and public transit OD demands; M_{pvt}, M_{pt} denote the number of sensors associated with private vehicle and public transit trips and $H_{M_{pvt}}, H_{M_{pt}}$ denote the corresponding number of time intervals for which measurements are available. Thus, we can express W within the same framework in (18) as:

$$W = \begin{bmatrix} w_{1,1} & \dots & w_{1,M} & \dots & w_{1,M * H_M} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{P,1} & \dots & w_{P,M} & \dots & w_{P,M * H_M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{P * H_P,1} & \dots & w_{P * H_P, M} & \dots & w_{P * H_P, M * H_M} \end{bmatrix} \quad (13)$$

where, $P = \{1, \dots, P_{pvt} * P_{pvt} + 1, \dots, P_{pvt} + P_{pt}\}$; $M = \{1, \dots, M_{pvt}, M_{pvt} + 1, \dots, M_{pvt} + M_{pt}\}$; $H_P = \{1, \dots, H_{p_{pvt}}, 1, \dots, H_{p_{pt}}\}$; $H_M = \{1, \dots, H_{M_{pvt}}, 1, \dots, H_{M_{pt}}\}$. Recall that M_{pvt} and M_{pt} in this study are the vehicle loop count (VLC) and bus passenger count (BPC) respectively. Thus, the components that represent the correlations between the private vehicle trips with the BPC ($w_{\{1:P_{pvt} * H\}, \{(M_{pvt} * H + 1):(M_{pvt} * H + M_{pt} * H)\}}$) and the public transit trips with the VLC ($w_{\{(P_{pvt} * H + 1):(P_{pvt} * H + P_{pt} * H)\}, \{1:M_{pvt} * H\}}$) would be zero.

The practical method to estimate the weight matrix for demand is to calculate the fraction of trips for each mode passing (or boarding/alighting) each individual sensor, commonly termed the assignment matrix. This relative correlation is calculated as in (18):

$$w_{p+P \times (h_1-1), m+M \times (h_2-1)} = \frac{d_{p, h_1 m, h_2}}{\sum_{i=1}^M \sum_{j=1}^H d_{p, h_1, i, j}} \quad (14)$$

where, $d_{p, h_1 m, h_2}$ denotes the number of trips between the p th OD pair departing at h_1 that influence m at h_2 . As noted previously, the weight matrix excludes uncorrelated measurements in the deviation vector during the gradient approximation. With the introduction of the weight matrix, Eq.12 is modified as:

$$\hat{g}_k(\hat{\theta}_k) = \frac{Z\left(\left(|\hat{\theta}_k| + \frac{1}{2} \text{sgn}(\Delta_k)\right) + \frac{1}{2} \Delta_k\right) - Z\left(\left(|\hat{\theta}_k| - \frac{1}{2} \text{sgn}(\Delta_k)\right) - \frac{1}{2} \Delta_k\right)}{\Delta_k} W_k \quad (15)$$

Specifically, the gradient for i^{th} parameter is calculated using the OLS (ordinary least squares) based deviations between observed and simulated measurements for the plus and minus side (ϵ_{Mkj}^{\pm}) which gives:

$$\hat{g}_{ki}(\hat{\theta}_k) = \frac{\sum_{j=1}^D \left[\left(\epsilon_{Mkj}^+ \right)^2 - \left(\epsilon_{Mkj}^- \right)^2 \right]}{\Delta_{ki}} W_{ki} \quad (16)$$

where, D is the number of measurements ($M * H_M$); W_i is the i^{th} row vector of W at iteration k . The W-DSPSA algorithm described here is applied to solve a large-scale demand calibration problem in the following sections.

4. IMPLEMENTATION

In this section, the W-DSPSA is applied to the demand calibration of the SimMobility ST model for a subnetwork in Singapore. The experimental setup and implementation procedure are first described followed by a discussion of calibration and validation results.

4.1 Experimental Setting

Study site and data

The network topology for Singapore is shown in Fig.1. The study area under consideration is the shaded area which consists of 153 TAZs (Traffic Analysis Zone) covering an area of 74.36 km² including business, commercial, and residential zones. The road network consists of 715 nodes and 1768 links. The traffic signals are modeled at the node level to replicate the SCATS-like signal system while bus stops are modeled at a segment/lane level (including bus dwelling, passenger boarding/alighting).

