Issue 20(4), 2020, pp. 152-172
https://doi.org/10.18757/ejtir.2020.20.4.4463

TU Delft OPEN

EJTIR

ISSN: 1567-7141
http://ejtir.tudelft.nl/

Towards agent-based travel demand simulation across all mobility choices – the role of balancing preferences and constraints

Wolfgang Scherr[1]
Swiss Federal Railways (SBB), Switzerland.

Patrick Manser[2]
Swiss Federal Railways (SBB), Switzerland.

Chetan Joshi[3]
PTV America Inc., USA.

Nathalie Frischknecht[4]
Swiss Federal Railways (SBB), Switzerland.

Denis Métrailler[5]
Swiss Federal Railways (SBB), Switzerland.

This article presents an agent-based travel demand model, where agents react to transport supply across all mobility choices. Long-term choices include mobility tool ownership and work locations. Daily travel patterns are simulated at the individual level by sequentially combining activity frequency, activity durations and destinations as well as a rule-based time-of-day scheduling. A key to success in this novel approach is balancing individual preferences of travelers with system constraints. The model incorporates two types of constraints: 1) capacity constraints of the transport infrastructure. 2) natural time and space constraints during the execution of individual 24-hour day plans. Model results are validated against empirical observations of travel demand in Switzerland. The article concludes with a perspective for further research and development in the field of applied agent-based modeling.

Keywords: *travel forecasting, national model, public transport, decision support, microscopic model, agent-based model, activity-based model, SIMBA MOBi, MATSim.*

[1] A: Gardistrasse 2, 3000 Bern 65, Switzerland E: wolfgang.scherr@sbb.ch
[2] A: Gardistrasse 2, 3000 Bern 65, Switzerland E: patrick.manser@sbb.ch
[3] A: 1530 Wilson Boulevard, Suite 510, Arlington, VA 22209, USA E: chetan.joshi@ptvgroup.com
[4] A: Gardistrasse 2, 3000 Bern 65, Switzerland E: nathalie.frischknecht@sbb.ch
[5] A: Gardistrasse 2, 3000 Bern 65, Switzerland E: denis.metrailler@sbb.ch

1. Introduction

1.1 Motivation

Travel demand models play an essential role in transportation planning. They are applied to forecast traffic volumes and user benefits for future infrastructure projects and future service concepts. Since the beginning of computer-based travel demand models in the 1960s until today, the majority of transportation planning models uses a *macroscopic* (i.e. aggregated) approach. Mobility is simulated as a set of isolated trips that are neither connected in tours nor in activity schedules (Boyce and Williams, 2015). Some spatio-temporal constraints of travel are thus ignored. Today, the state of the art in travel demand modeling follows *microscopic* approaches, i.e. simulating each traveler as an autonomous decision-making unit.

This article presents a travel demand model that is *microscopic* and combines *activity-based* and *agent-based approaches*. The novelty of this methodology is that all mobility decisions of each traveler – including long-term and daily decisions – are modeled while ensuring that all of these individual decisions are sensitive to transport supply and to time-space constraints. This model property is crucial for the evaluation of infrastructure and service planning projects, where the complete reaction of travel demand to supply changes needs to be computed.

1.2 Previous related work

Two important approaches in microscopic travel demand modeling have advanced to real-world application, commonly referred to as *activity-based* and *agent-based* models. These approaches are discussed with recent examples in this section.

Activity-based models

The concept of activity-based transport models emerged and was put into practice in the 1990s. For a historical review see Bowman (2009) or Rasouli and Timmermans (2014).

In North America, microscopic activity-based models follow an econometric approach. A comprehensive presentation of the methodology is given by Castiglione et al. (2015). Important pioneers who managed to put the methodology into practice for several major U.S. cities in the 2000s are Bowman and Ben Akiva (2001), Vovsha et al. (2004) and Bhat et al. (2004). Individual day plans are built for each person or household, by a set of discrete-choice models for mode ownership, tour generation, activity selection, mode choice, location and destination choice.

Another activity-based approach by Roorda et al. (2008) adds rules to the individual activity and travel decisions. This approach allows verification of feasibility of individual patterns and builds consistent 24-hour day plans for each traveler.

While the common practice of the activity-based models presented above simulate 24 hours of a day, mobiTopp (Mallig et al., 2013) simulates an entire week, where long-term and short-term activity needs of all persons in a household are scheduled and simulated over seven days.

