A framework for the data-consistent deployment of urban microsimulations

Gunnar Flötteröd * Michel Bierlaire *

September 4, 2010

Report TRANSP-OR 100904
Transport and Mobility Laboratory
Ecole Polytechnique Fédérale de Lausanne
`transp-or.epfl.ch`

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

1 Introduction

Microsimulation-based models of urban systems have proven to be powerful tools for prediction and scenario analysis, with a particular yet continuously expanding focus on transportation and land use. They bring along a high level of detail, but they also come at the cost of enormous data needs for their estimation. This article develops a framework for the continuous deployment of a microsimulation-based urban model that integrates existing and emerging data sources.

The text is structured in two parts. The first part, which consists of Sections 2 and 3, focuses on modeling and simulation: First, Section 2 defines the considered urban processes and casts them into a basic formal framework. Second, Section 3 adopts a microsimulation-based perspective on these processes.

The integration of real data into an urban microsimulation is the topic of the second part, which consists of Sections 4 through 6: First, Section 4 defines the respective terminology and indicates two important and new data sources. Second, Section 5 discusses the urban state estimation problem, with a focus on the different time scales on which an urban system unfolds. Third, Section 6 elaborates on the parameter estimation problem for urban models, with a focus on the advantages and difficulties of estimating interacting model components in an integrated manner. Finally, Section 7 summarizes the article.

2 Urban systems

An urban system consists of several interacting components, which are outlined in the following. See Wegener (2004) for a more comprehensive introduction to integrated transportation and land used models and Ghauche (2010) for a recent review with an activity-based modeling focus (Bowman and Ben-Akiva, 1998). Three processes are crucial to a microsimulation-based urban modeling approach: activity participation, transportation, and relocation. Strongly related to these, one may account for energy consumption, the economy, environmental effects, and social interactions. The outline given below and depicted in Figure 1 focuses on the mutual interactions of the basic processes activity participation, transportation, and relocation.

Relocation. The relocation model takes as exogenous inputs the socio-economics of households and firms, their long-term needs and strategies, the development of the building infrastructure, and possible regulations regarding its use. Its endogenous input are the accessibility measures Z_{acc} obtained from

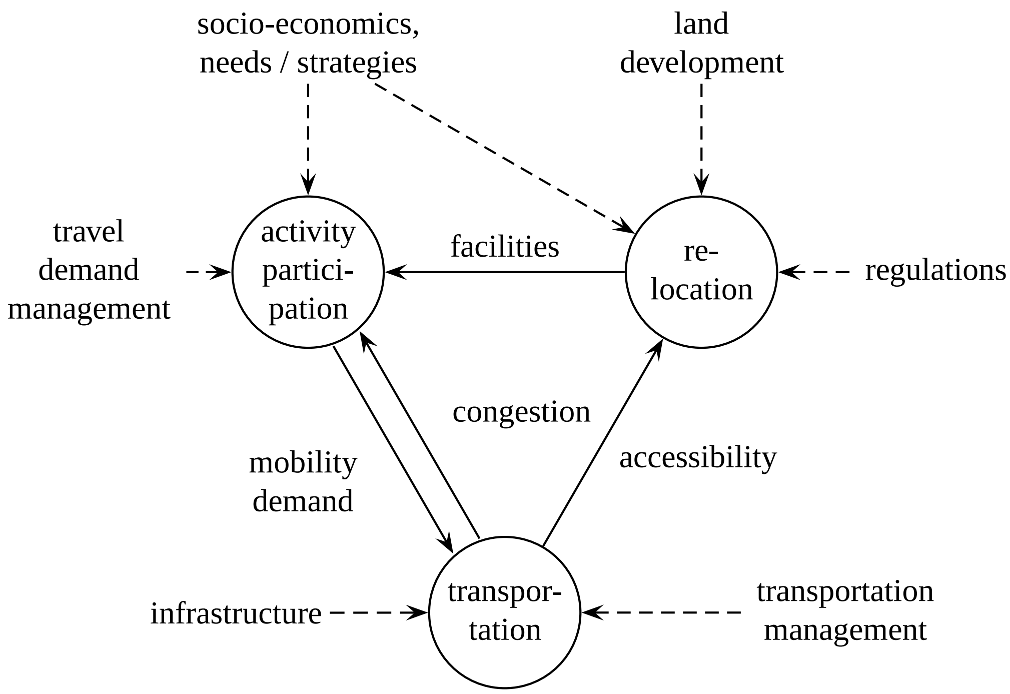

Figure 1: Interactions between activity participation, transportation, and relocation

a transportation model. The relocation model captures how households select their dwellings, how businesses select their offices, how land prices adapt in reaction to this (this may involve a separate economic model), and how these in turn affect the relocation decisions of all involved actors. We denote by X_{reloc} the allocation of all households and firms to all buildings in the system. The output of the relocation model are the facilities Z_{fac}, which provide activity opportunities to households and firms.

Activity participation. The activity participation model takes as exogenous inputs the socio-economics of households and firms, their long-term needs and strategies, and possible travel demand management measures. Its endogenous input are the facilities Z_{fac} defined by the relocation model. Household members conduct activities in different places, including working, regular shopping, and spontaneous leisure activities, all of which may require them to travel to the respective facilities. Firms obtain production inputs such as raw materials or components to be assembled and deliver their products. We denote by X_{act} the activity and travel plans of all households and firms in the system. The output of the activity participation model(s) are sequences of (desired) trips through the transportation network for different modes (or mode combinations) and demand sectors, which we collectively refer to as the mobility demand Z_{mob}.

