Bayesian demand calibration for dynamic traffic simulations

Gunnar Flötteröd [*]   Michel Bierlaire [*]   Kai Nagel [†]

January 5, 2010

---

[*]Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, {gunnar.floetteroed, michel.bierlaire }@epfl.ch

[†]Transport Systems Planning and Transport Telematics Laboratory, Berlin Institute of Technology, Germany, nagel@vsp.tu-berlin.de

Abstract

We present an operational framework for the calibration of demand models for dynamic traffic simulations. Our focus is on disaggregate simulators that represent every traveler individually. We calibrate, at a likewise individual level, arbitrary choice dimensions within a Bayesian framework, where the analyst's prior knowledge is represented by the dynamic traffic simulator itself and the measurements are comprised of time-dependent traffic counts. The approach is equally applicable to an equilibrium-based planning model and to a telematics model of spontaneous and imperfectly informed drivers. It is based on consistent mathematical arguments, yet applicable in a purely simulation-based environment, and, as our experimental results show, capable of handling large scenarios.

1 Introduction

There is a broad consensus about the adequacy of microsimulators to the modeling of urban transportation systems, and a wide scope of such simulation systems has been put forward, e.g., (Ben-Akiva et al., 2001a; Mahmassani, 2001; Raney and Nagel, 2006; Waddell et al., 2007). The arguably most prominent advantage of microsimulators is their superior expressiveness because of their arbitrarily fine-grained model structure. However, increasing the resolution of a model also increases its degrees of freedom, which calls for more interactions to be modeled and more parameters to be identified. That is, the potentially greater expressiveness of a microsimulator is faced with a likewise increased need for modeling, data, and calibration. Typically, the calibration of a (nontrivial) model is cast in a statistical framework and is carried out by some numerical procedure. The mathematical convenience of the model under consideration, e.g., in terms of continuity, differentiability, normality or ergodicity, defines the computational feasibility of this approach. A microsimulator easily reaches a level of detail at which most of these features are lost.

In this article, we present a mathematically consistent and computationally efficient framework for the calibration of microsimulation-based travel demand models in the context of dynamic traffic assignment (DTA). Specifically, we show how to calibrate a microscopic motorist demand simulator from time-dependent traffic counts that are obtained at a limited set of network locations. The problem is solved in a Bayesian setting, where the a priori assumption about every individual's choice distribution is combined with the available measurements' likelihood into an estimated posterior choice distribution. The method is entirely simulation-based in that it only requires a simulation system to represent the behavioral prior

distribution and only generates realizations from the behavioral posterior distribution. The approach is applicable both in stochastic equilibrium conditions and in non-equilibrium conditions. We present experimental results that demonstrate the method's applicability to systems with ten thousands of network links and hundred thousands of travelers.

The calibration of both DTA simulators and disaggregate demand models has received much attention in the literature, which is detailed in the following. However, we are not aware of any work that estimates individual-level travel behavior within a DTA simulation system from aggregate sensor data on a practically relevant scale. All of the subsequently reviewed approaches consider either simplified or partial versions of this problem.

The most frequently adopted method for demand calibration from traffic counts is origin-destination (OD) matrix estimation. An OD matrix models the demand of a given time interval in terms of flows from every origin to every destination of a traffic system. The originally static problem was to estimate such a matrix given a linear assignment mapping of demand on link flows. Various methods such as entropy maximization and information minimization (van Zuylen and Willumsen, 1980), Bayesian estimation (Maher, 1983), generalized least squares (Bell, 1991; Bierlaire and Toint; Cascetta, 1984), and maximum likelihood estimation (Spiess, 1987) were proposed to solve this task. Nonlinear assignment mappings were incorporated by a bilevel-approach that iterates between the nonlinear assignment and a linearized estimation problem (Maher et al., 2001; Yang, 1995; Yang et al., 1992) until a fixed point of this mutual mapping is reached (Bierlaire and Crittin, 2006; Cascetta and Posterino, 2001). The combined estimation of OD matrices in subsequent time slices was demonstrated in (Cascetta et al., 1993), and many originally static methods were applied to dynamic problems in this vein, e.g., (Ashok, 1996; Bierlaire, 2002; Sherali and Park, 2001; Zhou, 2004).

Since a time-dependent OD matrix maps (origin, destination, departure time) tuples on demand levels, it represents destination and departure time choice on an aggregate level. Route choice, however, constitutes no additional degree of freedom but is a function of demand that is defined through the DTA system's modeling assumptions. Path flow estimators (PFEs) overcome this confinement.

The seminal PFE is a macroscopic one-step network observer that estimates static path flows from link volume measurements based on a multinomial logit stochastic user equilibrium (SUE) modeling assumption in a congested network (Bell, 1995; Bell et al., 1997). The estimation problem is transformed into one of smooth optimization, which is iteratively solved. The model was enhanced by multiple user classes and a simple analytical queuing model to represent traffic flow dynamics (Bell et al., 1996) and was successfully implemented in various research

and development projects (Bell and Grosso, 1999). The PFE's non-stochastic user equilibrium counterpart had been proposed in (Sherali et al., 1994, 2003) and was further advanced in (Nie and Lee, 2002; Nie et al., 2005). PFEs also serve as OD matrix estimators since an OD flow is the sum of the path flows between its OD pair.

All PFEs and OD matrix estimators are confined to their underlying modeling assumptions. PFEs only consider static demand per time slice and rely on particular assumptions about route choice behavior. Time-dependent OD matrix estimators represent demand correlations across subsequent time slices in a simplified and aggregate way, e.g., by auto-regressive processes or polynomial trends (Ashok, 1996; Zhou, 2004). These approaches disregard many aspects of real travel behavior, which results from highly individual activity patterns and likewise complex constraints (Bowman and Ben-Akiva, 1998; Kitamura, 1988, 1996; Vovsha et al., 2004). That is, even if a PFE or an OD matrix estimator is applied to a fully microscopic DTA simulator, the aggregate estimator is unable to account for those facets that amount to the microscopic modeling approach.

Random utility models (RUMs) capture travel behavior at the individual level, and sophisticated calibration procedures for this class of models are available (Ben-Akiva and Lerman, 1985; Bierlaire, 2003; Train, 2003). However, in order to maintain tractability, their calibration procedures require a mathematically well-behaved link between observations and model parameters. Here, this link is given through a DTA microsimulator. We are not aware of any work that calibrates a RUM in such conditions.

A calibration of the UrbanSim microsimulator in a Bayesian setting is reported in (Sevcikova et al., 2007), where a sampling importance resampling (SIR) type algorithm is applied to the estimation of almost 300 model parameters. However, concerns regarding the computation times for larger problems are mentioned.

The remainder of this article is organized as follows. The disaggregate demand calibration is incrementally developed in Sections 2 through 4: First, Section 2 derives a macroscopic and static version of the calibration. Second, Section 3 carries this result over to a fully disaggregate DTA microsimulation. Finally, Section 4 discusses the operational aspects of the calibration and summarizes the conceptual developments with a specification of the interactions between the calibration and a DTA microsimulator. A large real-world case study is presented in Section 5. Section 6 concludes the article and gives an overview of ongoing and future research topics.

2 Aggregate path flow estimation

This section develops a new solution to the familiar problem of estimating aggregate path flows between a set of OD pairs from traffic counts. For simplicity, the time dimension is omitted and homogeneous travelers are assumed. The next section generalizes this result for a broad class of DTA microsimulations, which naturally account for both dynamics and heterogeneity in the population. However, since these properties can also be incorporated in the macroscopic framework considered here, the result of this section is a novel PFE in its own right.

2.1 Specification

A network of nodes and links is considered, where some or all nodes constitute demand origins and/or destinations. There are N OD pairs. The largest possible number of trips between OD pair n is denoted by d_n, the symbol C_n represents the set of available paths that connect OD pair n, and d_{ni} is the number of trips on path $i \in C_n$, where $d_n = \sum_{i \in C_n} d_{ni}$. Variations in the total OD flows can be enabled by adding one fictitious path to every OD pair that bypasses the physical network (Sheffi, 1985).

The share of travelers in OD relation n that choose path i is denoted by $P_n(i|\mathbf{x}(\mathbf{d}))$ where $\mathbf{d} = (d_{ni})$ is the vector of all path flows and $\mathbf{x}$ is the vector of network conditions, which depend on the path choice in the entire population. An SUE in this system is defined as a path flow pattern that solves

$$d_{ni} = P_n(i|\mathbf{x}(\mathbf{d}))d_n \ \forall n = 1 \ldots N, \ i \in C_n, \tag{1}$$

which states that the path flows, when loaded on the network, result in path choice fractions that reproduce these path flows (Daganzo and Sheffi, 1977). Appendix A shows that this model can be reformulated as the problem of finding path flows **d** that maximize the **prior entropy function**

$$W(\mathbf{d}) = \sum_{n=1}^{N} \sum_{i \in C_n} [d_{ni} \ln P_n(i|\mathbf{x}(\mathbf{d})) - d_{ni} \ln d_{ni}]$$
$$\text{s.t.} \sum_{i \in C_n} d_{ni} = d_n \ \forall n = 1 \ldots N, \tag{2}$$

which represents for a large population the logarithm of the probability that, for given **prior route choice fractions** $P_n(i|\mathbf{x}(\mathbf{d}))$ at the microscopic level, the path flows **d** occur at the macroscopic level.

Given the traffic counts $\mathbf{y}$ that are observed on some or all links of the network, the calibration should adjust the path flows in a way such that these counts are reproduced to a reasonable degree. For this purpose, the path flows $\mathbf{d}$ that maximize the **posterior entropy function**

$$W(\mathbf{d}|\mathbf{y}) = \ln p(\mathbf{y}|\mathbf{x}(\mathbf{d})) + W(\mathbf{d})$$
$$\text{s.t.} \sum_{i \in C_n} d_{ni} = d_n \ \forall n = 1 \ldots N \tag{3}$$

are identified, where the likelihood $p(\mathbf{y}|\mathbf{x}(\mathbf{d}))$ is the probability of observing the measurements $\mathbf{y}$ given the network conditions $\mathbf{x}$ that result from the path flows $\mathbf{d}$. The posterior entropy models, again for a large population, the logarithm of the probability that a certain aggregate path flow pattern $\mathbf{d}$ occurs given both the prior route choice model $P_n(i|\mathbf{x}(\mathbf{d}))$ and the measurements $\mathbf{y}$.

Appendix B shows that a maximization of $W(\mathbf{d}|\mathbf{y})$ yields the following **posterior route choice fractions**:

$$P_n(i|\mathbf{x}(\mathbf{d}), \mathbf{y}) = \frac{\exp(\Lambda_{ni} + \Gamma_{ni})P_n(i|\mathbf{x}(\mathbf{d}))}{\sum_{j \in C_n} \exp(\Lambda_{nj} + \Gamma_{nj})P_n(j|\mathbf{x}(\mathbf{d}))} \tag{4}$$

where

$$\Lambda_{ni} = \frac{\partial \ln p(\mathbf{y}|\mathbf{x}(\mathbf{d}))}{\partial d_{ni}} \tag{5}$$

$$\Gamma_{ni} = \sum_{m=1}^{N} \sum_{j \in C_m} \frac{d_{mj}}{P_m(j|\mathbf{x}(\mathbf{d}))} \frac{\partial P_m(j|\mathbf{x}(\mathbf{d}))}{\partial d_{ni}}. \tag{6}$$

This result follows from the first order necessary optimality conditions. Without further assumptions about the functions $P_n(i|\mathbf{x}(\mathbf{d}))$ and $p(\mathbf{y}|\mathbf{x}(\mathbf{d}))$, it is not guaranteed to be a global maximizer of the posterior entropy function. However, for a concave likelihood function and fixed path choice fractions (which result in a concave prior entropy), the posterior entropy is concave as well and the above solution is the unique maximizer.

The specification (4) – (6) is at the heart of the disaggregate demand calibration procedure presented in the next sections. It requires to scale the choice fractions of every path i of every OD pair n by $\exp(\Lambda_{ni} + \Gamma_{ni})$ and to re-normalize. Λ_{ni} captures the effect of the path flow d_{ni} on the log-likelihood, i.e., on the measurement reproduction. Γ_{ni} essentially describes how a change in d_{ni} affects all path flows $\mathbf{d}$ through the network conditions $\mathbf{x}$.

The presented approach constitutes a generic PFE in that it makes, apart from differentiability, no assumptions about the deployed route choice and network loading model, and it functions with arbitrarily few measurements, the precision of

which can be accounted for through an arbitrary likelihood function. This is an important advantage over all PFEs reviewed in Section 1, which require special route choice and network loading models and do not deal with incomplete and inconsistent measurements in the integrated and statistically consistent manner a generic likelihood function provides. However, the arguably most important advantage of the proposed PFE is its transferability to a broad class DTA microsimulations, which constitutes the main objective of this article. Further applications to formal mathematical models are therefore left as a subject of future research.

The following subsection illustrates the workings of the new PFE in terms of an academic example, which is revisited in a microsimulation setting in Section 4.4.