The signalized intersections are equipped with inductive loop detectors for aggregated VLC and public transit data is collected through the contactless smart card payment system for public transport fare collection in Singapore (called EZ-Link card). We extracted the aggregated BPC that includes passenger boarding and alighting over time at each bus stop. The Land Transport Authority (LTA) of Singapore provided both intersection loop count data and EZ-Link data collected from Aug/2013.



FIGURE 1 Network topology for study area in Singapore (Shaded).

Experimental setting

The simulation time period is 1 hour (along with an initial simulation warm-up period of 15min) from the morning peak during typical weekdays. The seed (apriori) OD demand matrix is obtained from the island-wide simulation of SimMobility MT which uses a synthetic population (28) and behavioral models estimated from the household interview travel survey data (HITS 2012). We extracted the relevant trips by classifying the trips into four types: internal, incoming, outgoing, and passing through trips. The private and public-transit ODs are defined at the level of network nodes. In case of the public-transit demand, a route choice model translates an origin-destination trip into a specific transit route which includes an access leg from the origin node to a specific bus or train stop, a transfer leg (if necessary), and an egress leg from a bus/train stop to the destination node. Accordingly, we have 7,289 and 8,819 OD pairs resulting in 29,156 and 105,828 time-dependent OD pairs for private (15min time intervals) and public transit trips (5min time intervals) respectively. Although in theory, we would expect 5 minute OD intervals to better capture temporal OD patterns and hence a yield better fit even if the measurements are aggregated at 15 minute intervals (as was the case with the available loop detector data), this comes at the cost of a significantly increased problem dimension and a potentially a larger number of iterations required for convergence. Moreover, it was observed in preliminary experiments that the 15 minute OD intervals yielded a reasonably good fit to the sensor counts and hence the choice of 15 minute intervals for the private vehicle ODs. In case of the public transit ODs however, it was observed that using 5 minute OD intervals resulted in a significantly better fit to both boardings and alightings compared to 15 minute OD intervals (possibly due to greater stochasticity in public transit route choice involving different bus lines, bus routes, departure time, crowdedness etc.) hence the choice of 5 minute intervals for transit ODs despite the increase in problem dimension. Other simulation parameters including driving (car-following, lane-changing, intersection behavior, etc.) and travel behavior (route choice) were calibrated previously for the Singapore context (26).

The proposed solution in Section 3 with the algorithm parameters (12) is applied to calibrate demand parameters using the sensor measurements (788 *VLCs* and 1,510 *BPCs*). The

discrepancy between the observed and simulated measurements is measured using the normalized root mean square error, RMSN. Note that the simulated measurements are subject to certain supply constraints including bus operation (controlled with the fixed schedule and route for 195 bus lines), bus capacity (number of seaters), road capacity, speed limits, etc. The details of the overall experimental setting are summarized in Table 1.

TABLE 1 Experimental Setting

Items	Details
Simulator	SimMobility Short-term
Study area	153 TAZs in Singapore (74.36 km ²) - 715 nodes, 1768 links
Simulation parameters	29,156 Time-dependent private vehicle OD flows (15min interval) 105,828 Time-dependent public transit OD flows (5min interval) Calibrated driving behavior (β) and travel behavior (δ)
Measurements	788 <i>VLC</i> measurements (15min interval) 1,510 <i>BPC</i> measurements (30min interval)
Goodness-of-fit (GoF)	Fit of observed (y_n) and simulated (\hat{y}_n) measurement. $RMSN = \sqrt{\frac{N \sum_{n=1}^N (y_n - \hat{y}_n)^2}{\sum_{n=1}^N y_n}}$
Calibration parameters	$\alpha = 0.602$, $a = 0.16$, $A = 50$

4.2 Implementation Process

This section describes the implementation of the calibration procedure for the SimMobility ST simulator. The procedure involves three main steps as shown in Fig.2: Initialization, Calibration, and Post-processing. Although the implementation process is complicated, we note that the initialization and post-processing steps are not computationally intensive in comparison with the solution algorithm itself (computational advantages of the SPSA algorithm are discussed in Section 1.1).