Transportation. The transportation model takes as exogenous inputs the transportation infrastructure and possible transportation management measures. Its endogenous input is the mobility demand Z_{mob} from the activity participation model. The transportation model represents the physical world of mobility and describes how the mobility demand is served by the transportation infrastructure. The transportation model captures congestion, the temporary over-utilization of the system. We denote by X_{transp} the state of the transportation system, including all mobile entities (vehicles, buses, trains, ...), their occupations, and possibly the internal states of intelligent control mechanisms (such as adaptive traffic lights). The transportation model has two outputs: it feeds back congestion information Z_{cong} to the activity participation model, and it feeds back the resulting changes in location accessibility Z_{acc} to the relocation model.

Formally, we denote by $Z_{mob} = G_{mob}(X_{act})$ the mapping of activity participation on mobility demand, by $Z_{cong} = G_{cong}(X_{transp})$ and $Z_{acc} = G_{acc}(X_{transp})$ the mapping of the transportation system's state on congestion and accessibility, and by $Z_{fac} = G_{fac}(X_{reloc})$ the mapping of the building infrastructure's use on the availability of facilities. Collecting all $Z_{...}$, $X_{...}$, and $G_{...}$, one obtains the **process interaction equations**

$$Z = G(X). \tag{1}$$

The activities $X_{act} = F_{act}(Z_{cong}, Z_{fac})$ are a function of congestion and available facilities. The transportation system's state $X_{transp} = F_{transp}(Z_{mob})$ evolves depending on the mobility demand. The building usage $X_{reloc} = F_{reloc}(Z_{acc})$ is a function of the accessibilities. Collecting all $F_{...}$, this yields the **process state equations**

$$X = F(Z). \tag{2}$$

Combining (1) and (2), one obtains

$$X = F(G(X)), \tag{3}$$

which fully specifies the state X of the urban model in terms of a fixed-point relationship, which states that all processes evolve consistently with each other. We deliberately omit exogenous factors for notational simplicity and postpone the introduction of model parameters to Section 4.

An explicit introduction of the time dimension into this model is postponed to Section 5. For now, we observe that the presented notation allows for both a static equilibrium model or a dynamic (either equilibrium or out-of-equilibrium) model: One may think of X and Z as time-independent long-term average values and of F and G as likewise static functions. This turns (3) into an equilibrium model. One may also think of X and Z as time-dependent values $X = \{x(t)\}_t$ and $Z = \{z(t)\}_t$ with t being the time dimension. In this case, (3) does not necessarily call for an equilibrium but at least for the mutually consistent dynamic evolution of all processes.

3 Microsimulation

A large number of urban microsimulations has evolved in the last decades. To name a few, there are the transportation microsimulations DynaMIT (Ben-Akiva et al., 1998) and MATSim (Raney and Nagel, 2006), the activity participation simulators TASHA (Miller and Roorda, 2004) and Albatross (Arentze and Timmermans, 2004), and the more comprehensive land use simulators UrbanSim (Waddell and Ulfarsson, 2004) and ILUTE (Salvini and Miller, 2005). The software platform OPUS is a recent effort to provide a technical framework for integrated urban microsimulations (Waddell et al., 2005).

Microsimulation can be seen both as a modeling paradigm and a model solution technique, and both perspectives apply in the context of urban models.

Microscopic modeling. If the system under consideration consists of interacting entities, then a modeling approach that captures these entities individually is structurally consistent. This holds in particular if the system is (i) coarse-grained in that a continuous-limit perspective that aggregates individual entities into real-valued quantities is not appropriate and/or (ii) heterogeneous in that the entities differ too much from each other to be represented by a limited number of homogeneous groups. If these properties do not apply then a macroscopic model may sometimes be preferable, for example in thermodynamics. In land use and transportation, however, there is broad agreement that both the coarse granularity of and the differences between the involved actors favor a microscopic modeling approach (Nagel and Axhausen, 2001). Last but not least, microscopic models deal with entities that have counterparts in the real world, which makes them more intuitive and easier to communicate than abstract systems of equations.

Microscopic simulation. Even if macroscopic modeling is feasible, it usually is uncertain and hence involves distributional assumptions about quantities that cannot exactly be determined. The solution of such models requires to solve possibly complicated integrals over these distributions. This mere computational problem can be solved by simulation in the numerical sense (Ross, 2006): instead of evaluating the integral directly, a number of random realizations is generated, the resulting indicators are calculated, and their average is used as an approximation of the integral. The substantial uncertainty clinging to land use and transportation models in combination with the impossibility to evaluate them in closed form motivates a simulation-based approach from this technical perspective as well. In particular, the uncertainty of microscopically modeled behavior calls for a probabilistic analysis, which for all but the most simple models can only be conducted through simulation.

The microscopic approach is essentially characterized by disaggregation. However, there may be different degrees of disaggregation. We make this observation formally concrete for the individually simulated actors in the system. These **agents** are indexed by $n = 1 \ldots N$, where N is the size of the simulated population. Consistently with the framework of Section 2, the state X_n of agent n consists of its activity state $X_{act,n}$, its transportation state $X_{transp,n}$, and its relocation state $X_{reloc,n}$.

$X_{act,n}$ represents the activity and travel plan of the agent. $X_{transp,n}$ describes if, where, and how the agent is currently mobile in the transportation system. $X_{reloc,n}$ defines the dwelling of the agent (housing for a household and, e.g., office space for a firm). If the agent represents more than one individual (members of a household, employees of a firm), then the respective state variables represent all of these individuals. The process states X_{act}, X_{transp}, and X_{reloc} comprise the individual-level components for all members of the population but may contain additional information, depending on the scope of the whole simulation system.