2.2 Example: two-route network

A simple network that consists of two unidirectional, identical, and parallel links (1 and 2) that connect a single OD pair is considered. For simplicity, the OD index is omitted in this example. The demand amounts to $d = 1000$ travelers in the considered analysis period. Either link constitutes a feasible routing alternative. The travel times on either path result from identical link performance functions

$$t(d_i) = \left(\frac{d_i}{750}\right)^2, \quad i = 1, 2 \tag{7}$$

that depend on the flow d_i (in vehicle units) on the respective path. Keeping with the full notation of the previous subsection, a three-dimensional vector of relevant network conditions is specified:

$$\mathbf{x}(\mathbf{d}) = \begin{pmatrix} x_1(\mathbf{d}) \\ x_2(\mathbf{d}) \\ x_3(\mathbf{d}) \end{pmatrix} = \begin{pmatrix} t(d_1) \\ t(d_2) \\ d_1 \end{pmatrix} \tag{8}$$

where the first two components, the route travel times, are needed for feedback into the route choice model and the third component is used to specify a likelihood function further below.

Route choice is captured by the logit model

$$P(i|\mathbf{x}(\mathbf{d})) = \frac{\exp(-t(d_i))}{\exp(-t(d_1)) + \exp(-t(d_2))}, \quad i = 1, 2. \tag{9}$$

The symmetry of this setting implies prior route flows of 500 vehicle units on either path in SUE conditions. The concrete values in this example are chosen in

Table 1: System responses to different path flows

path flows	$d_1 = 500, d_2 = 500$	$d_1 = 250, d_2 = 750$
t_1, t_2 according to (7)	0.44, 0.44	0.11, 1.0
$P(1), P(2)$ according to (9)	0.5, 0.5	0.71, 0.29

order to obtain clear system responses that facilitate the discussion. For illustration, some numbers are given in Table 1.

A single flow sensor is located on link 1, which counts y_1 vehicle units during the analysis period. Writing $\mathbf{y} = (y_1)$, the likelihood function is specified as

$$p(\mathbf{y}|\mathbf{x}(\mathbf{d})) \propto \exp \frac{(d_1 - y_1)^2}{2\sigma_1^2} \tag{10}$$

where σ_1 (in vehicle units) is the standard deviation of the sensor data.

The posterior entropy of this simple scenario is strictly concave and has a unique maximum. Observing that $d_2 = d - d_1$, the posterior route choice fraction $P(1|\mathbf{x}(\mathbf{d}), \mathbf{y})$ can be expressed as a single nonlinear equation by substitution of (7) – (10) into (4) – (6), which in this setting guarantees global optimality. However, the resulting expression is fairly unwieldy and therefore given only in graphical terms.

Figure 1 shows the estimated flows on path 1 over measurements y_1 and variances σ_1^2 that are varied between 0 and d. The results are consistent with what one would intuitively expect: The smaller σ_1, the more belief is put on the measurement and the better it is reproduced. For large σ_1 values, the estimator becomes independent of the sensor data and falls back to the prior path flows. Between these extremes, there is a smooth transition that reflects the PFE's ability to interpolate between the prior information contained in the model and the measurements.

In the full PFE, the Γ coefficients require to calculate the derivatives of all path choice fractions with respect to all path flows, where the coupling of these quantities is given through the network loading in that the interactions of all path flows generate network conditions that in turn are evaluated in the route choice model. These derivatives are available in simple settings, but they may be hard to obtain for generic demand and supply models. This difficulty is not specific to this PFE but applies more generally to all instances of the OD matrix estimation problem in congested conditions, where the most widespread solution is to assume a "proportional assignment" that essentially assumes fixed route choice fractions (Cascetta and Nguyen, 1988) and to account for their actual dependency on the network conditions in a heuristic, iterative fashion, e.g., (Lundgren and Peterson,
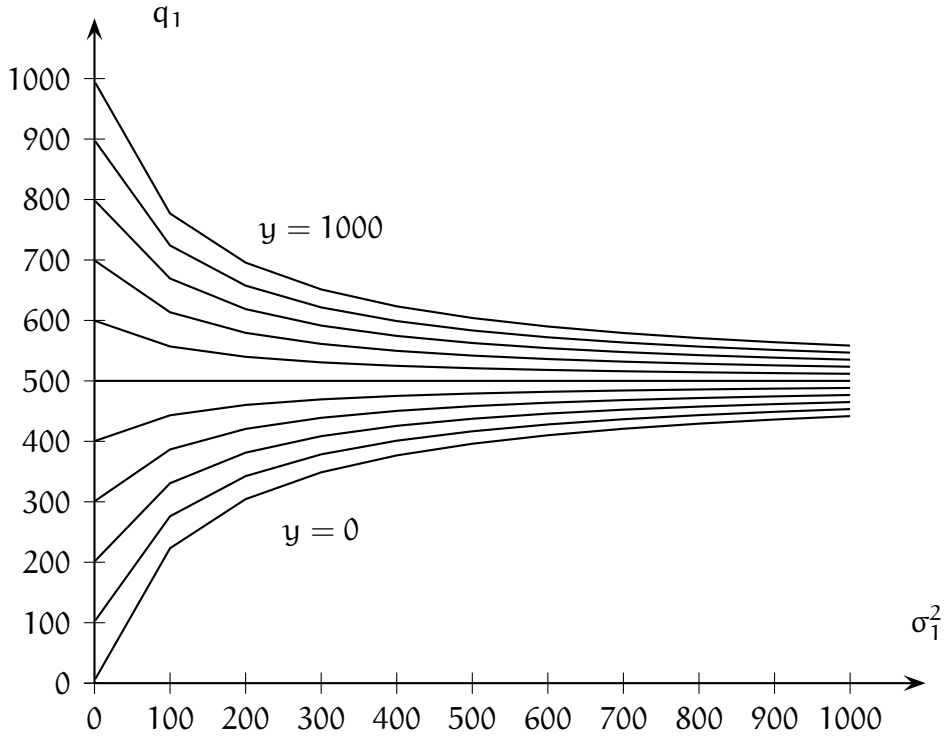
Figure 1: Calibration results for two routes example

2008). This coincides with the statement of zero derivatives of route shares with respect to path flows and hence implies that zero Γ coefficients may be an operationally attractive simplification. Even for zero Γ coefficients, congestion is accounted for in (4) and (5) through the dependency of both the route choice model and the likelihood function on the network conditions.

Figure 2 demonstrates the effect of this simplification on the estimation results. It plots the difference between the exactly estimated route flows and their approximations for zero Γ coefficients. The bias attains a maximum value of ca. $\pm 7\%$ of the total demand around $\sigma_1^2 = 100$ for $y_1 = 0$ and $y_1 = d$. For very small and very large variances, the bias ceases: In the first case, the Λ coefficients absolutely dominate (4), whereas in the second case the calibration falls back to the prior model. Since the bias is of moderate magnitude, it appears justified to choose zero Γ coefficients in favor of the operational advantages this brings along. This course of action is chosen in the remaining experiments of this article. However, accounting more precisely for the SUE feedback effects, which here are represented by the Γ coefficients, is an important subject of ongoing and future research (Lundgren and Peterson, 2008). Note that related progress in the field of OD matrix estimation is likely to be transferable to the methodology proposed here.

Summarizing, this section introduces a new PFE that makes much milder assumptions about the underlying model components and the amount and quality of avail-
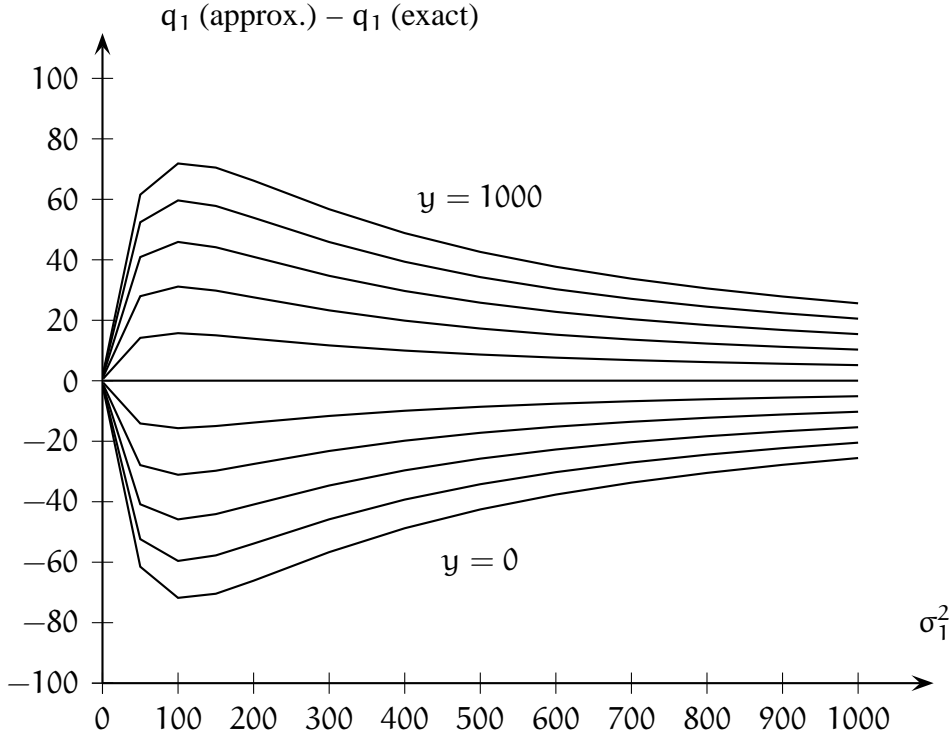
9

Figure 2: Bias of simplified calibration for two routes example

able sensor data than the PFEs presented so far in the literature. Its functioning is demonstrated through an academic example, and some intuition about an operationally advantageous simplification is provided.

3 Disaggregate demand calibration

This section carries the macroscopic PFE over to the calibration of DTA microsimulations. It is organized in two parts. First, the considered type of DTA simulator is described. Second, the considerations that enable a mathematically consistent application of the PFE to this type of simulation are discussed.

Throughout this article, probability density functions are denoted by a lowercase p and discrete probability functions by an uppercase P. Instead of noting the probability that random variable X takes value x by some expression of the form $P(X = x)$, $P(x)$ is briefly written and ambiguities are avoided by self-explanatory variables.

3.1 Considered DTA simulator

This specification builds on the seminal model of Cascetta (1989), which it simplifies in some regards and extends in others. The notation of the previous section

10

is in large parts re-defined here in a microsimulation context. The most important changes in the new setting are that (i) it is fully disaggregate in that every traveler is modeled as an individual entity and (ii) it is fully dynamic both on the demand side and the supply side.

Agents and plans

We assume a microsimulation-based approach where every traveler is modeled as an individual **agent** $n = 1 \dots N$. At every point in simulated time, every agent n disposes of a **plan** i_n that describes the intended travel behavior of that agent. A typical plan comprises a sequence of trips that connect intermediate stops during which activities are conducted, including all associated timing information. We subsequently write $\{i\}$ as a shortcut for the whole population's plan set $\{i_1, \dots, i_N\}$.

A plan constitutes a fully dynamic demand specification that captures arbitrary choice dimensions such as route choice, departure time choice, and mode choice. An informal example of a plan would be "Leave home by car for work at 7 am with a planned arrival at 7:30 am, taking the habitual route; work until 5 pm; then take the highway to get to the local mall for one hour of shopping; finally return home for the rest of the day, again using the habitual route."

Supply simulator

The **supply simulator** executes the plans of all agents simultaneously on the network. It models the physical interactions of the agents, including congestion. The result of such a **dynamic network loading** are the **dynamic network conditions** $\mathbf{x}$, which comprise all time-dependent, aggregate network characteristics (such as flows, densities, velocities) that are relevant to the decision making of the agents. (No time index is used here for notational simplicity; one may think of $\mathbf{x}$ as a large vector in which time-dependent $\mathbf{x}(k)$ vectors are stacked, where k is the simulation time step index.)

Formally, the supply simulator draws from a distribution $p(\mathbf{x}|\{i\})$ of the time dependent network conditions $\mathbf{x}$ that result from the dynamic network loading of a particular plan set $\{i\}$ in the population. In its most widespread form, this distribution is implicitly defined through a stochastic supply microsimulator. However, a deterministic, macroscopic supply simulator where $p(\mathbf{x}|\{i\})$ collapses into a singleton is just as feasible.

Demand simulator

The **demand simulator** models the decision making of travelers. It maps, for every agent $n = 1 \ldots N$ individually, the expected network conditions $\bar{\mathbf{x}}$ on a plan i_n the agent chooses in these conditions. $P_n(i_n|\bar{\mathbf{x}})$ is the probability that plan i_n is chosen by agent n given the expected network conditions $\bar{\mathbf{x}}$, and C_n denotes agent n's **choice set** of available plan alternatives.

It is assumed that the agents' plan choice distributions are independent once the expected network conditions are given. That is,

$$P(\{i\}|\bar{\mathbf{x}}) = \prod_{n=1}^{N} P_n(i_n|\bar{\mathbf{x}}), \tag{11}$$

which implies that the agents do not interact directly but only through the aggregate network conditions. This is a reasonable assumption for large-scale and/or time-critical simulations where traffic flow dynamics are typically represented by aggregate laws of motion ("mesoscopic simulators") instead of vehicle-by-vehicle interactions ("car-following models") (Astarita et al., 2001; Ben-Akiva et al., 2001a; De Palma and Marchal, 2002; Mahmassani, 2001; Nökel and Schmidt, 2002).

The choice distributions $P_n(i_n|\bar{\mathbf{x}})$ and the choice sets C_n are arbitrary and entirely transparent to the proposed calibration approach. The demand simulator is only required to generate realizations of these distributions.