The strength of these activity-based approaches is that individual travel decisions are sensitive to travel time and cost in many mobility choices. A shortcoming of most activity-based models in practice is that the microscopic approach is given up once travel demand is established, and trips are aggregated to origin-destination (OD) matrices that are fed into aggregated network assignment models to simulate route choice and network flow. To overcome this shortcoming, some activity-based models have been extended with dynamic highway assignment, for example by Vovsha et al. (2016), but up to this day the activity-based concept is not extended into network simulation for the entire demand with all modes of travel.

In recent years, several researchers have developed methods to optimize activity-based day plans with the aim to provide mobility planning tools to individual travelers (Arentze, 2013; Sierpinski, 2016; Hilgert et al., 2016; Estergár-Kiss, 2017). While this work does not aim for system-wide traffic forecasting, there is a potential for synergy with activity-based or agent-based travel demand modeling.

Agent-based models

Within the microscopic approaches, *agent-based simulation* treats also vehicles as autonomous units (in addition to traveling persons) and keeps a memory for each unit (i.e. for each agent). Using this memory, agents learn and update their travel patterns iteratively. Agent-based simulation aims at a user equilibrium of all agents across the entire transport system.

Agent-based models emerged in the 2000s with a focus on large-scale microscopic traffic simulation. A first approach was TRANSIMS (Cetin et al., 2002). Today, the open source software MATSim (Horni et al., 2016) is used in many academic applications. MATSim connects supply and demand in a network equilibrium, where individual travelers (agents) start with pre-defined day plans, search routes through the networks and adapt choices of travel mode and time of day. An important feature of MATSim is its integrated end-to-end simulation of persons and vehicles. Vehicle flow is modeled with a queue-based traffic flow model. Compared to activity-based models, the microscopic approach is more detailed as each agent makes decisions based on individual travel conditions which are computed with high resolution in both time and space. MATSim automatically enforces natural arrival/departure time constraints to 24-hour day plans. As a result, a person cannot depart for travel from an activity location before having arrived and dwelled for the required activity time. In the standard software version of MATSim, only the following mobility decisions can be simulated: mode choice, time-of-day choice and route choice.

To include other mobility decisions in the simulation, several researchers have developed extensions of MATSim for the destination choice of secondary activities, (Horni et al., 2011; Ordonez, 2016; Hörl et al., 2019). Still, neither long-term mobility decisions (mobility tool ownership and primary activity locations), nor activity choices have been included. Hence in all MATSim variations we know, the day plans of agents remain largely fixed and do not react to changes in travel supply. Another interesting extension of MATSim has been developed by Hörl et al. (2018), by replacing the random mode choice by a discrete choice model. For each tour or subtour the mode is chosen by a multinomial LOGIT, as input for the agent-based traffic flow simulation.

In Switzerland, there is more than a decade of experience with agent-based modeling using MATSim. Since Meister et al. (2008), a model at national scale had been developed. This model has gone through several updates and extensions (e.g. Bösch et al., 2016) and has been applied over a multi-year research program to advance agent-based modeling to address various research questions. In this model, generation of day plans and long-term location choices are not responsive to transport supply. The work presented in this article was inspired by this model. This model also provided a strong foundation for the starting point of this work, since a lot of technical and administrative ground had been broken for agent-based modeling in Switzerland during the development of this model.

Combination of the activity-based and agent-based approaches

In recent years, several researchers have combined activity-based modeling with agent-based simulation in MATSim, obtaining synergy from both approaches.

Ziemke at al. (2015) compute a set of day plans for each agent using the CEMDAP activity-based model, subsequently the agents choose between these different plans in MATSim, which also performs traffic flow simulation. While this approach presents a complete integration of activity and agent-based modeling through all model stages, this experiment does not include a complete calibration, as CEMDAP parameters for Los Angeles are combined with MATSim parameters for Berlin, Germany.

Moeckel et al. (2019) developed the MITO model for Munich, Germany, combining MATSim with an activity-based demand model which simplifies the construction of activity schedules. It explicitly uses travel time budgets as constraints of destination choice.

Briem et al. (2019) combine the activity-based model mobiTopp with MATSim. They limit the use of MATSim to traffic flow simulation, without using the MATSim features of mode and time choice adaptation, but agent-based learning is successfully combined with a genuine activity-based travel demand model.

1.3 Objectives and contributions

The model effort presented in this article is agent-based and applied at a large scale. It simulates 24 hours of a weekday and covers the entire population of Switzerland. It uses a combination of two major approaches mentioned above: First, activity-based modeling with discrete choice models (day pattern approach) is used to generate initial day plans at the individual level. Next, network-wide agent-based traffic simulation with MATSim is performed. This allows the agents to learn from their individual travel experience and to adjust mode and time of travel accordingly. To integrate the two approaches, a new methodology is presented to build 24-hour day plans for each agent. These day plans are consistent in time and space. Time consistency is imposed using time budgets. A novel plan-building heuristic is developed for this purpose. The objective of this methodology is that all elements of an agent’s plan are responsive to changes in network level of service and to constraints in the transport system. Consequently, the overall agent-based simulation obeys more realistic constraints to inform real-world transportation planning decisions.