The disaggregate activity participation, transportation, and relocation models for agent n are written in the following way:

$$X_{act,n} = F_{act,n}(Z_{transp,n}, Z_{reloc,n}, Z_{cong}, Z_{fac}) \quad (4)$$
$$X_{transp,n} = F_{transp,n}(Z_{act,n}, Z_{cong}) \quad (5)$$
$$X_{reloc,n} = F_{reloc,n}(Z_{fac}, Z_{acc}). \quad (6)$$

Equation (4) states that the activity and travel plans $X_{act,n}$ of an agent n depend on the congestion status Z_{cong} of the transportation network, the available facilities Z_{fac}, and the transportation and relocation state

$$Z_{transp,n} = G_{transp,n}(X_{transp}) \quad (7)$$

and

$$Z_{reloc,n} = G_{reloc,n}(X_{reloc}) \quad (8)$$

of this very agent. (An explanation of the $G_n(X)$ notation follows immediately. For now it may be read as $G_n(X) = X_n$.) Equation (5) expresses the transportation state $X_{transp,n}$ of agent n as a function of the congestion state Z_{cong} of the transportation system and its activity and travel plan

$$Z_{act,n} = G_{act,n}(X_{act}). \tag{9}$$

Finally, (6) expresses the relocation state $X_{reloc,n}$ of agent n as a function of the available facilities Z_{fac} for relocation and their accessibilities Z_{acc}. Recall that all of these models may be static or dynamic, as explained in the last paragraph of Section 2.

The degree of model disaggregation, which has important implications for the consistency of an individual agent's state variables, differs between mesoscopic and truly microscopic models.

Mesoscopic models. In the mesoscopic approach, disaggregation takes place within the processes, but the interactions between processes are still based on aggregate information. This disconnects individual entities in different processes in that the $G_n(X)$ functions in (7)-(9) anonymously sample/infer the state of an agent in one process when feeding it into another process:

- $G_{transp,n}(X_{transp})$ reconstructs what agent n experiences in the transportation system, but without reference to a particular entity in that system. Typically, this is done by following the agent's path based on aggregate travel time information.
- $G_{reloc,n}(X_{reloc})$ assigns a relocation state to agent n based on the population's distribution in the relocation model.
- $G_{act,n}(X_{act})$ infers agent n's activity and travel plan from the distribution of all plans in the activity participation process. This is typically done by (i) breaking down the activity patterns into trip sequences (ii) aggregating these trips into origin/destination (OD) matrices (this would be Z_{mob} in the process-based perspective), and (iii) re-sampling individual trip-makers from these matrices.

Because of their aggregate process interactions, mesoscopic models can integrate macroscopic model components relatively naturally. Their major deficiency is their limited ability to relate individual-specific information obtained in one process to individuals in other processes.

Microscopic models. Microscopic models maintain the integrity of the simulated entities, both in the processes and their interactions. Here, the $G_n(X)$ functions in (7)-(9) are true identities: $G_{transp,n}(X_{transp})$, $G_{reloc,n}(X_{reloc})$, and

$G_{act,n}(X_{act})$ extract those components of X_{transp}, X_{reloc}, and X_{act} that uniquely belong to agent n.

While the microscopic approach guarantees consistency between the agent representations in different processes, it does not keep disaggregate model components from interacting through macroscopic quantities. For example, the decision of a household to move into a certain region may be based on aggregate characteristics like shopping facility density and noise levels, the decision of a land developer to construct a new building may depend on the average propensity of the targeted household segment towards this type of dwelling, or the mobility behavior of an individual may depend on average travel times in the transportation network.

The notation $Z_{transp,n}$, $Z_{reloc,n}$, and $Z_{act,n}$ in (4)-(6) allows to treat mesoscopic and microscopic models within the same formal framework. Unless stated otherwise, the following discussion therefore applies to both model classes. Furthermore, all statements in terms of the process-based notation (1), (2) can be mapped on mesoscopic or microscopic models through appropriate composition of the state (interaction) variables X and Z.

4 Estimation

This section consists of two parts. First, Subsection 4.1 distinguishes the notions of parameter estimation and state estimation and introduces some basic notation. Second, Subsection 4.2 presents two emerging data sources of particular relevance for the estimation of urban microsimulations.

4.1 Formal framework

We distinguish between the estimation of parameters and states. Parameters are by definition time-independent. The parameter estimation problem is to identify temporally stable system properties that identically apply in the future and for different scenarios. States, on the other hand, evolve over time. The state estimation problem is to identify a complete configuration of the system's endogenous variables. In either case, the estimation combines structural model information with observations from the real system.

Parameter estimation. The process model is now assumed to depend on the parameters β, i.e., (2) is augmented into

$$X = F(Z|\beta). \tag{10}$$

β comprises components β_{act}, β_{transp}, and β_{reloc} for the respective processes. The parameter estimation problem is to infer a β^* that is most consistent with the model structure and all available data Y. Denoting this estimator by $\mathcal{B}$, we write

$$\beta^* = \mathcal{B}(Y). \tag{11}$$

Typical methods implemented in $\mathcal{B}$ are Bayesian or Maximum Likelihood estimation (Greene, 2003). $\mathcal{B}$ also comprises all available information about the model structure, in particular the interplay of the state (interaction) variables X and Z through the process (interaction) functions F and G.

State estimation. Even if the model parameters are well calibrated, some uncertainty about the model state X remains. The measurements Y can be used to reduce this uncertainty. Denoting the respective state estimator by $\mathcal{X}$, we write

$$X^* = \mathcal{X}(Y|\beta) \tag{12}$$

for the estimated state X^*. Typical methods implemented in $\mathcal{X}$ are Kalman Filtering or Bayesian inference (Arulampalam et al., 2002; Chui and Chen, 1999). Again, the estimator $\mathcal{X}$ comprises all available information about the model structure.