Initialization

The initializer extracts the trip-activity tables (which contain a set of trip chains TC_0) from the PostgreSQL database and prepares the initial run of the simulation by loading the network (Shape files), simulation configuration (with extensible markup language format), and sub-functions (the vector generator, evaluation function). The trip chain (TC_0) is transformed into the OD vector (θ_0) by aggregating the number of trips for each OD pair and time interval. Following this, SimMobility ST ($k = 0$) is executed with TC_0 and generates three measurement vectors: C_{sim} , B_{sim} , and A_{sim} associated with *VLC* and *BPC* for boarding and alighting respectively.

The vector generator merges the simulated vectors with observed data and generates the vector for each measurement: M_C , M_B , and M_A . This module is also responsible for computing the weight matrix (W) based on the assignment matrices (Eq.13) and to update the link travel-times to reflect the informed individual route choice decisions for the next iteration (after the initial ($k = 0$)). Next, the calibration loop is initiated followed by the estimation of the objective value (Z_0) with Eq.5 in the evaluation step.

Calibration loop

The calibration loop starts with calculating the perturbed vectors (θ_k^\pm) using Eq.11 based on the gain sequence (a_k) determined with calibration parameters (A, a, α) at the k^{th} iteration. With the perturbations, we run two simulations in parallel to generate two objective values from plus and minus side (Z_k^\pm) with the measurement vectors. Next, we estimate the updated vector (θ_{k+1}) in Eq.6 using by the two-sided gradient ($g_k(\theta_k)$) approximated by Eq.16. Finally we run the simulation (at k) with θ_{k+1} to update W and link travel-time for $k + 1$ and check the convergence criteria. Note that the vectors (θ_k^\pm, θ_k) are mapped to the trip-activity tables and are used to update the tables in the PostgreSQL through a killing and cloning process (26). The iterative process continues until the convergence condition is satisfied with the current objective value (Z_k). If convergence has been achieved (or the maximum number of iterations has been reached), we move forward to the post-processing and validation step.

Post-processing

Testing is required to determine the reliability and predictive power of the derived solution (θ_k^*). For validation purposes, the simulated and observed travel-times between origin and destinations (and the ride-time between stops boarding/alighting) will be compared to assess the area-wide accuracy. After that we generate and update the trip-activity schedule in the database which concludes the calibration procedure.