2. SIMBA MOBi: model architecture and methodology

2.1 The context of travel modeling at SBB

Travel demand modeling at SBB (Swiss Federal Railways) is aimed at supporting management decisions about future service concepts and investments in infrastructure and rolling stock (Scherr et al., 2018; Scherr et al., 2019). For over 15 years, the SBB passenger division has used a macroscopic rail-only model. SBB has developed the agent-based model “SIMBA MOBi” – complementary to this existing macroscopic model over the past three years (2016 to 2019). SIMBA MOBi is presented in this article. It has been developed to analyze future mobility schemes, including disruptive changes, new technologies and new intermodal services. The model is sensitive to changes in transport supply across all modes and can thus be reliably applied in corporate decision-making. SIMBA MOBi applies standard software tools (MATSim, Biogeme, PTV Visum), but many software extensions and workflow tools have been programmed in Python and Java, where standard solutions where not available.

2.2 General model architecture

SIMBA MOBi has three major behavioral modules (see Figure 1). The first module, MOBi.synpop generates a resident population for the existing state and allows modeling future synthetic populations. The synthetic population, including households, persons, businesses and other institutions is the starting point of the demand model. The methodology is well described in Bodenmann et al. (2019) and hence not discussed in more detail. The focus of this article are the other two modules.

The second module MOBi.plans constructs 24-hour activity and travel plans for each agent in the synthetic population during the following choice steps:

- *Long-term decisions* (Section 2.3): Mobility tool ownership and primary locations.
- *Daily mobility preferences* (2.4): Number and type of activities a person wishes to perform during a day, and the sequence of tours and the pattern of the activities within each tour. Also, destinations of the secondary activities, preliminary modes for each tour and durations for each activity.
- *Rule-based plan-building* (2.5): Confrontation of the daily preferences with time- and space-constraints based on time budgets and building of a time- and space-consistent 24-hour day plan for each agent.

The rule-based plan-building step is the key to success in the overall model structure as it brings together the individual preferences of the activity-based methodology with the agent-based methodology which requires time- and space-consistent agent plans.

The third module is MOBi.sim, which simulates 24-hour dynamic network flows of cars and public transport using the agent-based software MATSim (2.6). In MOBi.sim, the agents choose their final modes, route and departure times for each trip based on traffic condition and a full-day public transport schedule.

A feedback loop from MOBi.sim back to MOBi.plans then informs agents' decisions about travel conditions and service quality experienced across the transport network. This makes – for example – destinations in heavily congested areas less attractive for car-oriented persons.

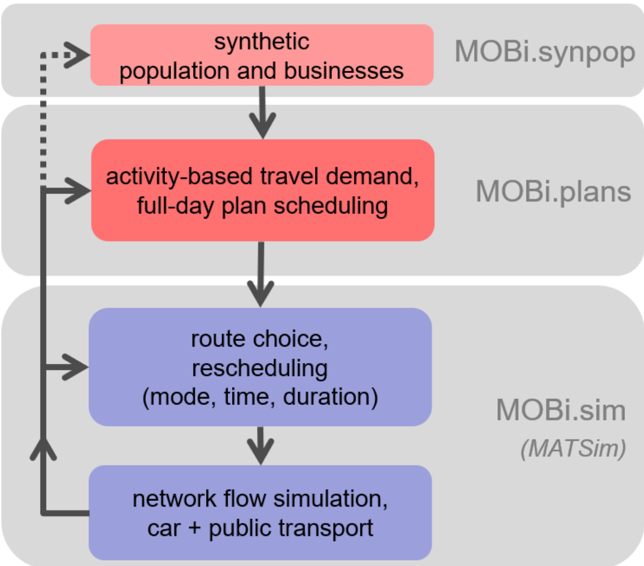

Figure 1. The main behavioral modules of SIMBA MOBi

SIMBA MOBi is a person-based model. Interaction between household members is not modeled explicitly. However, household properties, such as size, presence of children or car ownership, are included into the person-based decision models. Hence, linkages between household members are modeled implicitly.