From (12), it is clear that the state estimation problem is solved conditionally on the parameter estimation problem. The converse setting, where the parameters are estimated conditionally on the estimated states, also has some practical relevance and is visited later in Subsection 6.2.

4.2 New data sources

The amount of data needed to calibrate a model depends on its granularity. Macroscopic models that function in terms of aggregate quantities can be estimated based on aggregate data alone. Microscopic models of individual behavior need to be estimated from disaggregate data. This turns urban microsimulations into data-hungry systems, and instruments for the affordable provision or substitution of such data are essential for their estimation. In the following, we indicate two emerging and particularly relevant data sources, smart phones and vehicle identification systems. Note, however, that all established data sources, ranging from postal surveys that query complete activity and travel patterns to inductive loops that merely count vehicles on roads, should be deployed in combination with these new technologies.

Smart phones. These devices collect a wealth of information about their users' environment, travel, and activities. This includes GPS (global positioning system) tracks and the MAC (media access control) addresses of nearby devices as well as all communications and running applications on the phone, and it can go as far as taking visual and acoustic samples of the environment. Methodological work is underway for the identification of the user's current travel and activity from smart phones, and it is reasonable to anticipate that the smart phones of selected individuals will soon serve as reliable travel and activity sensors in the urban system (Bierlaire et al., 2010; Bohte and Maat, 2009; Hato, 2010; Hurtubia et al., 2009; Raj et al., 2008; Schüssler and Axhausen, 2009).

Vehicle identification systems. The identification of a vehicle at one or several locations in the network reveals individual-level information about the chosen destination, route, and departure time of the driver. Vehicle identification systems usually rely on cameras and/or transponder-based short range communications. These systems are crucial for electronic toll collection systems, and hence this data source can be accessed wherever such a system is installed. The estimation of travel behavior from vehicle identification systems is an active field of research that has already resulted in the implementation of operational prototypes (Antoniou et al., 2006; Vaze et al., 2009; Zhou, 2004).

Both data sources continuously reveal individual-level behavior at a relatively low cost once the system is installed. This makes them attractive not only for the estimation of model parameters but also for real-time state estimation purposes. Note, however, that in the urban simulation context, the objective is to estimate disaggregate behavior without a one-to-one mapping from simulated to real actors. This differs from other applications of the same sensor technology that require person-specific estimates. For example, a smart phone may internally keep track of its user's activity preferences in order to provide customized, context specific information.

5 State estimation

The continuous tracking of the urban state allows to manage the system more effectively in response to its most recent internal changes. An important problem in this context is that urban processes evolve at vastly different time scales. Based on an analysis of these time scales in Subsection 5.1, a rolling horizon state estimation framework is developed in Subsection 5.2.

5.1 Time scales of an urban microsimulation

We distinguish short-, medium-, and long-term dynamics.

Short-term dynamics. This refers to dynamics within a day. The physical transportation system evolves on the time scale of minutes or even seconds. Its state on subsequent days may be considered as decoupled if congestion does not persist over night and vehicles are parked in the same location every night. Travel behavior and activity participation are, within limits, also variable within a single day, either in reaction to exogenous events or in reaction to the transportation system's performance.

Medium-term dynamics. This refers to dynamics across a limited number of days. Many aspects of activity participation and the resulting travel behavior are linked across a number of days. Households and firms schedule their maintenance activities across weeks or even months. Travel behavior is based on anticipated network conditions, which are extrapolated from information collected during many previous days. Relocation is also relevant on medium-term time scales in that it continuously changes the facilities that are available for activity participation.

Long-term dynamics. This refers to dynamics in the order of years. It is the time scale of the relocation model. However, even if individual relocations are unlikely to take place more than once per year, population relocation is a continuous process that affects activity and travel behavior also on medium time scales. The accessibility feedback from the transportation system on relocation, however, occurs with such an inertia that the relocation model can be considered as decoupled from the transportation system on short and medium time scales.

This classification leaves out all but the three central processes identified in Section 2. Apart from transportation and short-term activity participation, communication is another important short-term process. This comprises centralized information distribution systems (radio, Internet) as well as direct communications along the edges of social networks. Energy consumption is to a large extent derived from activity participation and hence occurs on short time scales as well. Ecological and economical processes and the evolution of social networks, however, may safely be constrained to medium and long time scales.

Most of the existing literature on state estimation in the urban context focuses on the tracking of the physical transportation system's state $\mathsf{X}_{\mathsf{transp}}$, e.g., (Chrobok et al., 2003; Tampere and Immers, 2007; Wang and Papageorgiou, 2005). The real-time tracking of the behavioral states X_{act} is mainly constrained to limited aspects of the derived travel patterns such as OD matrices or path flows (Ben-Akiva et al., 1998; Bell et al., 1997; Zhou and Mahmassani, 2007). A mentionable exception is Flötteröd (2008); Flötteröd et al. (2010), where full-day activity and travel plans are estimated from traffic counts and supplementary model information. The relocation state X_{reloc} may be considered as completely measurable based on sufficient data access rights and enough time to process it.

5.2 Rolling horizon framework

The simulation-based nature of an urban microsimulation model renders the application of computationally and mathematically convenient recursive filtering techniques infeasible. We therefore opt for a rolling horizon framework. Here, it is advantageous to consider transportation, activity participation, and relocation separately.