Iterative simulation logic

So far, the DTA simulator is defined in terms of a supply simulator and a demand simulator. A solution to the DTA problem represents a situation in which demand and supply are consistent with each other. It typically is impossible to simulate this situation directly, but it is possible to alternately execute the supply simulator and the demand simulator. After a burn-in period, these draws can be tested for convergence towards a stationary distribution, and their continuation in stationary conditions allows to extract the relevant characteristics of mutually consistent demand and supply (Balijepalli et al., 2007; Cascetta and Cantarella, 1991; Nagel et al., 1998; Watling and Hazelton, 2003).

To clarify the causal structure of this logic, an **iteration cycle counter** c is introduced. In a given iteration c, the demand simulator first draws plans from $P(\{i\}^c|\bar{\mathbf{x}}^c)$ conditional on expected network conditions $\bar{\mathbf{x}}^c$ that are inferred from the simulated network conditions of previous iterations, and then the supply simulator draws network conditions that result from an execution of these plans from $p(\mathbf{x}^c|\{i\}^c)$.

Algorithm 1 Iterative dynamic traffic assignment
---
1. Initialize cycle counter $c = 0$.

2. Choose initial network conditions $\mathbf{x}^0, \mathbf{x}^{-1}, \ldots$ (e.g., free-flow conditions).

3. Repeat for as many iterations as necessary to extract relevant characteristics in stationary conditions:

   (a) Increase c by one.

   (b) Calculate expected network conditions $\bar{\mathbf{x}}^c$ from $\mathbf{x}^{c-1}, \mathbf{x}^{c-2}, \ldots$.

   (c) Replanning. For $n = 1 \ldots N$, draw plan i_n^c from $P_n(i_n^c | \bar{\mathbf{x}}^c)$.

   (d) Network loading. Draw network conditions $\mathbf{x}^c$ from $p(\mathbf{x}^c | \{i\}^c)$.

The loop is closed by a model component that infers the expected network conditions $\bar{\mathbf{x}}^c$ from the previously simulated network conditions $\mathbf{x}^{c-1}, \mathbf{x}^{c-2}, \ldots$. Possible realizations of this filter are a moving average over a number of previous iterations, e.g., (Liu, 2005), an autoregressive process, e.g., (Ben-Akiva et al., 2001b; Raney and Nagel, 2006), or the method of successive averages (MSA), e.g., (Liu et al., 2007). For the calibration, it only is required that the expected network conditions attain a low variability as c becomes large. This requirement is made more precise further below.

Algorithm 1 summarizes the workings of this approach. It constitutes a stochastic process that eventually stabilizes at a stationary distribution of plan choices and resulting network conditions that constitute the simulation-based solution of the DTA problem. It is called the **prior solution** of the model because it incorporates no sensor data. (The existence of a unique stationary distribution depends on the involved model components. It can, for example, be guaranteed if the simulation process is designed as an ergodic Markov chain (Ross, 2006).)

Denoting by π a continuous and by Π a discrete stationary probability distribution, the prior solution can be formally given in terms of the following system of equations:

$$\Pi_n(i_n) = P_n(i_n | \bar{\mathbf{x}}), \, n = 1 \ldots N \tag{12}$$

$$\Pi(\{i\}) = \prod_{n=1}^{N} \Pi_n(i_n) \tag{13}$$

$$\pi(\mathbf{x}) = p(\mathbf{x} | \{i\} \sim \Pi(\{i\})) \tag{14}$$

$$\bar{\mathbf{x}} \approx E\{\mathbf{x} | \mathbf{x} \sim \pi(\mathbf{x})\}. \tag{15}$$

Equation (12) specifies the individual-level prior choice distribution of every agent n. Equation (13) states that the population prior choice distribution $\Pi(\{i\})$ results from the independent choices of all agents (where the mutual interactions are fully captured through the expected network conditions $\bar{\mathbf{x}}$). The prior distribution of the network conditions is defined in (14), and the expected prior network conditions are given in (15).

The requirement (15) that the agents replan based on (an approximation of) the expected network conditions is motivated as follows. The macroscopic PFE solves the calibration problem through an adjustment of all choice distributions *in equilibrated conditions*. The counterpart of these distributions in a microsimulation are the stationary choice distributions, which are implicitly defined through the iterative dynamics of the stochastic simulation process. If, however, the expected network conditions $\bar{\mathbf{x}}$ eventually stabilize at constant values, then the transition distributions $P_n(i|\bar{\mathbf{x}})$ and the stationary choice distributions $\Pi_n(i)$ coincide and the calibration problem can be tackled by a modification of the operationally more accessible transition distributions only. (The subscript n of a plan i_n is subsequently omitted when the agent the plan refers to is not of relevance.)

The transition distributions and the stationary choice distributions coincide well even if some variability in the expected network conditions $\bar{\mathbf{x}}$ is left in that they are distributed according to some distribution $\pi(\bar{\mathbf{x}})$ in stationary conditions. To see this, the stationary plan choice distribution (12) is rewritten as

$$\Pi_n(i) = \int P_n(i|\bar{\mathbf{x}})\pi(\bar{\mathbf{x}})d\bar{\mathbf{x}}. \tag{16}$$

If the expectation of $\pi(\bar{\mathbf{x}})$ equals the expectation $E\{\mathbf{x}|\mathbf{x} \sim \pi(\mathbf{x})\}$ of the simulated network conditions and if the distribution $\pi(\bar{\mathbf{x}})$ is tight enough to allow for a linearization of $P_n(i|\bar{\mathbf{x}})$ around $\bar{\mathbf{x}}^0 = E\{\mathbf{x}|\mathbf{x} \sim \pi(\mathbf{x})\}$ then

$$\Pi_n(i) \approx \int \left[P_n(i|\bar{\mathbf{x}}^0) + \frac{\partial P_n(i|\bar{\mathbf{x}}^0)}{\partial \bar{\mathbf{x}}^0}(\bar{\mathbf{x}} - \bar{\mathbf{x}}^0) \right] \pi(\bar{\mathbf{x}})d\bar{\mathbf{x}} = P_n(i|\bar{\mathbf{x}}^0), \tag{17}$$

which implies that the stationary plan choice distribution and the transition distribution coincide well even if (15) is implemented through a filter that maintains some variability in the expected network conditions. Also, the expected network conditions may differ for individual agents within the aforementioned limits. However, for notational convenience the model will subsequently be specified in terms of an approximation of the expected network conditions only, as it is expressed in (15) by the "$\approx$" symbol.

The iterative assignment logic is equally applicable to simulate an SUE-based planning model and a telematics model where drivers are spontaneous and imperfectly informed. From a simulation point of view, the only difference between

these two models is that an SUE demand simulator typically utilizes all information from the most recent network loadings, whereas a telematics demand simulator generates every elementary decision of a plan only based on such information that could have actually been gathered up to the according point in simulated time. The filtering of the expected network conditions has different meanings in either approach: In an equilibrium model, it can be seen as a learning mechanism through which travelers remove random fluctuations from their observations. For a non-equilibrium model, the same mechanism can be employed to stabilize the iterative solution procedure, but no behavioral interpretation is available in this case (Bottom, 2000; Bottom et al., 1999). To keep the terminology simple, the remaining presentation is given only in terms of an SUE planning model.

3.2   Disaggregate application of the calibration

The macroscopic PFE developed in Section 2 is now carried over to the previously described DTA microsimulator. Essentially, the OD pairs are replaced by agents and the routes are replaced by plans. That is, $n = 1 \ldots N$ now represents the agent population instead of the OD pairs, C_n represents the choice set of agent n instead of the route set connecting OD pair n, and $i \in C_n$ indicates a plan available to agent n instead of a route that connects OD pair n. The transition from a static specification that only considers paths to a dynamic specification that accounts for full plans is feasible because a time-dependent network can be equivalently modeled as a time-expanded static network and a full-day plan constitutes a simple path in the expanded network (Bierlaire, 2002; Flötteröd, 2008; van der Zijpp and Lindveld, 2001).

The basic assumption of this approach is that the macroscopic SUE model of Section 2 captures the average conditions in the microsimulation such that the macroscopic PFE can be deployed to adjust the average conditions in the microsimulation as well. This requires to clarify the notions of "average network conditions" and "average agent behavior" in the considered class of DTA microsimulators.

Average network conditions. The macroscopic PFE assumes that the network conditions result from a deterministic network loading of the continuous-valued demand. The microscopic model is based on an expectation of stochastic network conditions. Since the network loading is in general a nonlinear operation, the expected network conditions differ from the result of a deterministic network loading of the expected demand levels. This deviation between aggregate SUE assignments and stochastic microsimulations has been identified by Cascetta (1989), who concludes that "in the limiting case of a number of remembered costs tending to infinity with uniform weights, users tend to base their choices on average

15

Table 2: Microscopic redefinition of macroscopic PFE entities

symbol	macroscopic	microscopic
$n = 1 \ldots N$	OD pairs	agents
C_n	routes connecting OD pair n	plans available to agent n
$i \in C_n$	a route connecting OD pair n	a plan of agent n
d_n	number of trips in OD pair n	number of times agent n chooses a plan per iteration (= one)
d_{ni}	number of trips on route $i \in C_n$	stationary probability that agent n chooses plan i

costs, which are still different from costs computed for average flows in the case of nonlinear cost functions. Also in this case [the iterated microsimulation] and SUE expected flows are only approximately equal." However, he also shows that "in general, however, they can be considered coincident within the limits of a first-order approximation".

Average agent behavior. Every agent n chooses one plan in every iteration of the microsimulation. This implies that d_n, which previously was the number of trips in OD relation n, now is one. A natural re-interpretation of d_{ni}, which previously was the number of trips in OD relation n along path i, is possible in terms of a **continuous limit** that results when agent n is (only hypothetically) replaced by $Z \to \infty$ identical agents of size $1/Z$ that all draw independently from the original agent's plan choice distribution. In the continuous limit, d_{ni} becomes agent n's probability $P_n(i|\cdot)$ of choosing plan i. This observation is relevant because the macroscopic PFE maximizes entropy, which assumes a large population of decision makers. The continuous limit behavior can be evaluated by the considered class of DTA microsimulations in stationary conditions, where every agent n replans based on stable expected network conditions $\bar{\mathbf{x}}$ such that repeated instantaneous choices of the same agent follow the same distribution as a sequence of choices over several iterations. That is, the entropy maximization approach of the macroscopic PFE can still be applied to a microsimulation in stationary conditions with $d_n = 1$ and d_{ni} being the according stationary choice probability of plan i.

Table 2 gives a summary of these re-definitions. Based on these considerations, the macroscopic PFE (4) – (6) can be combined with the solution (12) – (15) of

the simulation-based DTA model into the following specification:

$$\Pi_n(i|\mathbf{y}) \;=\; \frac{\exp(\Lambda_{ni} + \Gamma_{ni})P_n(i|\bar{\mathbf{x}}_{|\mathbf{y}})}{\sum_{j \in C_n} \exp(\Lambda_{nj} + \Gamma_{nj})P_n(j|\bar{\mathbf{x}}_{|\mathbf{y}})}, \; n = 1 \ldots N \qquad (18)$$

$$\Pi(\{i\}|\mathbf{y}) \;=\; \prod_{n=1}^{N} \Pi_n(i_n|\mathbf{y}) \qquad (19)$$

$$\pi(\mathbf{x}|\mathbf{y}) \;=\; p(\mathbf{x}|\{i\} \sim \Pi(\{i\}|\mathbf{y})) \qquad (20)$$

$$\bar{\mathbf{x}}_{|\mathbf{y}} \;\approx\; E\{\mathbf{x}|\mathbf{x} \sim \pi(\mathbf{x}|\mathbf{y})\}, \qquad (21)$$

where (18) and (19) now specify the stationary posterior plan choice distributions in the population symmetrically to (4), and the (expected) posterior network conditions are defined in (20) and (21). Λ_{ni} and Γ_{ni} are defined in (5) and (6), only that they are now evaluated in expected posterior network conditions $\bar{\mathbf{x}}_{|\mathbf{y}}$ and with the path flows d_{ni} being replaced by the stationary posterior choice distributions $\Pi_n(i|\mathbf{y})$.

Recall that (4) – (6) only specify a stationary point of the posterior entropy function but not necessarily a global maximum. If there are several stationary points then additional measures are necessary to ensure a proper maximization, e.g., by running the above model several times and comparing the results. However, our present experience with this specification is that it unambiguously converges towards a single, plausible solution.

The model (18) – (21) can be solved by the same iterative simulation approach that is used to solve (12) – (15), the only difference being that the plan choice distribution of every replanning agent is now scaled by the exponential of the according Λ and Γ coefficients. This is a computationally very efficient specification because it only affects the agent behavior at the individual level, which turns the joint demand calibration problem for N agents into N individual-level calibration problems, where all interactions are captured through the iterations of the simulation.

Algorithm 2 outlines, as for now only conceptually, how the calibration is applied to a generic DTA microsimulator. Clearly, the applicability of this calibration logic is very broad.

In order to make the calibration operational, two more questions need to be answered: how to calculate the Λ and Γ coefficients in Step 2a and how to implement the scaling of the choice probabilities in Step 2b for a generic microsimulation that can only be expected to generate *realizations* of the choice distributions and network conditions. This is discussed in the next section.