2.3 Long-term mobility preferences in MOBi.plans

Car availability and public transport subscription

A first model step determines car availability and public transport subscriptions (i.e. yearly or monthly public transport passes) for each person in the population. Both are important person attributes and have a strong impact on mode choice and destination choice. In the existing case, these attributes are determined by choice models and then corrected based on small-area control variables. At the time of the publication of this article, the model by Danalet et al. (2018) is applied. It will soon be replaced by a more advanced model by Hillel et al. (2020), which will include accessibility as an explanatory variable of mode and subscription ownership making these choices responsive to transport supply and capacity constraints in the transport network.

Location choice

Location choice is defined as a choice of permanent or long-term locations of activities such as work or school of each individual agent. This decision is mainly constrained by the home location, person specific preferences and transport supply such as travel times, parking costs, and public transport schedule (Table 2). In a first step, location choice probabilities are computed at an aggregate level, i.e. on traffic analysis zones. SIMBA MOBi uses 8'000 zones defined by the Federal Office for Spatial Development (2020), averaging 1'000 inhabitants per zone. Zone-based location choice has the advantages of simplicity and of allowing for the use of established calibration methods. It also obviates the need for sampling of alternatives. In a second step, each individual agent chooses a specific zone based on the aggregated probabilities. In a third step, a specific facility is selected for the corresponding activity type in the chosen zone based on a weighted random draw (with the weight being the attraction of the facility, e.g. the number of jobs).

Sensitivity to changes in the transport supply is a crucial requirement of the model. Hence, a nested model structure informs the location choice about the resulting "logsum" (or maximum expected utility) of the trips-based mode choice, capturing effects of the network supply. One major challenge is the replication of the nonlinear decrease with travel distance with the exponential form of the LOGIT model. As done in earlier studies, a piecewise-linear form of the distance term is used to capture this nonlinear effect more accurately. The LOGIT formula of location choice is identical to the one used in destination choice (Section 2.4). Additional additive utility terms (shadow prices) in the higher level of the location choice explain preferences, which cannot be explained by physical attributes of travel nor by the transport supply. In Switzerland, a classic example is the language barrier which presents a substantial barrier in traveling as well. Lastly, a shadow price at the target allows considering capacity constraint (e.g. number of jobs in a zone).

2.4 Daily mobility preferences in MOBi.plans

The methodology to determine daily mobility choices follows the North American approach of activity-based models as described by Castiglione et al. (2015).

Tour and stop frequency - activity choices

The aim of tour- and stop frequency generation in combination with activity choice is to determine how many and which activities an agent performs during the day, as well as how those activities are combined in tours. A tour is a sequence of activities and trips that begin at home and end at home. A stop is an intermediate activity, that is performed during a tour.

Table 1 gives an overview of the tour- and stop frequency models used in MOBi.plans. Tour and stop frequency choices are organized as a sequence of multinomial LOGIT models (MNL). The dependent variable in the LOGIT model is the number of tours, subtours or stops, respectively. The endogenous variables are socio-economic variables plus spatial and accessibility measures. The MNL coefficients were estimated using Biogeme (Bierlaire, 2016), based on data of the national travel diary survey (Federal Statistical Office, 2017). The most important property of the generation models is their ability to forecast changes in the mobility of individuals based on mode availability, changes in transport supply (by means of accessibility) and demographic shifts (by means of the age variable).

Four different types of tours are modeled, the first two being considered primary tours (main tour purpose is a primary activity) and the latter two secondary tours (tour consists of secondary activities only):

1. Work tour (primary tour)
2. Education tour (primary tour)
3. Business tour (secondary tour)
4. Other tour (secondary tour)

Table 1. Overview: Multinomial LOGIT models of Tour- and Stop Frequency

	Tour frequency				**Subtour frequency**	**Stop frequency**		
	No. of prim. Tours		No. of sec. tours		Has subtour	No. of stops prim. tour		No. of stops sec. tour
	Work	Educ.	Bus.	Other		Outbound	Inbound	
Constant	X	X	X	X	X	X	X	X
Employment level	X	X	X	X	X	X	X	X
Main occ. is in education		X		X	X			
Age	X	X	X	X	X	X	X	X
Is in management			X					
Children in HH	X			X	X	X	X	X
Car available	X		X	X	X	X	X	X
Public transp. Subscript.	X	X	X	X	X	X	X	X
Dist. to primary location	X	X	X			X	X	
No. of tours					X	X	X	X
No. of primary tours				X				
Is a work tour					X	X	X	
Is a business tour								X
Accessibility home loc.	X	X		X		X	X	X
Accessibility primary loc.			X		X			

The stop frequency choice model determines the number of activity stops made within a tour, in addition to the primary activity. For primary tours, stop frequency is segmented into:

1. An outbound stop model, which predicts the number of stops made between leaving home and the first primary activity (work or education).
2. A subtour frequency model, that predicts whether there is a primary location-based subtour or not. A subtour is a sequence of trips that begins and ends at the primary location without going home. An example of a subtour is "work – leisure – work".
3. An inbound stop model, that predicts the number of stops made on the way back home from the primary location.