Estimation of X_{transp}. A estimation of the transportation system's state requires, on the most disaggregate level, to track individually simulated (yet anonymous) transportation units (vehicles, pedestrians, ...). This requires to account for high-resolution dynamics in the order of seconds or minutes. A reasonable length of the estimation time horizon is typically between 30 and 60 minutes. Formally, the transportation system state estimator is written as

$$X^*_{transp} = \mathcal{X}_{transp}(Y_{transp}|Z_{mob}; \beta_{transp}) \tag{13}$$

where Y_{transp} comprises all sensor data that is relevant for the transportation state estimation problem. The estimator depends on the mobility demand Z_{mob} and the parameters β_{transp} of the transportation system. The congestion information Z_{cong} derived from the transportation system's state is at least locally visible to travelers and can be made more globally accessible through information distribution systems (radio, Internet) and communications in social networks.

Estimation of X_{act}. Activity and travel scheduling happen both within-day and day-to-day. However, it is not advisable to estimate daily activity schedules during a limited time window within a day because activity scheduling is not a temporally linear process. The complex internal logic of daily activity schedules, including their various constraints, require to schedule and estimate a day as a whole (Bowman and Ben-Akiva, 1998). Formally, the activity state estimator is written as

$$X^*_{act} = \mathcal{X}_{act}(Y_{act}|Z_{cong}, Z_{fac}; \beta_{act}) \tag{14}$$

where Y_{act} comprises all sensor data that is relevant for the activity estimation problem. The estimator may be conditional on the activity participations of previous days and depends on the congestion Z_{cong}, the available facilities Z_{fac} for activity participation, and the parameters β_{act}. (The activity and travel participations of firms may be more difficult to observe than those of individual volunteers for reasons of market competition, although they are likely to be more structured and better documented.)

Estimation of X_{reloc}**.** It is reasonable to assume that, given enough time and access to the necessary data bases, the relocation state of the urban system is directly measurable. However, this is possible only with a lag, and a real-time tracking of urban relocations appears infeasible. We may assume that relocations are measurable in yearly intervals and that the relocations of the upcoming year are planned with such a lag that that they are not affected by the events within that year. This allows to simulate all upcoming relocations at once at the beginning of the year, to derive a day-by-day relocation sequence from this, and to feed the resulting facility information Z_{fac} exogenously into the mid- and short-term state estimation processes. Directly observable facility changes such as openings of shopping malls can also be exogenously incorporated.

These considerations suggest to estimate the urban system state on a daily basis, where the transportation system's state is tracked with a rolling horizon within the day and the activity and travel behavior is estimated without a rolling horizon for the day as a whole, possibly conditional on the activities of previous days. Figure 2 gives an overview, which is detailed in the following.

The physical transportation system and the activity and travel behavior need to be estimated in mutual dependency; they are coupled through the mobility demand and the congestion information. The coupling between the relocations and the activity and travel behavior estimator is unidirectional in that relocation events are predicted infrequently and then disaggregated across the time line in order to allow for a continuous evolution of the boundary conditions for the activity and travel behavior.

We close this section with the observation that a microsimulation-based state estimator is unlikely to represent distributional information differently but through samples. Considering that a single sample represents an entire urban state, we are facing a computationally enormous problem. Unless randomness is artificially reduced, this is likely to require (loosely coupled) parallel computing efforts where a number of computers calculates one realization of the urban state each. The sample-based approach connects the urban state estimator to particle filtering

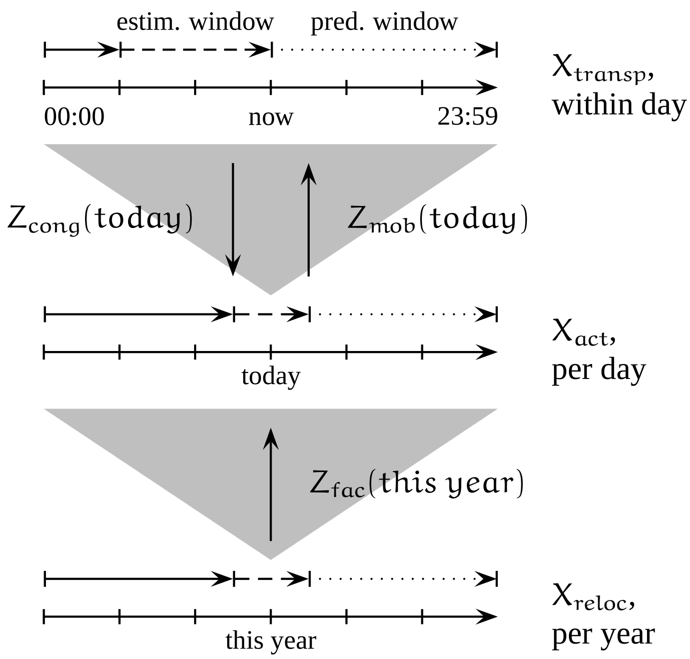

Figure 2: Rolling horizon urban state estimation

techniques (Arulampalam et al., 2002), however, on a vastly increased scale of computational complexity.

6 Parameter estimation

We now consider the problem of how to calibrate the structural model parameters β from some data set Y. As stated in Section 4.1, one may assume a comprehensive estimator $\mathcal{B}$ to be given that estimates all parameters β jointly from all available data Y. This, however, is a rather extreme case, and it is more likely to assume that different components of the whole model are estimated separately and possibly conditional on each other. In Subsection 6.1, we clarify this observation for the process-based decomposition of Section 2. Subsections 6.2 and 6.3 then discuss two techniques to approximately account for the process interactions when estimating submodel parameters.