17

Algorithm 2 Calibration of a generic DTA microsimulator
---

1. Initialize the calibration and the DTA simulator.

2. Repeat for as many iterations as necessary to extract relevant characteristics in stationary conditions:

   (a) Calculate all Λ_{ni} and Γ_{ni} coefficients.

   (b) For all agents $n = 1 \ldots N$, draw a new plan from a choice distribution that is scaled by $\exp(\Lambda_{ni} + \Gamma_{ni})$ for all $i \in C_n$.

   (c) Load all agents on the network.

---

4   Making the framework operational

This section details the technical steps that are necessary to apply the demand calibration to a DTA microsimulation. First, Subsection 4.1 clarifies how to calculate the Λ coefficients, given an arbitrary supply simulator. Second, Subsection 4.2 explains different methods to enforce the scaled plan choice distribution (18) in an arbitrary demand simulator. Third, Subsection 4.3 gives a step-by-step specification of how to apply the calibration to a generic DTA microsimulation. Finally, Subsection 4.4 clarifies the developments with a continuation of the two-routes example of Section 2.2.

As from now, the Γ coefficients in (18) are set to zero because of the operational reasons given in Section 2.2. If they are to be accounted for, they can be added to the corresponding Λ coefficients wherever the latter are used in the following to affect the simulated agent behavior.

4.1   Linearization of the log-likelihood

Stationary posterior conditions are assumed in this subsection, which means that all agents draw their plans from posterior choice distributions $\Pi_n(i|\mathbf{y})$. This is justified by the specification of the calibrated system state that relies on a linearization of the log-likelihood in posterior conditions. Since in stationary conditions the choices of all agents depend on stable $\bar{\mathbf{x}}_{|\mathbf{y}}$ values and hence are not affected by the particular realizations of $\mathbf{x}$ in recent network loadings, the iteration counter c is omitted in this subsection.

According to (5), a calculation of the Λ coefficients requires to differentiate the log-likelihood function $\ln p(\mathbf{y}|\mathbf{x}(\mathbf{d}))$ with respect to d_{ni}, which in the microscopic

case carries over to a differentiation with respect to the according stationary choice probability $\Pi_n(i|\mathbf{y})$ in expected posterior network conditions $\bar{\mathbf{x}}_{|\mathbf{y}}$, cf. Section 3.2:

$$\Lambda_{ni} = \frac{\partial \ln p(\mathbf{y}|\mathbf{x}_{|\mathbf{y}})}{\partial \Pi_n(i|\mathbf{y})} = \left\langle \frac{\partial \ln p(\mathbf{y}|\mathbf{x}_{|\mathbf{y}})}{\partial \mathbf{x}_{|\mathbf{y}}}, \frac{\partial \mathbf{x}_{|\mathbf{y}}}{\partial \Pi_n(i|\mathbf{y})} \right\rangle \tag{22}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. The first vector, $\frac{\partial \ln p(\mathbf{y}|\mathbf{x}_{|\mathbf{y}})}{\partial \mathbf{x}_{|\mathbf{y}}}$, will turn out to be relatively easy to compute. The evaluation of the second vector, $\frac{\partial \mathbf{x}_{|\mathbf{y}}}{\partial \Pi_n(i|\mathbf{y})}$, however, requires some additional effort. For this purpose, the notion of a "proportional network loading" is introduced.

A **proportional network loading** describes a situation in which the time-dependent travel times on all links in the network are known and fixed. This implies that there are no interactions between the flows, which move through an exogenously specified network environment. The resulting flow on any link becomes a linear superposition of all path flows across that link. For a microsimulator, this implies that the agents linearly superpose on each link. In order to obtain a mathematically tractable relation between demand and resulting network conditions, the true dynamics of the supply simulator are captured by a linear network loading. Formally, this implies that the simulated traffic count $x_a(k)$ on link a in simulation time step k is written as

$$x_a(k) = \sum_{n=1}^{N} \mathbf{1}(ak \in i_n) \tag{23}$$

where $\mathbf{1}(\cdot)$ is the indicator function and $ak \in i_n$ indicates that plan i_n requires agent n to enter link a in time step k (where, for simplicity, it is assumed that the sensors are located at the upstream end of a link). This is an imperfect model of the actual network loading in that the assumption of constant travel times implies that the inflow of links at the capacity limit increases beyond this limit if the demand is increased. Consequently, (23) is an imperfect representation of the supply simulator in congested conditions.[1] An alternative approximation that captures congestion with greater precision is described in (Flötteröd and Bierlaire, 2009). However, for clarity only the simple case of a proportional network loading is considered in the following. The results carry over almost identically to the congested case.

---

[1] Note that a proportional assignment, which is widely and successfully assumed in the field of time-dependent OD matrix estimation, implies the same assumption of constant travel times. That is, although (23) is consistent only in uncongested conditions, the state of practice suggests its applicability even in the case of congestion.

Assuming (23) to be applicable, the vector $\mathbf{x}_{|\mathbf{y}}$ of expected posterior network conditions contains the elements

$$x_a(k)_{|\mathbf{y}} = \sum_{n=1}^{N} \sum_{i \in C_n} \mathbf{1}(ak \in i)\Pi_n(i|\mathbf{y}), \tag{24}$$

which yields when inserted into (22)

$$\Lambda_{ni} = \sum_{ak \in i} \frac{\partial \ln p(\mathbf{y}|\mathbf{x}_{|\mathbf{y}})}{\partial x_a(k)_{|\mathbf{y}}}. \tag{25}$$

This means that the Λ coefficients can be evaluated by summing up the derivatives of the log-likelihood with respect to the simulated traffic counts along all links that are contained in the considered plan.

In order to show that this is not a difficult task, univariate normal likelihood functions are considered as an example. Denoting the measured counterpart of $x_a(k)$ by $y_a(k)$ and maintaining the symbol $\mathbf{y}$ for the vector of all available measurements, one has

$$\ln p(\mathbf{y}|\mathbf{x}_{|\mathbf{y}}) = \text{const} - \sum_{ak} \frac{(x_a(k)_{|\mathbf{y}} - y_a(k))^2}{2\sigma_a^2(k)} \tag{26}$$

where the sum runs only over sensor-equipped links and $\sigma_a^2(k)$ is the variance of the sensor data on link a in time step k. In this case, an evaluation of (25) yields

$$\Lambda_{ni} = \sum_{ak \in i} \frac{y_a(k) - x_a(k)_{|\mathbf{y}}}{\sigma_a^2(k)} \tag{27}$$

where the expectation can be obtained by averaging the simulated traffic counts over many stationary iterations in the DTA simulator.

4.2 Affecting the agent behavior

The disaggregate demand calibration requires to scale the choice distribution $P_n(i|\cdot)$ of every replanning agent individually by $\exp(\Lambda_{ni})$ and to re-normalize. Given that the Λ coefficients are available from (25), a universally applicable method to realize this scaling is rejection sampling (Ross, 2006). Denote by

$$P_{\text{accept},n}(i) = \exp(\Lambda_{ni})/D_n \tag{28}$$

the acceptance probability for plan i from agent n's choice set C_n where D_n must be such that

$$D_n \geq \max_{i \in C_n} \exp(\Lambda_{ni}) \tag{29}$$

for (28) to be a proper probability. If repeated draws taken from $P_n(i|\cdot)$ are accepted with probability $P_{\text{accept},n}(i)$ and are rejected otherwise, then the first accepted draw constitutes a draw from the desired scaled choice distribution. The correctness of this approach is verified in Appendix C.

While the accept/reject estimator is arguably the most general method to affect agent behavior, it is by no means the only one. For example, if the demand simulator implements a multinomial logit (MNL) model (Ben-Akiva and Lerman, 1985) then a computationally more efficient approach is to affect the agent behavior by modifications of their utility functions. Appendix D shows that an MNL demand simulator immediately generates draws from the calibrated choice distributions if the according Λ_{ni} coefficients are added to the systematic utility of every considered alternative before the MNL model is evaluated. Note that this result carries over to path-size logit (Ben-Akiva and Bierlaire, 2003) and C-logit (Cascetta et al., 1996) models. It also is noteworthy that a heuristic application of this technique is possible even if the demand simulator does not implement an MNL choice distribution. Such an approach is based on a weaker theoretical foundation, but it may still produce practically useful results.

4.3 Algorithm

The definition of Λ_{ni} in (25) requires to calculate the according derivatives in average *posterior* network conditions, which, however, are a priori unknown. This constitutes a fixed-point problem that can be iteratively solved: Starting from the behavioral prior, successively improved linearizations are generated from iteration to iteration until a stable state is reached where the estimator draws from the behavioral posterior based on stable Λ coefficients that in turn are consistent with this very posterior.

For illustrative purposes, the method of successive averages (MSA) is applied to this problem in Algorithm 3, which affects the agents' choice behavior using the general rejection sampling technique as an example. This algorithm calibrates whatever choice dimensions are represented by the demand simulator, is compatible with an arbitrary supply simulator, and is fully consistent with the execution logic of a typical DTA microsimulator.

Algorithm 3 Linearization-based accept/reject estimator

---

1. Initialize cycle counter $c = 0$.

2. Choose initial network conditions $\mathbf{x}^0, \mathbf{x}^{-1}, \ldots$ (e.g., free-flow conditions).

3. Repeat for as many iterations as necessary to extract relevant characteristics in stationary conditions:

    (a) Increase c by one.

    (b) Calculate expected network conditions $\bar{\mathbf{x}}^c_{|\mathbf{y}}$ from $\mathbf{x}^{c-1}, \mathbf{x}^{c-2}, \ldots$.

    (c) Replanning. For $n = 1 \ldots N$, do:

        i. Run the demand simulator and obtain a plan i'.
        ii. Calculate $\Lambda_{ni'}$ according to (25) using $\bar{\mathbf{x}}^c_{|\mathbf{y}}$.
        iii. With probability $1 - P_{\mathrm{accept},n}(i')$ according to (28), goto step 3(c)i.
        iv. Retain the first accepted draw: $i^c_n = i'$.

    (d) Network loading. Draw $\mathbf{x}^c$ from $p(\mathbf{x}^c|\{i\}^c)$.

4.4 Example

This subsection exemplifies the workings of Algorithm 3 in terms of the two-routes example introduced in Section 2.2. The example is now microscopically simulated for a population of 1000 identical agents, each of which perceives travel time according to (7) and chooses a route according to (9). The expected travel times result from a moving average of the simulated travel times over five iterations.

For illustrative purposes, a measured flow of $y_1 = 250\,\mathrm{veh/h}$ with a standard deviation of $\sigma_1 = 10\,\mathrm{veh/h}$ is assumed. The calibration is run for 100 iterations. Note that in this setting the Λ_1 coefficient can be calculated according to (27) and that Λ_2 is zero because there is no sensor on route 2. Figures 3, 4, and 5 show, for a single calibration experiment, the flow q_1 on route 1, the expected travel time $\bar{t}_1$ on that route, and the Λ_1 coefficient, respectively. For comparison, the uncalibrated flows and travel times of a single simulation are added in dashed lines.

The prior flows fluctuate in a stable manner around $500\,\mathrm{veh/h}$, which is consistent with the symmetry of the scenario. After some overshooting, the posterior flows stabilize around $360\,\mathrm{veh/h}$, which constitutes the compromise the calibration identifies between the prior flows and the measured value of $250\,\mathrm{veh/h}$. Note
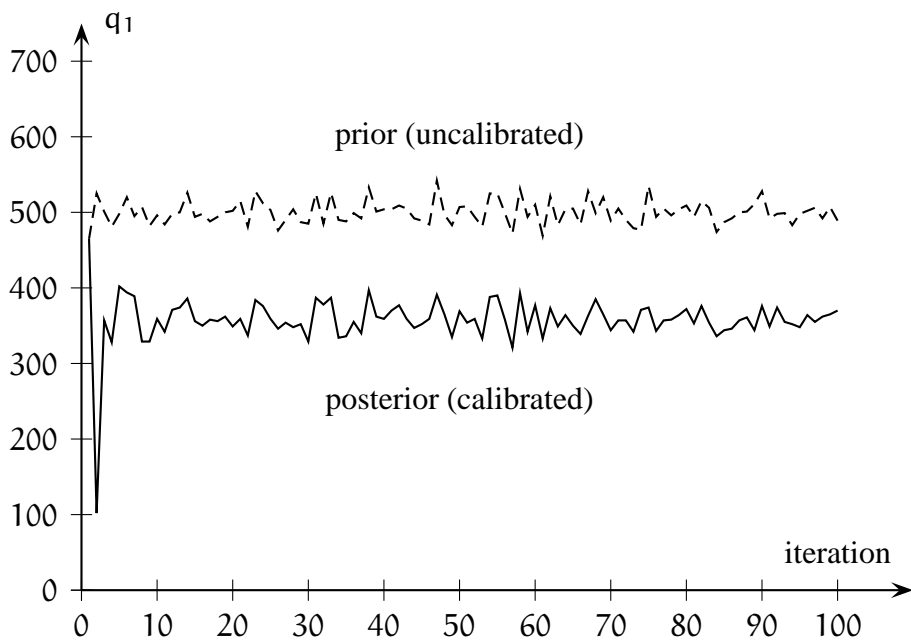
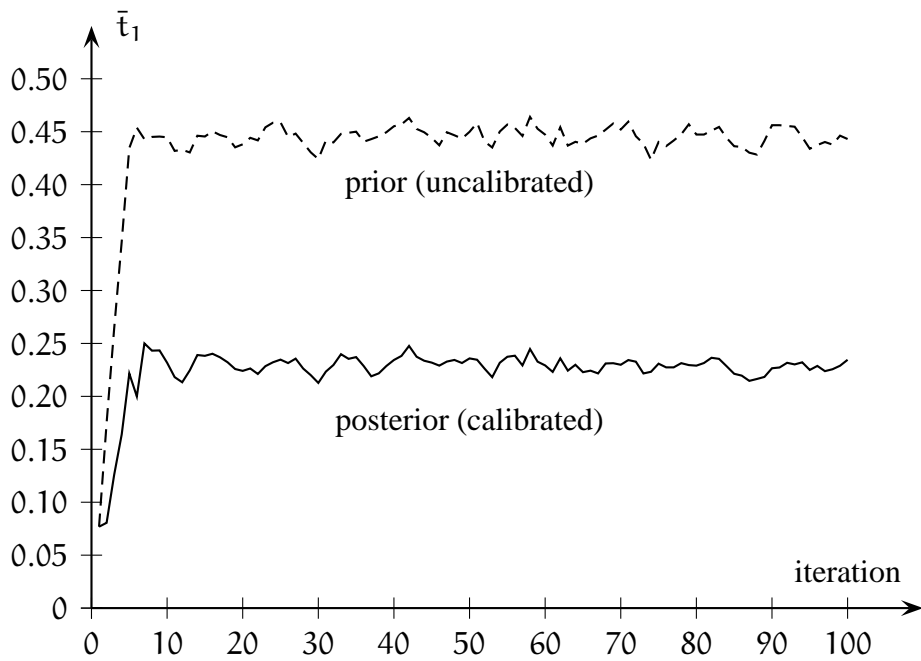Figure 3: Evolution of q_1 for two-routes example