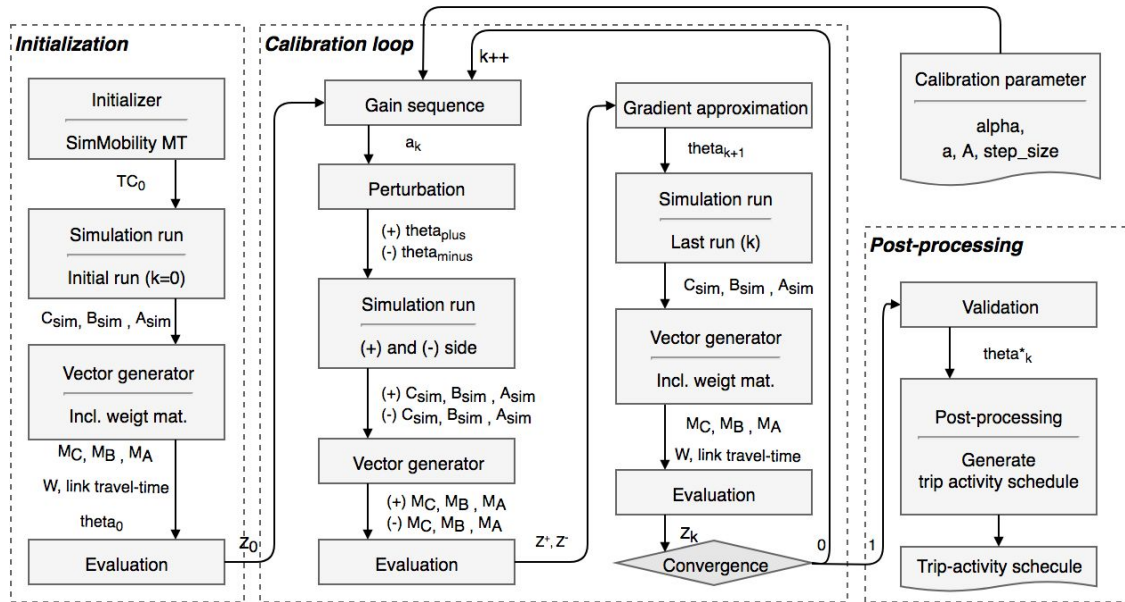


FIGURE 2 Calibration Implementation Procedure

4.3 Calibration Results

Calibration performance

The calibration results indicate that the Goodness-of-fit (GoF) measure decreases significantly over the iterations for both types of measurements (Fig.3). The RMSN improves from 0.81 and 3.5 to 0.32 and 0.56 at convergence for *VLC* and *BPC* respectively (an improvement of 60~80%). The initial seed matrix for the transit demands yield a significantly worse starting point for the *BPC* compared to the *VLC*, resulting in a larger number of iterations required for convergence (150 and 100 iterations, respectively). The relatively poor performance for the *BPC*s may be attributed to several reasons. First, although buses are operated based on a fixed

schedule, these are subject to change and the exact schedule on the days during which data is collected (which is not available) may differ from that used in the simulations. Second, although the PT route choice behavior was estimated using smart card data, these might require additional calibration. Finally, the smart card data does not provide exact origins and destinations and moreover, no information on waiting time which is an important factor affecting route choice behavior of transit travelers.

Fig.4 presents the fit-to-counts during the simulation period. On the left figure, the initial *VLCs* are underestimated compared to the field observations (grey) whereas at convergence, the overall measurements (black) show a high R-squared value of 0.858. The right figure shows a similar trend in *BPC*. The fits to the number of passengers boarding (blue) and alighting (red) show that the bus stops, where initially less trips (or none at all) are allocated, now fit the observed measurements significantly better (R-squared value: 0.986 (boarding) and 0.866 (alighting)).

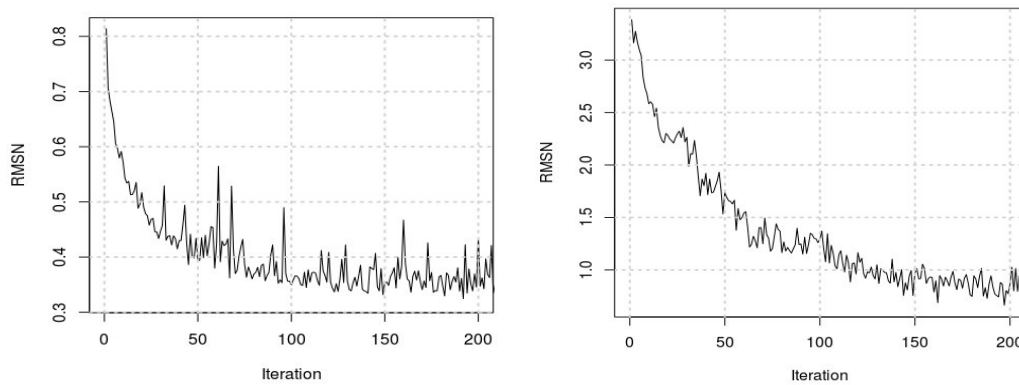


FIGURE 3 GoF over iterations (left: *VLC*, right: *BPC*).

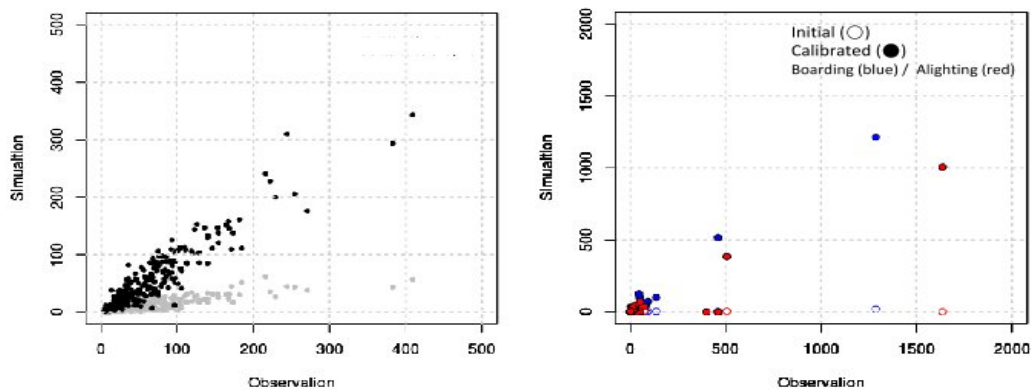


FIGURE 4 Fit-to-counts (left: *VLC* (veh/15min)), right: *BPC* (pax/30min)).