6.1 Parameter estimation for interacting processes

In a process-based decomposition, the activity and travel participation, transportation, and relocation parameters β_{act}, β_{transp}, and β_{reloc} are estimated from subsets Y_{act}, Y_{transp}, and Y_{reloc} of Y that are relevant to the respective processes, and the process boundaries Z are considered as given:

$$\begin{aligned}
\beta^*_{act} &= \mathcal{B}_{act}(Y_{act}|Z_{cong}, Z_{fac}) & (15)\\
\beta^*_{transp} &= \mathcal{B}_{transp}(Y_{transp}|Z_{mob}) & (16)\\
\beta^*_{reloc} &= \mathcal{B}_{reloc}(Y_{reloc}|Z_{acc}). & (17)
\end{aligned}$$

Since the boundaries of each process depend on the states of all adjacent processes, which in turn depend on the respective parameters, one does not face three independent parameter estimation problems but one large, coupled problem. The least one can do to account for this coupling is to repeatedly solve the individual estimation problems conditional on each other until a state of mutual consistency is attained.

Since we are dealing with a microsimulation, the estimation of behavioral models from individual-level data sources deserves particular attention. Denoting by $m = 1 \ldots M$ the observed individuals in reality, a typical parameter estimation approach is to define $\mathcal{B}$ as a maximum likelihood estimator

$$\mathcal{B}(Y) = \arg\max_{\beta} \mathcal{L}(\beta) \qquad (18)$$

with the log-likelihood function

$$\mathcal{L}(\beta) = \sum_{m=1}^{M} \ln p(Y_m|\beta). \qquad (19)$$

Again, this estimator can be decomposed by process. Assuming that the individual-level observations Y_m are related to the parameters β only through the individual-level state X_m, (4)-(6) yield

$$\begin{aligned}
\mathcal{L}(\beta_{act}) &= \sum_{m=1}^{M} \ln p(Y_{act,m}|Z_{transp,m}, Z_{reloc,m}, Z_{cong}, Z_{fac}; \beta_{act}) & (20)\\
\mathcal{L}(\beta_{transp}) &= \sum_{m=1}^{M} \ln p(Y_{transp,m}|Z_{act,m}, Z_{cong}; \beta_{transp}) & (21)\\
\mathcal{L}(\beta_{reloc}) &= \sum_{m=1}^{M} \ln p(Y_{reloc,m}|Z_{fac}, Z_{acc}; \beta_{reloc}). & (22)
\end{aligned}$$

In the context of choice models, the individual-level boundary conditions Z_m can be considered as person-specific attributes and choice set information, whereas the aggregate boundary conditions Z define further attributes of the alternatives. Together, they define the choice context of the observed individual.

The possibility to measure the choice context from smart phones or vehicle identification systems, in particular in terms of non chosen alternatives and perceived attributes of the alternatives, is limited. It depends on unobservable information the individual has gathered through experience in the urban environment and, due to the anticipatory nature of decision making, on attributes that are spatially and temporally remote and hence not accessible to the sensor. Since this information is crucial to the estimation of behavioral models (Ben-Akiva and Lerman, 1985; Train, 2003), it appears plausible to impute the context information within the urban simulation. This requires to estimate behavioral models conditional on their simulated environment – which in turn is defined through the estimated behavior.

Some previous research was conducted in this context. Balakrishna (2006) reports on the joint calibration of travel demand and traffic flow parameters in the DynaMIT traffic microsimulator. Sevcikova et al. (2007) calibrate the UrbanSim land use simulator, which comprises, amongst other components, a relocation model and an external transportation model. Methodologically, these efforts are constrained to the application of black box calibration techniques, which by definition exploit no problem structure. Also, they are limited to the time scales of their respective processes: DynaMIT operates in the order of days, whereas UrbanSim runs from year to year.

The remainder of this section discusses two integrated parameter estimation approaches that account for system structure and different time scales. First, Section 6.2 describes a combined state and parameter estimator that loosens the process interactions based on data. Second, Section 6.3 proposes response surfaces and metamodels as means to improve the tractability of integrated parameter estimation approaches.

6.2 Decoupling of submodels through state estimation

Any set of decoupled parameter estimators can be written as

$$\beta^* = \mathcal{B}(Y|Z) \tag{23}$$

where Z comprises the boundary conditions between the respective processes.

An integrated parameter estimation is enabled if the boundary conditions are computed conditional on the parameters, i.e., if (23) is solved jointly with

$$Z = G(X) \tag{24}$$

$$X = F(Z|\beta^*). \tag{25}$$

The main difficulty of solving the integrated estimation problem is that the mathematical intractabilities of simulation-based components (such as a traffic flow microsimulation) also enter other, themselves well-behaved estimation problems (such as the maximization of a log-likelihood function for a behavioral model).

Essentially, the process interactions Z in (24) result from a plausible combination of structural model information and the data Y. Observing that a very similar problem is solved by the urban state estimator (12), one may approximate the calibrated process equation (25) by that estimator:

$$F(Z|\beta^*) \approx \mathcal{X}(Y|\beta^0) \tag{26}$$

where β^0 is an initial guess of the process parameters used during the state estimation. The advantage of this approximation is that, given sufficient data Y, the state estimation computes process interactions Z that are close to those interactions that would result from a simulation based on the estimated parameters. Hence, these interactions need no adjustment during the parameter estimation, which decouples the respective processes.

An operational implementation of this combined state and parameter estimation approach is outlined in Figure 3. The urban state estimator is deployed continuously based on given parameters. In regular intervals (e.g., monthly) these parameters are re-estimated based on all data collected so far. All process interactions up to the present point in time are approximately known from the state estimator. (Very old data may be discarded, which results in a rolling-horizon parameter estimator.) After the parameter estimation, the urban state estimation is further deployed based on the updated parameters. This approach iterates between parameter and state estimation, where the iterations take place along the time line. It can be expected to result in increasingly consistent parameter and state estimates as time progresses.