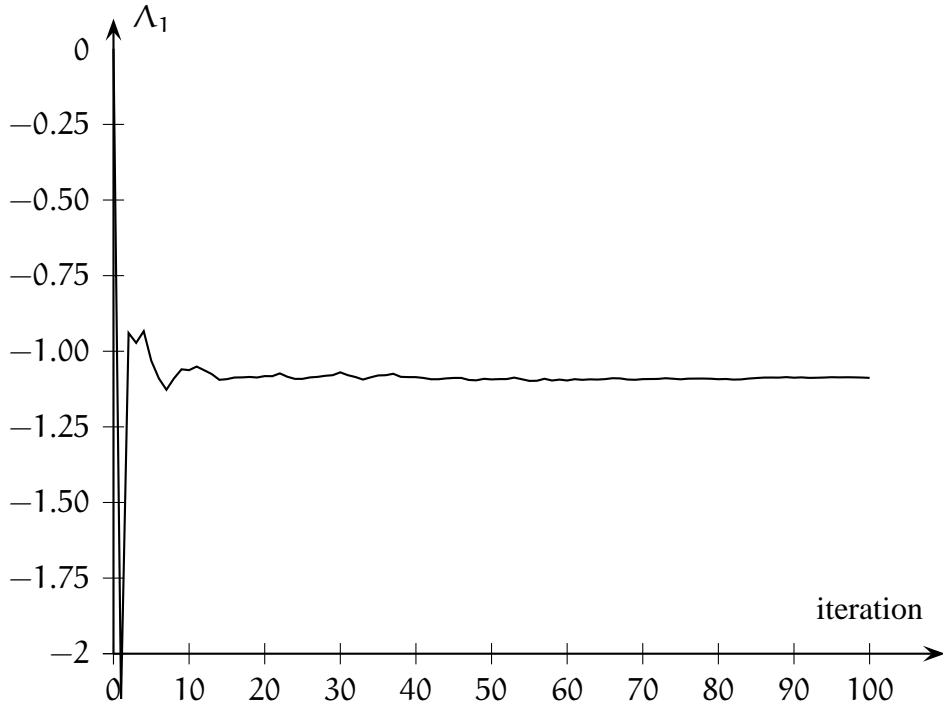Figure 4: Evolution of $\bar{t}_1$ for two-routes example

Figure 5: Evolution of Λ_1 for two routes example

that although the calibration has been derived in terms of average network conditions, the actually calibrated network conditions are still distributed in a way that is consistent with the stochasticity of the demand generator and (in general but not in this example) the supply simulator.

The average travel time on route 1 changes from 0.45 in prior conditions to 0.23 in posterior conditions. This constitutes an important driving force behind the interpolation of prior information and measurements: As the calibration removes more and more vehicles from path 1 in order to fit the measurement, the travel time on this path decreases, which in turn increases its attractiveness. Upon convergence, the calibration has compromised in a plausible Bayesian manner between these two effects.

Finally, the evolution of the Λ_1 coefficient shows how the calibration takes effect. After a few iterations of transient oscillations, the coefficient stabilizes around -1.1. This value is consistent with the theory: Inserting y_1, σ_1 and the average posterior flow of 360 veh/h in (27), one obtains the same value. The negative sign of Λ_1 indicates that there is too much simulated flow on route 1, which the calibration tries to reduce by scaling the choice probability of this route by $\exp(\Lambda_1) < 1$.

This type of detailed analysis is hard to conduct for the large real-world test case presented in the next section, which therefore resorts to more aggregate perfor-

24

mance measures. However, the conceptual workings of the calibration are the same as described in this example.

5 Zurich case study

This section presents results from an ongoing real-world case study for the city of Zurich (Flötteröd et al., 2009). First, the deployed simulation system is described in Section 5.1. Second, the Zurich scenario is presented in Section 5.2. Third, the interactions between simulation and calibration are investigated in Section 5.3. Finally, Section 5.4 reports on the validation of the calibrated simulation system.

5.1 Deployed simulation system

The MATSim ("Multi-agent transport simulation toolkit", MATSim, accessed 2009) DTA microsimulation is used for the purposes of this study. Its workings coincide well but not perfectly with the specification of Section 3.1. This situation is likely to be encountered in the calibration of other microsimulations as well. An important aspect of this study is therefore to show that the calibration is robust with respect to (mild) violations of its underlying assumptions.

Consistently with all assumptions of the calibration, MATSim consists of a microscopic and stochastic demand and supply simulator, which are iteratively executed until stationary conditions are attained. The supply simulator is based on a queueing model that is fully consistent with the assumptions of this work (Cetin et al., 2003). The choice dimensions accounted for in the demand simulator are route choice, departure time choice, and mode choice (car vs. no-car). The demand simulator has some unusual features that are discussed in the following. It is described in detail in (Raney and Nagel, 2006).

Continuous choice set generation. The choice set generation and the choice simulation are intertwined in MATSim. The rational behind this is that the choice set should be appropriate in equilibrated network conditions, which are not known a priori. The simulation therefore proceeds in two stages. In the first stage, as from now called the **choice set generation stage**, the choice set is continuously updated in that new plans are generated and other plans are discarded during the iterations. In the second stage, the **choice stage**, the choice set generation is turned off and the demand simulator operates based on fixed choice sets.

Implicit choice distribution. Agents make choices both in the choice set generation stage and the choice stage. In the choice set generation stage, a newly

25

generated plan is selected for execution with probability one. This is necessary because MATSim calculates the utility of a plan only after it is executed; this logic is discussed in the next paragraph. Since the generation of new plans is realized by random variations of existing ones, the guaranteed selection of a newly generated plan generates draws from the set of all plans that can be possibly created by random variations. If no new plan is generated for an agent, one of its existing plans is selected according to a multinomial logit model. In the choice stage, no new plans are generated and the demand simulator only applies the multinomial logit model.

Simulation-based utility function. MATSim uses an all-day utility function that consists of positive terms for the execution of activities and negative terms for travel (Charypar and Nagel, 2005). Utilities are not calculated based on average network conditions but as averages over the experienced utilities of executed plans, which from a calibration perspective implies the same type of approximation as discussed in Section 3.2 when comparing the result of a deterministic network loading of an average demand with the expected network conditions given the actual demand distribution. MATSim averages the experienced utilities by a recursive first-order filter with an innovation weight of 0.1.

Apart from these peculiarities, MATSim constitutes an iterative DTA microsimulator that complies with all assumptions of the proposed calibration.

5.2   Description of test case and uncalibrated simulation

Figure 6 shows the road network of the analysis zone. An all-of-Switzerland network with 60 492 links and 24 180 nodes is used. It is based on a Swiss regional planning network, which has been made ready for simulation purposes based on additional OpenStreetMap network data (Chen et al., 2008).

A synthetic population of travelers for all of Switzerland is available from a previous study (Meister et al., 2008). All travelers have complete daily activity patterns based on microcensus information (SFSO, 2006). The experiments consider only those agents who cross a 30 km (18.6 miles) circle around the center of Zurich at least once during their daily travel, including those agents who stay within that circle for the whole day. In order to obtain a high computational speed, a random 10 % sample is chosen for simulation, which consists of 187 484 simulated travelers. All agents iteratively adapt route choice, departure time choice, and mode choice. Public transit is simulated as described in (Grether et al., 2009), that is, it is assumed that it provides door-to-door connectivity at twice the free speed travel time by car.
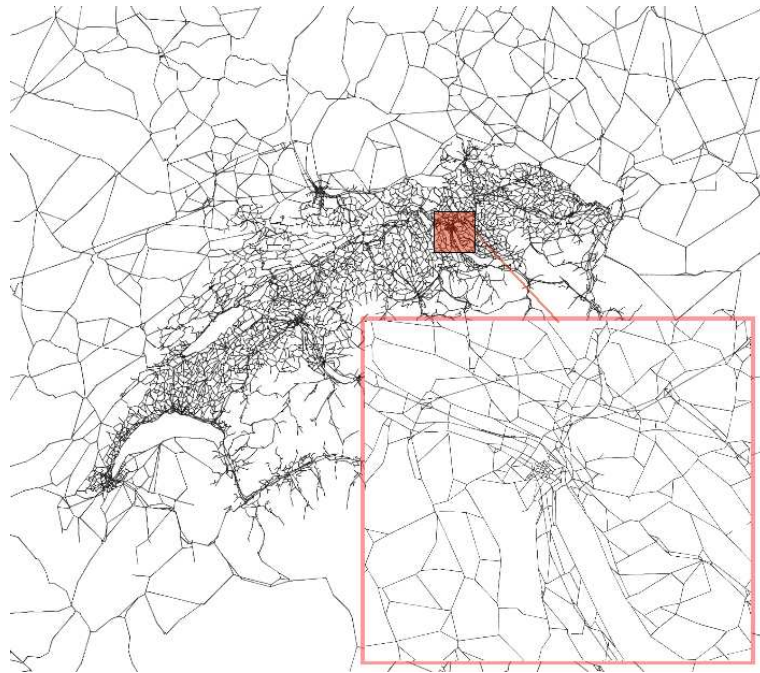
Figure 6: Zurich network

Hourly traffic counts from 161 inductive loop sensors are available from 06:00 to 20:00 of one day. The deviation between measured and simulated traffic counts is both graphically and quantitatively evaluated. For visual inspection, scatter plots such as those given in Figure 7 are used. Every point represents one pair of measured/simulated traffic counts, where the measured value defines the x-coordinate and the simulated value defines the y-coordinate. If all measurements were perfectly reproduced by the simulation, all points would lie on the diagonal with slope one. Deviations from that diagonal signalize inconsistencies between measurements and simulation.

Figure 7 shows scatter plots that are obtained after 500 iterations of uncalibrated simulation. The line above (below) the main diagonal represents simulation values of twice (half) the observed traffic counts (note that the plots are double-logarithmic). Most points are within this (admittedly loose) band, which indicates that the simulation captures the overall situation fairly well. However, there clearly is room for improvement.
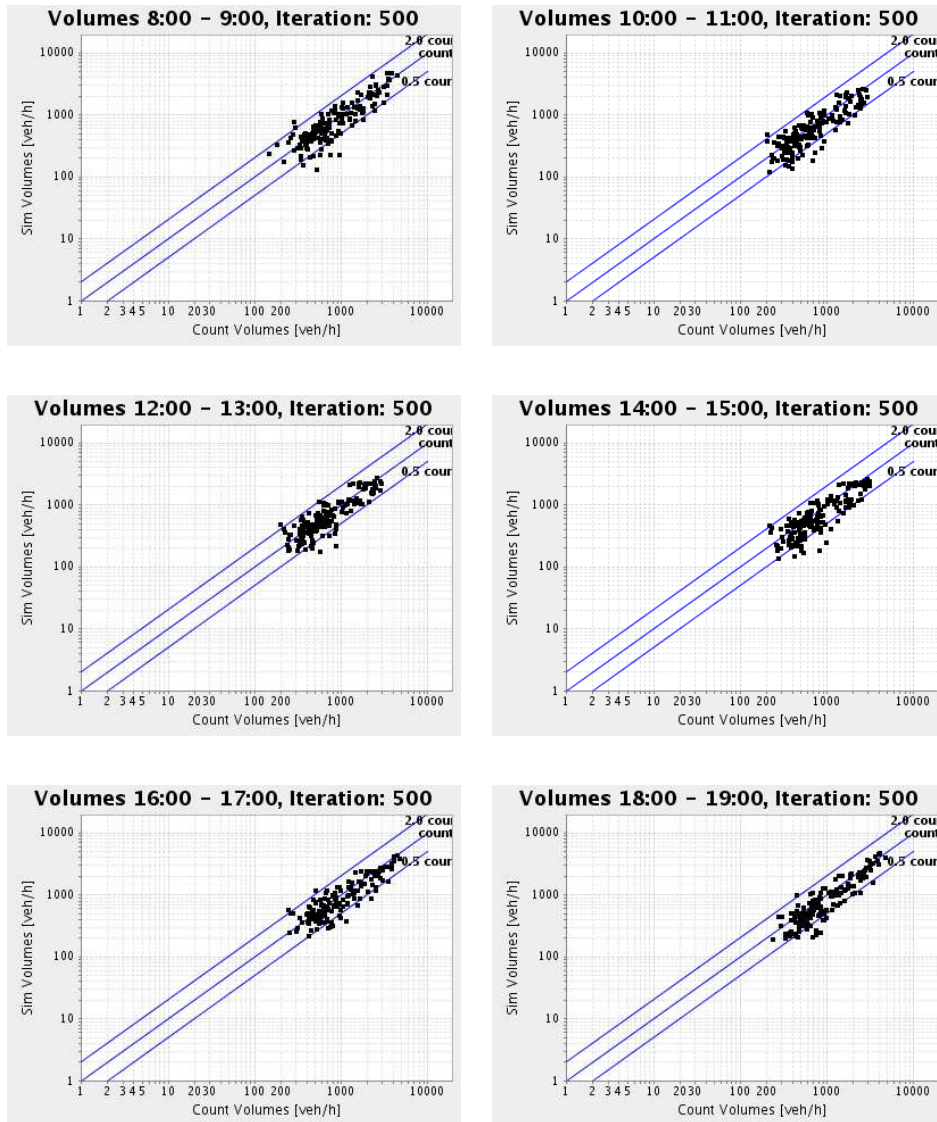
27