Characteristics of Calibrated OD Demands

Fig.5 shows the deviation of the calibrated OD demands from the initial points for *PVT* and *PT* and indicates that the initially underestimated seed OD demands have been appropriately tuned to achieve convergence. The extent of deviation with respect to the initial values have been

quantified in terms of RMSN and are 57.9 and 276.7 percent for *PVT* and *PT* respectively. The aggregated *PVT* demand gradually increases at the beginning of the peak period, while *PT* demand also depends on the bus dispatching frequency determining departure/arrival time of each line from/to each designated stop. The high demand reflected in the calibrated OD demands are consistent with the measurements. Through iterations, more trips have been allocated to those OD pairs relevant to the large traffic measured at sensors and a large number of boardings/alightings at bus stops and intersections located near major train stations of commercial centers within the study area.

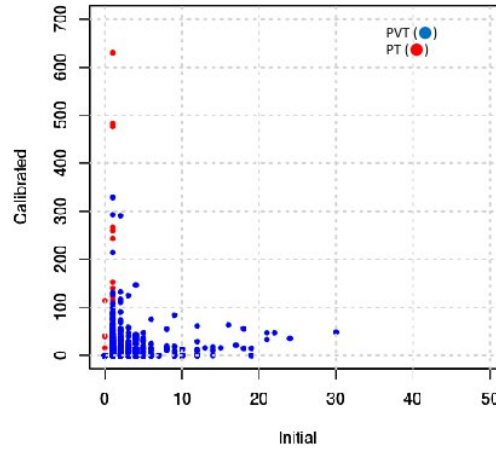


FIGURE 5 Calibrated ODs versus Initial values (red: *PVT*, blue: *PT*).

4.4 Validation

Using the estimated demand, this section reports results from a validation conducted using point-to-point travel-times (*PTT*) and passenger's stop-to-stop ride-time with public bus (*SRT*).

The field *PTT* is collected through taxi vehicles equipped with GPS sensors during Aug/2013. For the purposes of comparison, each GPS latitude and longitude (in world geodetic system) are mapped to network nodes to obtain travel-time measurements between origin and destination points. These travel-times are then compared against the corresponding OD travel times measured from the simulation, resulting in 1442 samples for each group. Fig.6 shows the overall distribution of *PTT* and gap between observed and simulated travel-time (Gap_{PTT}). To determine the statistical gap of two distributions, we conduct both parametric and nonparametric two-sided hypothesis tests. Based on the student's t-test, we fail to reject the null hypothesis defined with the sample mean ($H_0: \mu_{PTT_{Obs}} = \mu_{PTT_{Sim}}$), while the nonparametric test (using wilconxon's rank-sum) rejects the null hypothesis defined with sample median ($H_0: Q2_{PTT_{Obs}} = Q2_{PTT_{Sim}}$). We also measure the gap of two *PTTs* as: $Gap_{PTT} = PTT_{Obs_i} - PTT_{Sim_i} \forall i \in OD \text{ pair}$. It shows on average, a 1.58 min gap with about a 6min standard deviation (Table 2).

In terms of *SRT*, we match origin/destination bus stops for each sub-trip completed during the simulation with individual EZ-Link records. The overall distribution of *SRT* and its gap (Gap_{SRT}) is listed in Fig.7. The parametric and nonparametric statistical tests accept the null hypotheses ($H_0: \mu_{SRT_{Obs}} = \mu_{SRT_{Sim}}$ and $H_0: Q2_{SRT_{Obs}} = Q2_{SRT_{Sim}}$) in favor of the alternative hypothesis at the 99% confidence level. The ride-time gaps ($Gap_{SRT} = SRT_{Obs_i} - SRT_{Sim_i} \forall i \in Stop \text{ pair}$) on average are estimated at 2.25 min for 1143 samples (see Table 2).