A complete decoupling of all model components may not be desirable. In particular, if the estimates during the early deployment of the system are far from their true values, some of the interactions should be accounted for, but in a mathematically tractable setting. This is the topic of the following subsection.

6.3 Response surfaces and metamodels

Response surfaces and metamodels were originally developed in the context of simulation-based optimization, e.g., Osorio (2010).

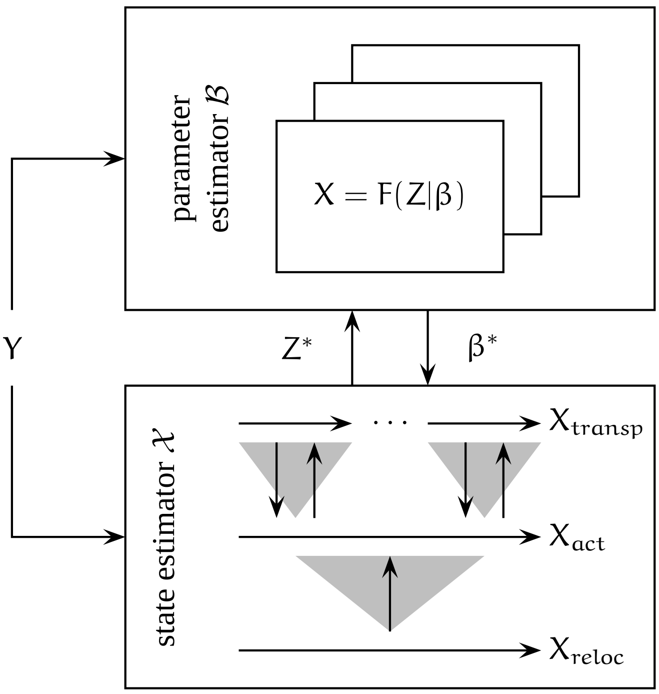

Figure 3: Integrated parameter and state estimation

Response surfaces. A response surface is typically a linear or quadratic polynomial that is fitted to the input/output signal of a complex simulation. Data availability is only limited by the computational effort to simulate it. However, a polynomial captures little to no structure about the simulation, and hence it may require a relatively large number of coefficients to reflect the simulation's relevant behavior.

Metamodels. Here, the polynomial is replaced by a mathematical model that structurally resembles the simulation. Metamodels can be expected to require less parameters than response surfaces for the same measurement fit because they contains more structural information. A drawback is that they are likely to loose the convenient linear-in-parameters form of polynomial models.

Both techniques can be used to approximately capture some interactions/processes in the integrated simulation when estimating its parameters. We exemplify this through a response surface and a metamodel for a vehicular traffic flow simulation. Assume that the mobility demand $Z_{mob} = (z_t^{od})$ consists of the number of vehicles z_t^{od} that want to travel between each OD pair in each time period t. The traffic flow simulation takes Z_{mob} as input and outputs the congestion information

$Z_{cong} = (z_{at})$ where z_{at} is the number of vehicles entering link a in time period t.

To obtain a response surface approximation, a matrix A is introduced that maps the OD flows Z_{mob} on *desired* link entrance flows $D = (d_{at}) = AZ_{mob}$ where d_{at} is the number of vehicles that plan to enter network link a during period t. A combines route choice information from the travel demand model with travel time information from the traffic flow model. A linear response surface $Z_{cong} = BD = BAZ_{mob}$ is then fitted to the traffic flow simulator, where the matrix B results from a regression as described above. An efficient approximation is to choose a diagonal B, which relates the flow across a link only to those vehicles that actually want to enter that link. The estimation of B from simulated data accounts for spillback in that some vehicle may be kept from entering their desired links in time because of congestion. This approach has been successfully applied to estimate disaggregate travel demand from traffic counts (Flötteröd and Bierlaire, 2009; Flötteröd et al., 2010).

The same problem can also be tackled based on the nonlinear metamodel described by Osorio and Bierlaire (2009a). This model goes structurally beyond a linear approximation in that it operates based on closed-form link state distributions and correctly accounts for spillback effects across network nodes. Its relative tractability is owed to the fact that it captures stationary conditions only, which may be a drawback in highly dynamic conditions. The flexibility of a response-surface approach and the structural power of a nonlinear metamodel can also be combined, which is demonstrated in Osorio and Bierlaire (2009b).

7 Summary

We presented a framework for the data-consistent deployment of urban microsimulations. In the first part of the article, we first adopted a process-oriented perspective on activity participation, transportation, and relocation and then refined this perspective in the microsimulation context. The second part of the article considered the parameter and state estimation problem. First, the different time scales of an urban system were identified and a rolling horizon framework for its continuous state estimation was developed. Second, the parameter estimation problem for an integrated urban microsimulation problem was investigated. The operational difficulty of jointly estimating all parameters of the urban model was met with two different approaches: the decoupling through estimated process interactions and the deployment of response surfaces and metamodels to mathematically approximate intractable, simulation-based processes.

References

Antoniou, C., Ben-Akiva, M. and Koutsopoulos, H. N. (2006). Dynamic traffic demand prediction using conventional and emerging data sources, *IEE Proceedings Intelligent Transport Systems* **153**(1): 97–104.

Arentze, T. and Timmermans, H. (2004). A learning-based transportation oriented simulation system, *Transportation Research Part B* **38**(7): 613–633.