Figure 7: Scatter plots for uncalibrated base case

5.3 Inserting the calibration into the simulation

The proposed calibration methodology is implemented in the free Cadyts ("Calibration of dynamic traffic simulations") software package (Cadyts, accessed 2009; Flötteröd, 2009). Cadyts is written with conceptual and technical flexibility in mind in that it offers various modes of interaction with different DTA microsimulations. All experiments reported in this section are based on an application of Cadyts to MATSim.

In this case study, the agent behavior is affected by modifying the utility of their available plans before they make their choices, cf. Section 4.2. The only exceptions are newly generated plans, which are always executed. This implies that these parts of the demand remain uncalibrated during the choice set generation stage and that the calibration takes full effect only in the choice stage.

The evolution of the calibrated simulation over the iterations is visualized in Figure 8, which shows the mean weighted square error (MWSE) of all measurements over the iteration number. This error measure is defined as

$$\text{MWSE} = \left\langle \frac{(y_a(k) - x_a(k))^2}{2\sigma_a^2(k)} \right\rangle_{ak} \tag{30}$$

where $\sigma_a(k)$ is the standard deviation assigned to the sensor data $y_a(k)$ on link a in hour k, $x_a(k)$ is its simulated counterpart, and $\langle \cdot \rangle_{ak}$ indicates an average over all sensor locations and hourly time intervals. This coincides with the log-likelihood function that is assumed in the calibration, which corresponds to the assumption of independent normally distributed measurement errors. The variance of a measurement is calculated as

$$\sigma_a^2(k) = 0.5 \cdot \max\{y_a(k), (25\,\text{veh/h})^2\}, \tag{31}$$

which reflects two considerations. First, there is the assumption that the variance of a measurement error is proportional to the measured value. Second, there is a positive lower bound on the variance, which ensures that very small measurements are not over-weighted and avoids numerical problems in the evaluation of (30). The numerical values used in this specification were experimentally obtained.

When applying the calibration, the system starts in an already equilibrated state that has been attained after 500 uncalibrated iterations. The calibrated simulation is then run for another 500 iterations, i.e., from total iteration number 500 to 1000. Running the calibration jointly with the simulation for another 500 iterations requires approximately $20\frac{1}{4}$ h on a 64 bit Intel Nehalem machine at 2.67 GHz using at most 10 GB of RAM. Not even 9 % of the computing time (approx. $1\frac{3}{4}$ h) are calibration overhead.
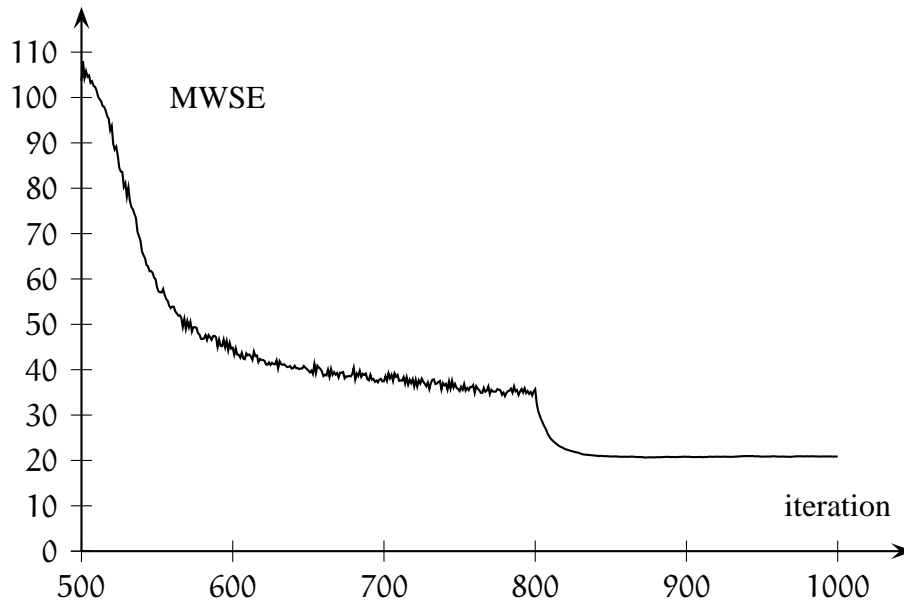
29

Figure 8: Mean weighted square error (MWSE) using all counting stations

Since the system starts already in an equilibrated state, all systematic changes of MWSE in Figure 8 can be attributed to the calibration. The MWSE is quickly reduced from more than 100 in iteration 500 to around 45 in iteration 600. After this, the curve flattens. It is plausible to assume that in the first iterations the calibration "fills up" the measurement locations by arbitrary plans and that in the following iterations the simulation rearranges the plans such that behaviorally more reasonable plans take the place of other plans that have been used by the calibration before.

The choice set generation stage finishes at iteration 800, which generates a jump in the system behavior: Since the immediate execution of newly generated plans is omitted, the calibration can affect the whole plan choice distribution, which results in another improvement of MWSE from around 35 to little more than 20. The variability of MWSE is reduced to almost zero after iteration 800, which is a consequence of the reduced variability in the executed plans once the choice set generation is turned off.

Figure 9 shows scatter plots that are obtained from the last iteration of the calibrated simulation, i.e., iteration 1000. A comparison with the uncalibrated scatterplots of Figure 7 shows that the data points are clearly more centered around the main diagonal. A quantitative evaluation of this effect is possible in terms of the MWSE of Figure 8: The MWSE at iteration 500 corresponds to the scatter plots of Figure 7, and the MWSE at iteration 1000 corresponds to those of Figure

9.

Overall, the calibration generates a clear improvement in measurement fit at an extremely low computational cost. However, this alone does not prove that the calibrated agent behavior becomes more realistic because there are many plausible and not-so-plausible combinations of plan choices that reproduce the measurements equally well. The next section provides cross-validation results that indicate that the calibrated demand is indeed more realistic.

5.4 Cross-validation results

While the previous section clearly demonstrates that the calibration improves the measurement reproduction, this section demonstrates that it does so in a way that also improves the realism of the global traffic situation. This is an important issue that applies to demand calibration from traffic counts in general because this problem is highly under-determined, which implies that there is a large number of demand configurations that reproduce the traffic counts equally well. Recall that the proposed calibration resolves this under-determination by taking the choice logic that is implemented in the simulation system itself as the prior information about the demand. The traffic counts are then added to this information in order to obtain an improved posterior choice distribution.

For cross-validation, the 161 sensor locations are randomly assigned to ten disjoint **validation data sets** of roughly equal size. For each validation data set, there is a corresponding **measurement data set** that contains the traffic counts from all sensors that are not represented by the respective validation data set. For every measurement/validation data set pair, one calibration is conducted, where only the measurement data is made available to the calibration and the corresponding validation data is used to evaluate how well the calibrated demand generates a spatiotemporal extrapolation of the traffic counts.

Figure 10 shows the MWSE trajectories of the measurement data for all ten experiments over the iterations, where all trajectories are normalized to their values at iteration zero for better comparability. Figure 11 shows the same type of curves for the validation data. The similar dynamics of the measurement MWSE values indicate that the calibrated simulation exhibits well-behaved dynamics and generates reproducible results. Overall, the measurement reproduction error is reduced by around 80 % in all cases.

The validation MWSE curves exhibit a greater variability, which can be explained by the lower number of measurements that enter the averaging in (30). Again, the variability is substantially decreased once the choice set generation is turned
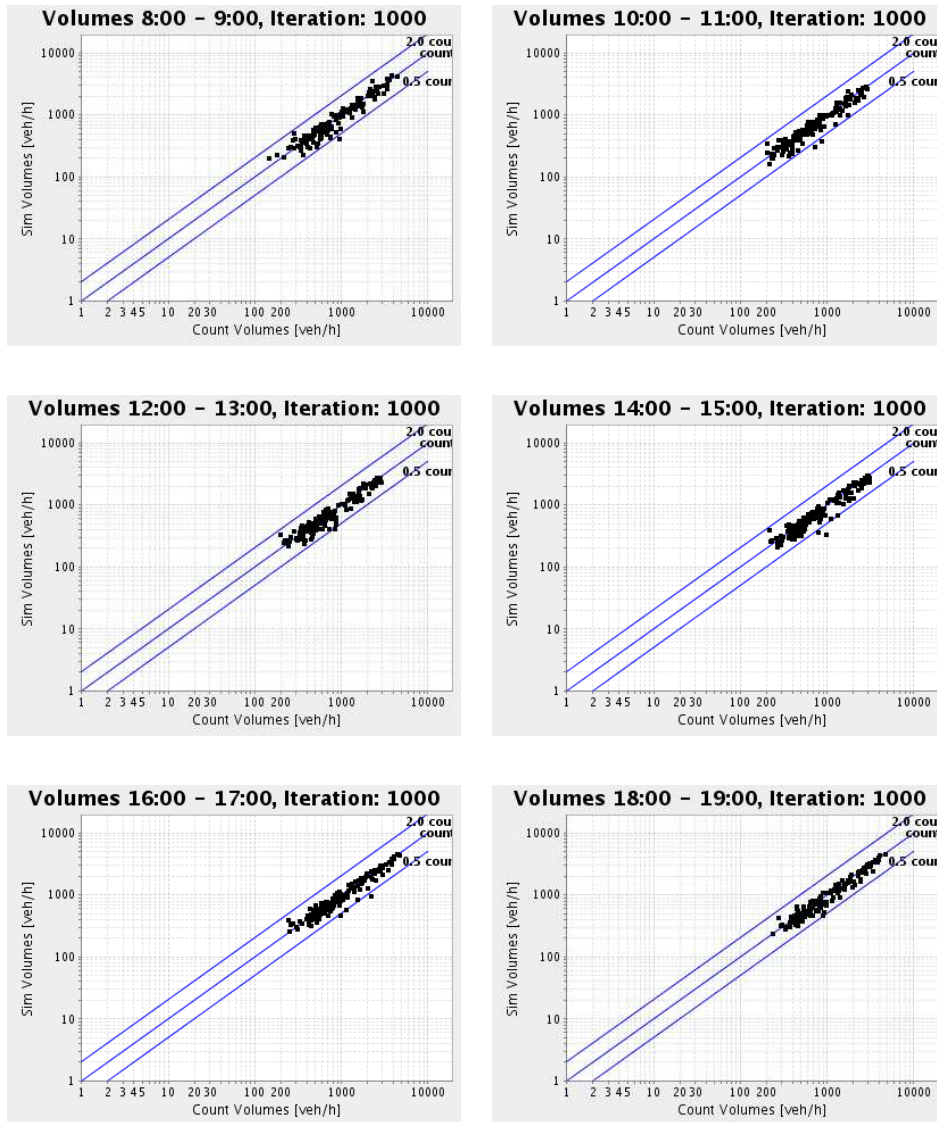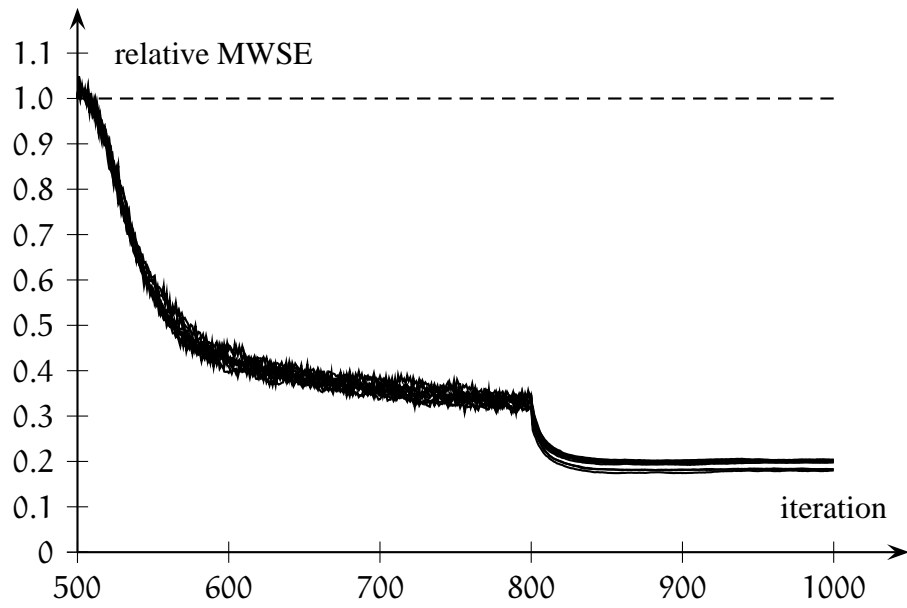
Figure 9: Scatter plots after calibration