TABLE 2 Statistical comparison on *PTT* and *SRT*

Type	<i>PTT</i>			<i>SRT</i>		
	Observed	Simulated	Gap	Observed	Simulated	Gap
Mean (min)	10.42	10.40	-1.58	11.66	10.41	2.25
Std (min)	6.59	6.55	6.17	6.95	6.56	1.97
Student's t test	p-value: 0.966			p-value: 0.0135		
Wilcoxon's rank-sum test	p-value: 7.418e-12 (≈ 0)			p-value: 0.0138		

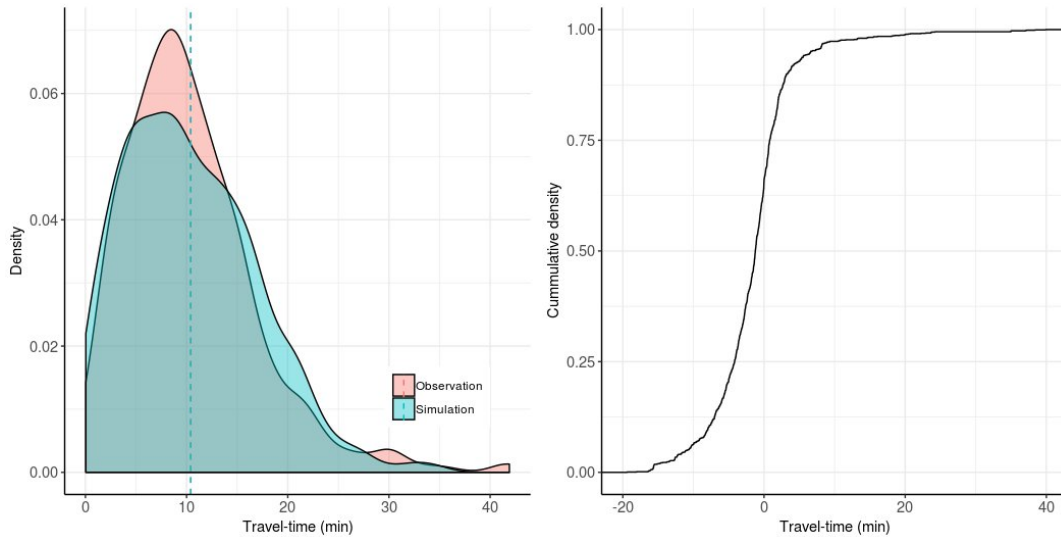


FIGURE 6 Distribution of *PTT* (left: Overall density, right: Cummulative density of gap).

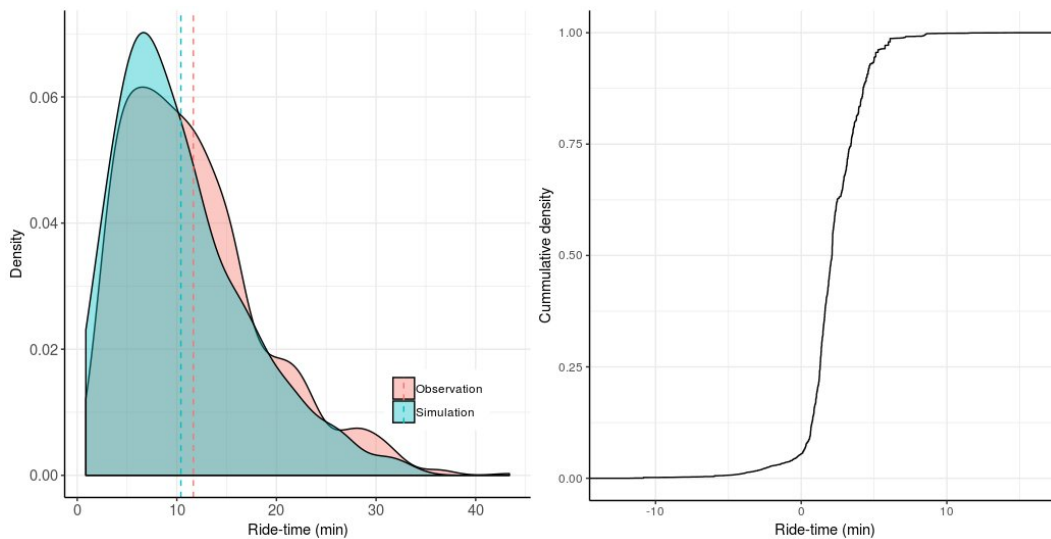


FIGURE 7 Distribution of *SRT* (left: Overall density, right: Cummulative density of gap).