Arulampalam, S., Maskell, S., Gordon, N. and Clapp, T. (2002). A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking, *IEEE Transactions on Signal Processing* **50**(2): 174–188.
URL: *citeseer.nj.nec.com/arulampalam02tutorial.html*

Balakrishna, R. (2006). *Off-line calibration of dynamic traffic assignment models*, PhD thesis, Massachusetts Institute of Technology.

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

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. and Mishalani, R. (1998). DynaMIT: a simulation-based system for traffic prediction, *Proceedings of the DACCORD Short Term Forecasting Workshop*, Delft, The Netherlands.

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

Bierlaire, M., Chen, J. and Newman, J. (2010). Using location observations to observe routing for choice models, *Proceedings of the 89. Annual Meeting of the Transportation Research Board*, Washington, DC, USA.

Bohte, W. and Maat, K. (2009). Deriving and validating trip purposes and travel models for multi-day GPS-based travel surveys: a large-scale application in the Netherlands, *Transportation Research Part C* **17**(3): 285–297.

Bowman, J. and Ben-Akiva, M. (1998). Activity based travel demand model systems, *in* P. Marcotte and S. Nguyen (eds), *Equilibrium and advanced transportation modelling*, Kluwer, pp. 27–46.

Chrobok, R., Pottmeier, A., Wahle, J. and Schreckenberg, M. (2003). Traffic forecast using a combination of on-line simulation and traffic data, *in* M. Fukui, Y. Sugiyama, M. Schreckenberg and D. Wolf (eds), *Traffic and Granular Flow '01*, Springer, pp. 345–350.

Chui, C. and Chen, G. (1999). *Kalman Filtering with Real-Time Applications*, Springer Series in Information Sciences, Springer.

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

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

Flötteröd, G., Bierlaire, M. and Nagel, K. (2010). Bayesian demand calibration for dynamic traffic simulations, *Technical Report TRANSP-OR 100105*, Ecole Polytechnique Fédérale de Lausanne & Berlin Institute of Technology.

Ghauche, A. (2010). Integrated transportation and energy activity-based model, Master thesis, Massachusetts Institute of Technology.

Greene, W. (2003). *Econometric Analysis*, Prentice Hall. 5th edition.

Hato, E. (2010). Development of behavioral context addressable loggers in the shell for travel-activity analysis, *Transportation Research Part C* **18**(1): 55–67.

Hurtubia, R., Flötteröd, G. and Bierlaire, M. (2009). Inferring the activities of smartphone users from context measurements using Bayesian inference and random utility models, *Proceedings of the European Transport Conference*, Noordwijkerhout, The Netherlands.

Miller, E. and Roorda, M. (2004). Prototype model of household activity-travel scheduling, *Transportation Research Record* **1831**: 114–121.

Nagel, K. and Axhausen, K. (2001). Microsimulation, *in* D. Hensher and J. King (eds), *The Leading Edge of Travel Behavior Research*, Pergamon, pp. 239–246.

Osorio, C. (2010). *Mitigating network congestion : analytical models, optimization methods and their applications*, PhD thesis, Ecole Polytechnique Fédérale de Lausanne. EPFL thesis 4615.

Osorio, C. and Bierlaire, M. (2009a). An analytic finite capacity queuing network model capturing the propagation of congestion and blocking, *European Journal of Operational Research* **196**(3): 996–1007.

Osorio, C. and Bierlaire, M. (2009b). A multi-model algorithm for the optimization of congested networks, *Proceedings of the European Transport Conference*, Noordwijkerhout, Netherlands.

Raj, A., Subramanya, A., Fox, D. and Bilmes, J. (2008). Rao-Blackwellized particle filers for recognizing activities and spatial context from wearable sensors, *in* B. Siciliano, O. Khatib and F. Groen (eds), *Experimental Robotics*, Springer, pp. 211–221.

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

Ross, S. (2006). *Simulation*, fourth edn, Elsevier.

Salvini, P. and Miller, E. (2005). ILUTE: an operational prototype of a comprehensive microsimulation model of urban systems, *Networks and Spatial Economics* **5**(2): 217–234.

Schüssler, N. and Axhausen, K. (2009). Processing raw data from global positioning systems without additional information, *Transportation Research Record* **2105**: 28–36.

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

Tampere, C. and Immers, L. (2007). An extended Kalman filter application for traffic state estimation using CTM with implicit mode switching and dynamic parameters, *Proceedings of the 10th IEEE Intelligent Transportation Systems Conference*, Seattle, USA, pp. 209–216.

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

Vaze, V., Antoniou, C., Wen, Y. and Ben-Akiva, M. (2009). Calibration of dynamic traffic assignment models with point-to-point traffic surveillance, *Transportation Research Record* **2090**: 1–9.

Waddell, P., Sevcikova, H., Socha, D., Miller, E. and Nagel, K. (2005). Opus: an open platform for urban simulation, *Proceedings of the Computers in Urban Planning and Urban Management Conference*, London, UK.

Waddell, P. and Ulfarsson, G. (2004). Introduction to urban simulation: design and development of operational models, *in* D. Hensher and K. Button (eds), *Transport Geography and Spatial Systems. Handbook 5 of the Handbook in Transport*, Pergamon/Elsevier Science, pp. 203–236.

Wang, Y. and Papageorgiou, M. (2005). Real-time freeway traffic state estimation based on extended Kalman filter: a general approach, *Transportation Research Part B* **39**(2): 141–167.

Wegener, M. (2004). Overview of land-use transport models, *in* D. Hensher and K. Button (eds), *Transport Geography and Spatial Systems. Handbook 5 of the Handbook in Transport*, Pergamon/Elsevier Science, pp. 127–146.

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

Zhou, X. and Mahmassani, H. (2007). A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework, *Transportation Research Part B* **41**(8): 823–840.