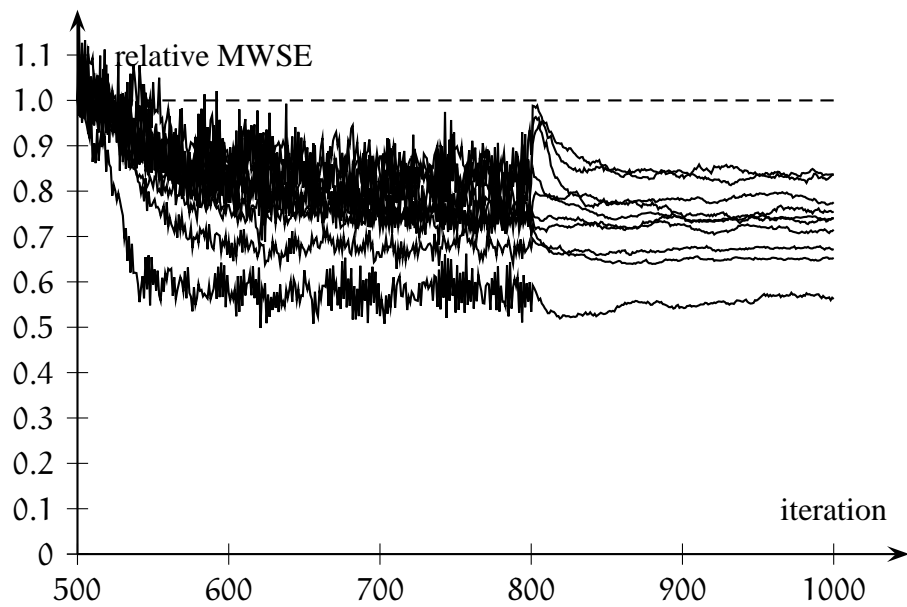Figure 10: Validation results – measurement reproduction

Figure 11: Validation results – measurement extrapolation

33

off. The different experiments attain different MWSE values because disjoint sets of sensor data are evaluated. Overall, an improvement of 15 % to 45 % is attained. This clearly indicates that the local information that is contained in the measurement data is used by the calibration in a way that affects the network-wide agent behavior such that more realistic global network conditions result. One also should keep in mind that the relative positioning of the sensors affects the validation results in that the extrapolation power of the calibration is limited by the spatiotemporal correlations in the network conditions: If the validation sensors are too far away, they simply are not affected any more by the calibration, no matter how well it performs.

These results show clearly that the calibration conducts demand modifications that are structurally meaningful in that they do not only fit the sensor data well but also lead to a global improvement in the system's realism. At this point, the difficulty of the calibration problem that is solved here needs to be stressed. The calibration adjusts simultaneously the route choice, mode choice, and departure time choice of hundreds of thousands of individual travelers in a purely simulation-based environment on a network with many ten thousand links. The number of iterations required to obtain stable and realistic results is in the order of a plain simulation, and the computational overhead introduced by the calibration is almost negligible. The authors are not aware of any other calibration technique that comes close to such results.

6   Summary and outlook

We present a new calibration framework that overcomes many of the simplifying modeling assumptions typically adopted in the calibration of dynamic traffic simulators. Our approach allows for the estimation of arbitrary behavioral patterns at the individual level in a Bayesian setting where traffic counts are combined with a simulation-based representation of the analyst's prior knowledge. The approach is compatible with both an equilibrium-based modeling assumption and a telematics model where drivers are spontaneous and imperfectly informed. Experimental results for a large real-world test case are presented that demonstrate the effectiveness and adequacy of the proposed method. A software implementation of the methodology is freely available on the Internet (Cadyts, accessed 2009).

Our current work focuses on the calibration of behavioral model *parameters* (such as the coefficients of a utility function) from traffic counts. Since this is likely to reach the limits of what can be inferred from this type of measurements, the incorporation of additional sensor data is another important research topic. The free software implementation of the calibration is continuously applied to different

DTA microsimulations, which yields important insights on how to improve the system's conceptual and technical flexibility.

Finally, the joint calibration of demand and supply is a challenge that eventually needs to be tackled. The current demand calibration assumes the supply simulator to be modeled without bias (an assumption it shares with all PFEs and OD matrix estimators that treat the network loading as a deterministic mapping), which should be relaxed in future research.

7   Acknowledgments

A   Maximization of prior entropy

Denote by d_n the total demand of OD pair n and by d_{ni} the demand for path $i \in C_n$, where C_n is the path set of OD pair n. If the demand was integral then the path flows $\mathbf{d} = (d_{ni})$ would be distributed according to

$$P(\mathbf{d}) = \prod_{n=1}^{N} d_n! \frac{\prod_{i \in C_n} (P_n(i|\mathbf{x}(\mathbf{d})))^{d_{ni}}}{\prod_{i \in C_n} d_{ni}!}, \tag{32}$$

where, differently from a standard multinomial distribution, the event probabilities are not fixed but themselves random variables because they depend on the path flows through the network conditions $\mathbf{x}$. Taking the logarithm and applying Stirling's approximation ($\ln Z! \to Z \ln Z - Z$ for large Z), one obtains the **prior entropy function**

$$W(\mathbf{d}) = \ln P(\mathbf{d}) = \sum_{n=1}^{N} \left[d_n \ln d_n + \sum_{i \in C_n} d_{ni} \ln P_n(i|\mathbf{x}(\mathbf{d})) - \sum_{i \in C_n} d_{ni} \ln d_{ni} \right]. \tag{33}$$

(Note that this specification of $W(\mathbf{d})$ differs from (2) in the main text by the addend $\sum_n d_n \ln d_n$, which affects only the maximum value of $W(\mathbf{d})$ but not the

according path flows.) In order to show the equivalence of the global maxima of $W(\mathbf{d})$ (subject to the flow conservation constraints $\sum_{i \in C_n} d_{ni} = d_n \, \forall n$) with the SUE flows, the following observations are made.

1. The maximum value of $W(\mathbf{d})$ subject to the flow conservation constraints is zero: For *fixed* path choice fractions $P_n(i) \, \forall n, i$, $W(\mathbf{d})$ is strictly concave and its maximization subject to the flow conservation constraints yields the path flows $d_{ni} = P_n(i) d_n \, \forall n, i$ and an objective function value of zero. Now consider any candidate combination of *variable* path choice fractions and path flows. Fixing the path choice fractions at their given values, a maximization with respect to the path flows again yields a unique maximum with a zero value of $W(\mathbf{d})$.

2. Every SUE flow is a global maximizer of $W(\mathbf{d})$ subject to the flow conservation constraints: A substitution of the SUE flows $d_{ni} = P_n(i|\mathbf{x}(\mathbf{d})) d_n \, \forall n, i$ yields $W(\mathbf{d}) = 0$, which is the global maximum value.

3. Every global maximizer of $W(\mathbf{d})$ subject to the flow conservation constraints is an SUE flow: Assume that there was a global maximizer $\mathbf{d} = (d_{ni})$ where at least one $d_{ni} \neq P_n(i|\mathbf{x}(\mathbf{d})) d_n$. Fixing the path choice fractions at $P_n(i) = P_n(i|\mathbf{x}(\mathbf{d})) \, \forall n, i$, $W(\mathbf{d})$ is maximized if and only if $d_{ni} = P_n(i) d_n \forall n, i$, which contradicts the assumption.

Items 2 and 3 establish the equivalence of SUE flows and global maxima of $W(\mathbf{d})$ subject to the flow conservation constraints. Note also that the possible existence of multiple global maxima can only result from non-unique SUE flows, which would indicate a modeling problem rather than a flaw in the equivalent maximization problem.

B Maximization of posterior entropy

Before maximizing the posterior entropy function

$$W(\mathbf{d}|\mathbf{y}) = \ln p(\mathbf{y}|\mathbf{d}) + W(\mathbf{d}), \tag{34}$$

the additional requirement of constant demand levels d_n per OD pair n is introduced in the Lagrangian

$$L(\mathbf{d}|\mathbf{y}) = W(\mathbf{d}|\mathbf{y}) + \sum_{n=1}^{N} u_n \left(\sum_{i \in C_n} d_{ni} - d_n \right) \tag{35}$$

where the u_n are the Lagrangian multipliers. Using (33), the derivative of $L(\mathbf{d}|\mathbf{y})$ with respect to d_{mj} (where m is an OD pair and $j \in C_m$) becomes

$$
\begin{aligned}
\frac{\partial L(\mathbf{d}|\mathbf{y})}{\partial d_{mj}} &= \frac{\partial \ln p(\mathbf{y}|\mathbf{x}(\mathbf{d}))}{\partial d_{mj}} + \ln \frac{P_m(j|\mathbf{x}(\mathbf{d}))}{d_{mj}} \\
&+ \sum_{n=1}^{N} \sum_{i \in C_n} \frac{d_{ni}}{P_n(i|\mathbf{x}(\mathbf{d}))} \frac{\partial P_n(i|\mathbf{x}(\mathbf{d}))}{\partial d_{mj}} - 1 + u_m.
\end{aligned}
\tag{36}
$$

Setting this to zero and solving for d_{mj} yields

$$
d_{mj} = \exp(u_m - 1) \exp(\Lambda_{mj} + \Gamma_{mj}) P_m(j|\mathbf{x}(\mathbf{d}))
\tag{37}
$$

where

$$
\Lambda_{mj} = \frac{\partial \ln p(\mathbf{y}|\mathbf{x}(\mathbf{d}))}{\partial d_{mj}}
\tag{38}
$$

$$
\Gamma_{mj} = \sum_{n=1}^{N} \sum_{i \in C_n} \frac{d_{ni}}{P_n(i|\mathbf{x}(\mathbf{d}))} \frac{\partial P_n(i|\mathbf{x}(\mathbf{d}))}{\partial d_{mj}}.
\tag{39}
$$

The $\exp(u_m - 1)$ terms result from a substitution of (37) in $d_m = \sum_{i \in C_m} d_{mi}$:

$$
\exp(u_m - 1) = \frac{d_m}{\sum_{i \in C_m} \exp(\Lambda_{mi} + \Gamma_{mi}) P_m(i|\mathbf{x}(\mathbf{d}))}.
\tag{40}
$$

Inserting this in (37) finally results in the posterior choice probabilities

$$
P_m(j|\mathbf{x}(\mathbf{d}), \mathbf{y}) = \frac{d_{mj}}{d_m} = \frac{\exp(\Lambda_{mj} + \Gamma_{mj}) P_m(j|\mathbf{x}(\mathbf{d}))}{\sum_{i \in C_m} \exp(\Lambda_{mi} + \Gamma_{mi}) P_m(i|\mathbf{x}(\mathbf{d}))},
\tag{41}
$$

which hence prevail at every maximum of the posterior entropy function (subject to the flow conservation constrains $d_n = \sum_{i \in C_n} d_{ni} \, \forall n, i$).

C   Derivation of accept/reject estimator

Given the acceptance probabilities $P_{\text{accept},n}(i)$ defined in (28), the overall probability of a single rejection for agent n is

$$
P_{\text{reject},n} = 1 - \sum_{i \in C_n} P_{\text{accept},n}(i) P_n(i|\cdot).
\tag{42}
$$

Consequently, the probability that i is the first accepted draw can be expressed as

$$\sum_{z=0}^{\infty} (P_{\text{reject},n})^z P_{\text{accept},n}(i) P_n(i|\cdot)$$

$$= \frac{P_{\text{accept},n}(i) P_n(i|\cdot)}{1 - P_{\text{reject},n}} \tag{43}$$

$$= \frac{P_{\text{accept},n}(i) P_n(i|\cdot)}{\sum_{j \in C_n} P_{\text{accept},n}(j) P_n(j|\cdot)},$$

which coincides with the definition in (18) (for zero Γ coefficients).