5. CONCLUSIONS

This paper presents an application of the W-DSPSA algorithm to solve the multimodal demand calibration problem whose domain is defined over discrete sets. The W-DSPSA algorithm is a discrete version of the continuous W-SPSA algorithm with the gradient defined using a symmetric discrete perturbation and multimodal weight matrices. The application of the

algorithm is demonstrated using a case study involving the calibration of the SimMobility ST simulator on a sub-network of Singapore. The results indicate that the algorithm at convergence achieves a significant improvement of 60-80% (with respect to the initial seed OD parameters) in the normalized root mean squared error between simulated and observed vehicle loop counts (*VLC*) and bus passenger counts (*BPC*). Further, a validation indicates that the calibrated model suitably replicates point-to-point and stop-to-stop travel time distributions.

Future research includes extending the current framework to the hybrid (discrete-continuous) problem which includes discrete and continuous valued parameters. Further, the current demand calibration is limited to an aggregate OD estimation which may not be suitable for calibrating trip-chains and tours in the context of activity-based demand models. This promises to be a challenging area for future research.

ACKNOWLEDGEMENTS

This research was supported in part by the National Research Foundation (NRF) Singapore through the Singapore MIT Alliance for Research and Technology's Future Urban Mobility IRG research program, the Ministry of National Development (MND), in collaboration with the Ministry of Transport (MOT), Land Transport Authority (LTA) and Urban Redevelopment Authority (URA), Singapore through the Land and Liveability National Innovation Challenge (L2NIC) grant.

AUTHOR CONTRIBUTION STATEMENT

S. Oh, R. Seshadri, C. L. Azevedo, and M. E. Ben-Akiva contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

REFERENCES

1. NGSIM, <https://ops.fhwa.dot.gov/trafficanalysisistools/ngsim.htm>
2. Pavone, M. Autonomous mobility-on-demand systems for future urban mobility. In *Autonomous Fahren*, Springer Vieweg, Berlin, Heidelberg. 2015, pp. 399-416.
3. Ben-Akiva, M. SMART-Future Urban Mobility. *JOURNEYS Sharing Urban Transport Solutions*, 2010. 5: 30-37.
4. Antoniou, C., Balakrishna, R., Koutsopoulos, H. N., and Ben-Akiva, M. Calibration methods for simulation-based dynamic traffic assignment systems. *International Journal of Modelling and Simulation*, 2011. 31(3): 227-233.
5. Cascetta, E., Inaudi, D., and Marquis, G. Dynamic estimators of origin-destination matrices using traffic counts. *Transportation science*, 1993. 27(4): 363-373.
6. Hazelton, L. Statistical inference for time varying origin-destination matrices. *Transportation Research Part B: Methodological*, 2008. 42(6): 542-552.
7. Frederix, R., Viti, F., Corthout, R., and Tampère, C. New gradient approximation method for dynamic origin-destination matrix estimation on congested networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2263: 19-25.
8. Djukic, T., Lint, J. W. C., and Hoogendoorn, S. P. Application of Principal Component Analysis to Predict Dynamic Origin-Destination Matrices. *Transportation Research Record: Journal of the Transportation Research Board*, 2012. 2283(1): 81-89.
9. Tjalling J. Ypma, Historical development of the Newton-Raphson method, *SIAM Review*, 1995. 37(4): 531-551.