D Derivation of utility-modification estimator

The individual-level posterior choice distribution (18) constitutes the starting point of this development. It is restated here for ease of reference (with zero Γ coefficients):

$$\Pi_n(i|\mathbf{y}) = \frac{\exp(\Lambda_{ni}) P_n(i|\bar{\mathbf{x}}_{|\mathbf{y}})}{\sum_{j \in C_n} \exp(\Lambda_{nj}) P_n(j|\bar{\mathbf{x}}_{|\mathbf{y}})}. \tag{44}$$

It is assumed that the demand simulator implements an MNL prior choice model (which comprises path-size logit (Ben-Akiva and Bierlaire, 2003) and C-logit (Cascetta et al., 1996) specifications):

$$P_n(i|\bar{\mathbf{x}}) = \frac{\exp[V_n(i|\bar{\mathbf{x}})]}{\sum_{j \in C_n} \exp[V_n(j|\bar{\mathbf{x}})]} \tag{45}$$

where $V_n(i|\bar{\mathbf{x}})$ denotes the systematic utility of plan i as perceived by individual n given the expected network conditions $\bar{\mathbf{x}}$. A substitution of (45) in (44) yields

$$\Pi_n(i|\mathbf{y}) = \frac{\exp[V_n(i|\bar{\mathbf{x}}_{|\mathbf{y}}) + \Lambda_{ni}]}{\sum_{j \in C_n} \exp[V_n(j|\bar{\mathbf{x}}_{|\mathbf{y}}) + \Lambda_{nj}]}. \tag{46}$$

This posterior is structurally identical to the prior. The only difference is that Λ_{ni} is added to the systematic utility of every considered plan i. This utility modification allows to force a demand simulator that implements the prior (45) to immediately draw from the posterior (46), and it avoids the computational overhead of a possibly large number of rejections in the accept/reject procedure.

References

K. Ashok. *Estimation and Prediction of Time-Dependent Origin-Destination Flows*. PhD thesis, Massachusetts Institute of Technology, 1996.

V. Astarita, K. Er-Rafia, M. Florian, M. Mahut, and S. Velan. A comparison of three methods for dynamic network loading. *Transportation Research Record*, 1771:179–190, 2001.

N. Balijepalli, D. Watling, and R. Liu. Doubly dynamic traffic assignment – simulation modeling framework and experimental results. *Transportation Research Record*, 2029:39–48, 2007.

M. Bell. The estimation of origin-destination matrices by constrained generalised least squares. *Transportation Research Part B*, 25(1):13–22, 1991.

M. Bell. Stochastic user equilibrium assignment in networks with queues. *Transportation Research Part B*, 29(2):125–137, 1995.

M. Bell and S. Grosso. Estimating path flows from traffic counts. In W. Brilon, F. Huber, M. Schreckenberg, and H. Wallentowitz, editors, *Traffic and Mobility: Simulation–Economics–Environment*, pages 85–102. Springer, 1999.

M. Bell, W. Lam, and Y. Iida. A time-dependent multi-class path flow estimator. In J.-B. Lesort, editor, *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, pages 173–193, Lyon, France, July 1996. Pergamon.

M. Bell, C. Shield, F. Busch, and G. Kruse. A stochastic user equilibrium path flow estimator. *Transportation Research Part C*, 5(3/4):197–210, 1997.

M. Ben-Akiva and M. Bierlaire. Discrete choice models with applications to departure time and route choice. In R. Hall, editor, *Handbook of Transportation Science, 2nd edition*, Operations Research and Management Science, pages 7–38. Kluwer, 2003. ISBN:1-4020-7246-5.

M. Ben-Akiva and S. Lerman. *Discrete Choice Analysis*. MIT Press series in transportation studies. The MIT Press, 1985.

M. Ben-Akiva, M. Bierlaire, D. Burton, H. Koutsopoulos, and R. Mishalani. Network state estimation and prediction for real–time transportation management applications. *Networks and Spatial Economics*, 1:293–318, 2001a.

M. Ben-Akiva, H. Koutsopoulos, and J. Walker. DynaMIT-P: dynamic traffic assignment model system for transportation planning. In *Proceedings of the 2001 World Conference on Transport Research*, Seoul, Korea, 2001b.

M. Bierlaire. The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables. *Transportation Research Part B*, 36(9): 837–850, 2002.

M. Bierlaire. BIOGEME: a free package for the estimation of discrete choice models. In *Proceedings of the 3rd Swiss Transport Research Conference*, Monte Verita/Ascona, March 2003.

M. Bierlaire and F. Crittin. Solving noisy large scale fixed point problems and systems of nonlinear equations. *Transportation Science*, 40(1):44–63, 2006. doi: 10.1287/trsc.1050.0119.

M. Bierlaire and P. Toint. MEUSE: an origin-destination estimator that exploits structure. *Transportation Research Part B*, 29(1):47–60.

J. Bottom. *Consistent Anticipatory Route Guidance*. PhD thesis, Massachusetts Institute of Technology, 2000.

J. Bottom, M. Ben-Akiva, M. Bierlaire, I. Chabini, H. Koutsopoulos, and Q. Yang. Investigation of route guidance generation issues by simulation with DynaMIT. In A. Ceder, editor, *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, pages 577–600. Pergamon, Jerusalem, Israel, July 1999.

J. Bowman and M. Ben-Akiva. Activity based travel demand model systems. In P. Marcotte and S. Nguyen, editors, *Equilibrium and advanced transportation modelling*, pages 27–46. Kluwer, 1998.

Cadyts. Cadyts web site. http://transp-or2.epfl.ch/cadyts, accessed 2009.

E. Cascetta. Estimation of trip matrices from traffic counts and survey data: a generalised least squares estimator. *Transportation Research Part B*, 18(4/5): 289–299, 1984.

E. Cascetta. A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research Part B*, 23(1):1–17, 1989.

E. Cascetta and G. Cantarella. A day-to-day and within-day dynamic stochastic assignment model. *Transportation Research Part A*, 25(5):277–291, 1991.

E. Cascetta and S. Nguyen. A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research Part B*, 22(6):437–455, 1988.

E. Cascetta and N. Posterino. Fixed point approaches to the estimation of o/d matrices using traffic counts on congested networks. *Transportation Science*, 35(2):134–147, 2001.

E. Cascetta, D. Inaudi, and G. Marquis. Dynamic estimators of origin-destination matrices using traffic counts. *Transportation Science*, 27:363–373, 1993.

E. Cascetta, A. Nuzzolo, F. Russo, and A. Vitetta. A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks. In J.-B. Lesort, editor, *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, pages 697–711, Lyon, France, July 1996. Pergamon.

N. Cetin, A. Burri, and K. Nagel. A large-scale agent-based traffic microsimulation based on queue model. In *Proceedings of the 3rd Swiss Transport Research Conference*, Monte Verita/Ascona, March 2003.

D. Charypar and K. Nagel. Generating complete all-day activity plans with genetic algorithms. *Transportation*, 32(4):369–397, 2005.

Y. Chen, M. Rieser, D. Grether, and K. Nagel. Improving a large-scale agent-based simulation scenario. VSP working paper 08-15, Transport Systems Planning and Transport Telematics Laboratory, Berlin Institute of Technology, https://svn.vsp.tu-berlin.de/repos/public-svn/publications/vspwp/2008/08-15/, accessed 2009, 2008.

C. Daganzo and Y. Sheffi. On stochastic models of traffic assignment. *Transportation Science*, 11(3):253–274, 1977.

A. De Palma and F. Marchal. Real cases applications of the fully dynamic METROPOLIS tool-box: an advocacy for large-scale mesoscopic transportation systems. *Networks and Spatial Economics*, 2:347–369, 2002.

G. Flötteröd. Cadyts – a free calibration tool for dynamic traffic simulations. In *Proceedings of the 9th Swiss Transport Research Conference*, Monte Verita/Ascona, September 2009.

G. Flötteröd. *Traffic State Estimation with Multi-Agent Simulations*. PhD thesis, Berlin Institute of Technology, Berlin, Germany, 2008.

G. Flötteröd and M. Bierlaire. Improved estimation of travel demand from traffic counts by a new linearization of the network loading map. In *Proceedings of the European Transport Conference*, The Netherlands, October 2009.

G. Flötteröd, Y. Chen, M. Rieser, and K. Nagel. Behavioral calibration of a large-scale travel behavior microsimulation. In *Proceedings of 12th International Conference on Travel Behaviour Research*, Jaipur, India, December 2009.

D. Grether, Y. Chen, M. Rieser, and K. Nagel. Effects of a simple mode choice model in a large-scale agent-based transport simulation. In A. Reggiani and P. Nijkamp, editors, *Complexity and Spatial Networks. In Search of Simplicity*, Advances in Spatial Science, chapter 13, pages 167–186. Springer, 2009.

R. Kitamura. An evaluation of activity-based travel analysis. *Transportation*, 15: 9–34, 1988.

R. Kitamura. Applications of models of activity behavior for activity based demand forecasting. In *Proceedings of the Activity-Based Travel Forecasting Conference*, pages 119–150, New Orleans, LA, USA, June 1996.

H. Liu, X. He, and B. He. Method of successive weighted averages (MSWA) and self-regulated averaging schemes for solving stochastic user equilibrium problem. *Networks and Spatial Economics*, 2007. doi: 10.1007/s11067-007-9023-x.

R. Liu. The DRACULA dynamic traffic network microsimulation model. In R. Kitamura and M. Kuwahara, editors, *Simulation Approaches in Transportation Analysis: Recent Advances and Challenges*, pages 23–56. Springer, 2005.

J. Lundgren and A. Peterson. A heuristic for the bilevel origin-destination-matrix estimation problem. *Transportation Research Part B*, 42(4):339–354, 2008.

M. Maher. Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach. *Transportation Research Part B*, 17(6):435–447, 1983.

M. Maher, X. Zhang, and D. Van Vliet. A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows. *Transportation Research Part B*, 35(1):23–40, 2001.

H. S. Mahmassani. Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and Spatial Economics*, 1(3/4):267–292, 2001.

MATSim. MATSim web site. http://www.matsim.org, accessed 2009.

K. Meister, M. Rieser, F. Ciari, A. Horni, M. Balmer, and K. Axhausen. Anwendung eines agentenbasierten Modells der Verkehrsnachfrage auf die Schweiz. In *Proceedings of Heureka '08*, Stuttgart, Germany, March 2008.

K. Nagel, M. Rickert, P. Simon, and M. Pieck. The dynamics of iterated transportation simulations. In *Proceedings of the 3rd Triennial Symposium on Transportation Analysis*, San Juan, Puerto Rico, 1998.

Y. Nie and D.-H. Lee. An uncoupled method for the equilibrium-based linear path flow estimator for origin-destination trip matrices. *Transportation Research Record*, 1783:72–79, 2002.

Y. Nie, H. Zhang, and W. Recker. Inferring origin-destination trip matrices with a decoupled GLS path flow estimator. *Transportation Research Part B*, 39(6): 497–518, 2005.

K. Nökel and M. Schmidt. Parallel DYNEMO: meso-scopic traffic flow simulation on large networks. *Networks and Spatial Economics*, 2(4):387–403, 2002.

B. Raney and K. Nagel. An improved framework for large-scale multi-agent simulations of travel behavior. In P. Rietveld, B. Jourquin, and K. Westin, editors, *Towards better performing European Transportation Systems*, pages 305–347. Routledge, 2006.

S. Ross. *Simulation*. Elsevier, fourth edition, 2006.

H. Sevcikova, A. Raftery, and P. Waddell. Assessing uncertainty in urban simulations using Bayesian melding. *Transportation Research Part B*, 41(6):652–669, 2007.

SFSO. Ergebnisse des Mikrozensus 2005 zum Verkehr. Swiss Federal Statistical Office, Neuchatel, 2006.

Y. Sheffi. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, 1985.

H. Sherali and T. Park. Estimation of dynamic origin-destination trip tables for a general network. *Transportation Research Part B*, 35(3):217–235, 2001.

H. Sherali, R. Sivanandan, and A. Hobeika. A linear programming approach for synthesizing origin-destination trip tables from link traffic volumes. *Transportation Research Part B*, 28(3):213–233, 1994.

H. Sherali, A.Narayan, and R. Sivanandan. Estimation of origin-destination trip-tables based on a partial set of traffic link volumes. *Transportation Research Part B*, 37(9):815–836, 2003.

H. Spiess. A maximum likelihood model for estimating origin-destination models. *Transportation Research Part B*, 21(5):395–412, 1987.

K. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2003.

N. van der Zijpp and C. Lindveld. Estimation of origin-destination demand for dynamic assignment with simultaneous route and departure time choice. *Transportation Research Record*, 1771:75–82, 2001.

H. van Zuylen and L. G. Willumsen. The most likely trip matrix estimated from traffic counts. *Transportation Research Part B*, 14(3):281–293, 1980.

P. Vovsha, M. Bradley, and J. Bowman. Activity-based travel forecasting models in the United States: progress since 1995 and prospects for the future. In *Proceedings of the EIRASS Conference on Progress in Activity-Based Analysis*, Maastricht, The Netherlands, May 2004.

P. Waddell, G. Ulfarsson, J. Franklin, and J. Lobb. Incorporating land use in metropolitan transportation planning. *Transportation Research Part A*, 41(5): 382–410, 2007.

D. Watling and M. Hazelton. The dynamics and equilibria of day-to-day assignment models. *Networks and Spatial Economics*, 3(3):349–370, 2003.

H. Yang. Heuristic algorithms for the bilevel origin/destination matrix estimation problem. *Transportation Research Part B*, 29(4):231–242, 1995.

H. Yang, T. Sasaki, and Y. Iida. Estimation of origin-destination matrices from link traffic counts on congested networks. *Transportation Research Part B*, 26 (6):417–434, 1992.

X. Zhou. *Dynamic Origin-Destination Demand Estimation and Prediction for Off–Line and On–Line Dynamic Traffic Assignment Operation*. PhD thesis, University of Maryland, College Park, 2004.