10. Arfken, G. The Method of Steepest Descents. *Mathematical Methods for Physicists*, Orlando, FL: Academic Press, 1985, pp. 428-436.
11. Kiefer, J., and Wolfowitz, J. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 1952. 23(3): 462-466.
12. Spall, J. C. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 1998. 34(3): 817-823.
13. Balakrishna, R., Antoniou, C., Ben-Akiva, M., Koutsopoulos, H., and Wen, Y. Calibration of microscopic traffic simulation models: Methods and application. *Transportation Research Record: Journal of the Transportation Research Board*, 2007. 1999: 198-207.
14. Vaze, V., Antoniou, C., Wen, Y., and Ben-Akiva, M. Calibration of dynamic traffic assignment models with point-to-point traffic surveillance. *Transportation Research Record: Journal of the Transportation Research Board*, 2009. 2090: 1-9.
15. Lee, J. B., and Ozbay, K. New calibration methodology for microscopic traffic simulation using enhanced simultaneous perturbation stochastic approximation approach. *Transportation Research Record: Journal of the Transportation Research Board*, 2009. 2124(1): 233-240.
16. Toledo, T., and Kolehkina, T. Estimation of Dynamic Origin-Destination Matrices Using Linear Assignment Matrix Approximations. *IEEE Transactions on Intelligent Transportation Systems*, 2013. 14(2): 618-626.
17. Tympakianaki, A., Koutsopoulos, H. N., and Jenelius, E. c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin-destination matrix estimation. *Transportation Research Part C: Emerging Technologies*, 2015. 55: 231-245.
18. Lu, L., Xu, Y., Antoniou, C., and Ben-Akiva, M. An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models. *Transportation Research Part C: Emerging Technologies*, 2015. 51: 149-166.
19. Antoniou, C., Azevedo, C. L., Lu, L., Pereira, F., and Ben-Akiva, M. W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models. *Transportation Research Procedia*, 2015. 7: 233-253.
20. Shi, L. A new algorithm for stochastic discrete resource allocation optimization. *Discrete Event Dynamic Systems*, 2000. 10(3): 271-294.
21. Kleywegt, A. J., Shapiro, A., and Homem-de-Mello, T. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 2002. 12(2): 479-502.
22. Gerencsér, L., Hill, S. D., and Vágó, Z. Optimization over discrete sets via SPSA. In Proceedings of the 31st conference on Winter simulation: Simulation---a bridge to the future, ACM, 1999. 1: 466-470.
23. Hill, S. D., Gerencsér, L., and Vágó, Z. Discrete stochastic approximation via simultaneous difference approximations. Proceedings of American Control Conference, 2005. 307-308.
24. Wang, Q., and Spall, J. C. Discrete simultaneous perturbation stochastic approximation on loss function with noisy measurements. Proceedings of American Control Conference, 2011, 4520-4525.
25. Adnan, M., et al. SimMobility: A multi-scale integrated agent-based simulation platform. Presented at 95th Annual Meeting of the Transportation Research Board,

- Washington, D.C., 2016.
26. Azevedo, C. L., et al. Simmobility Short-term: An integrated microscopic mobility simulator. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2622: 13-23.
 27. Lu, Y., et al. Simmobility Mid-term simulator: A state-of-the-art integrated agent-based demand and supply model. Presented at 94th Annual Meeting of the Transportation Research Board, Washington, D.C., 2015.
 28. Le, D. T., Cernicchiaro, G., Zegras, C., and Ferreira J. Constructing a synthetic population of establishments for the SimMobility microsimulation platform. *Transportation Research Procedia*, 2016. 19: 81-93.
 29. Waddell, P. UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American planning association*, 2002. 68(3): 297-314.
 30. Ben-Akiva, M., et al. A dynamic traffic assignment model for highly congested urban networks. Presented at 90th Annual Meeting of Transportation Research Board, Washington, D.C., 2011.
 31. Ben-Akiva, M., et al. Traffic simulation with MITSIMLab. In *Fundamentals of Traffic Simulation*. Springer, New York, NY. 2010, pp. 233-268.
 32. Azevedo, C. L., et al. Microsimulation of Demand and Supply of Autonomous Mobility-on-Demand. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2564: 21-30.
 33. Basu R., et al. Automated Mobility-on-Demand vs. Mass Transit: A Multi-Modal Activity-Driven Agent-Based Simulation Approach. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 18-06186.
 34. Alho, A., et al. A multi-scale agent-based modelling framework for urban freight distribution. *Transportation Research Procedia*, 2017. 27: 188-196.
 35. Koh, K., et al. Simulation of demand and supply of urban rail in a multimodal environment. IEEE 20th International Conference on Intelligent Transportation Systems, 2017. 647-654.
 36. Adnan, M., et al. Evaluating Disruption Management Strategies in Rail Transit using SimMobility Mid-term Simulator: A study of Singapore MRT North-East line. Presented at 96th Annual Meeting of the Transportation Research Board, Washington, D.C., 2017.
 37. Lovric, M., et al. Evaluating off-peak pricing strategies in public transportation using an activity-based approach. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2544: 10-19.