For secondary tours ("other tours"), there is only a single stop frequency model that predicts the total number of stops.

Once the number of tours and their stops are calculated, the type of activity is added by applying probabilities segmented by person groups from travel diary observations. The following activity types are considered:

- Leisure (L)
- Shopping (S)

- Business (B)
- Education as secondary activity (EC)
- Accompany (A)
- Other (O)

After having made all the choices about number of tours as well as the number and type of stops during one tour, every person has a desired activity pattern containing a set of activities as well as their order within the tours. In the following steps, the activity patterns are assigned destinations, modes and time of day. Also, the order of the tours is defined when the day plan is built for each agent.

Destination and mode choice
Knowing the number and type of activities each agent performs in each tour, the next step is choosing the destination of each secondary activity and the mode for each trip between activities. For this purpose, destination probability matrices for each activity type and different person groups are estimated following the same method as in the location choice (Section 2.3).
The probability of choosing mode m on OD pair ij is:

$$P(m|ij) = \frac{exp(V_{ijm})}{\sum_k exp(V_{ijk})} \tag{1}$$

Where V_{ijm} is the utility of mode m on OD-pair ij

The "logsum" or expected maximal utility (*EMU*) of mode choice is then:

$$EMU_{ij} = ln\left\{\sum_m \left[exp(V_{ijm}/\theta)\right]\right\} \tag{2}$$

Mode choice depends on various variables such as travel time, distance and other level of service measures (Table 2). To inform destination choice about the level of service of all modes, mode choice is nested into destination, by including EMU_{ij} (the expected maximal utility of mode choice) into the destination choice utility:

$$V(j|i) = ln(A_j) + \theta \cdot EMU_{ij} + \lambda_j + \lambda_{ij} \tag{3}$$

Where:

- A_j: the socio-economic attraction of zone j
- λ_j: shadow price of destination j
- λ_{ij}: shadow price of OD-pair ij

In contrast to home-based location choice, choosing destinations for each intermediate stop is not as straight forward. Since an intermediate tour stop lies between two pre-defined locations, "*rubber banding*" is used to consider both trip origin i and primary location k (work or school place) in the choice of intermediate destinations. This is done with the aim of minimizing out of way travel. The rubber banding formula uses weights to balance the influence of primary locations and secondary destinations:

$$V(j|i) = ln(A_j) + \alpha \cdot (\theta \cdot EMU_{ij} + \lambda_j + \lambda_{ij}) + \beta \cdot (\theta \cdot EMU_{jk} + \lambda_j + \lambda_{jk}) \quad (4)$$

Where:

- α: weight of the trip origin
- β: weight of the trip destination

Finally, the probability of choosing destination j under the condition of starting in origin i is:

$$P(j|i) = \frac{exp(V(j|i))}{\sum_k [exp(V(k|i))]} \quad (5)$$

This approach does not include the interaction between the destination choice decisions of multiple tours. This might result in unrealistically high travel time which violates time budgets and hence plan integrity. This issue will be faced in the plan-building step (Section 2.5).

Table 2. Level of service measures used in mode choice utility

Mode	const.	travel time	parking search time	service frequency	no. of transfers	parking cost	distance
walk	X	X					
bicycle	X	X					
public transport	X	X	X	X	X		X
car - driver	X	X	X			X	X
car - passenger	X	X	X			X	X

The mode choice parameters calibrated for the nested mode and destination choice are trip-based. Parameters are estimated for the variables as depicted in Table 2. In the mode choice step, the constraints of mode used along the whole tour are considered. A mode is assigned to each tour and subtour based on the zonal level of service measures in MOBi.plans. The mode choice later informs the plan-building step (2.5) about the expected travel times. It is important to note that the tour-based mode calculated in this step is the starting point for the agent-based traffic flow simulation (2.6). In MOBi.sim, the agents can adjust their tour-based modes depending on the individual travel conditions.

Desired activity durations

In a similar approach to the one taken by Hörl (2017), the activity durations are determined with probability distributions derived from the national travel diary survey (Federal Statistical Office, 2017) as shown in Figure 2. The choice is then a weighted random draw for each activity based on the distribution. The distributions distinguish between multiple demand segments that are defined by type of activity, by socio-economic attributes of the person, and by the frequency of the activity in one plan (e.g. the workplace is visited once or twice). Figure 2 shows the distributions for a selection of market segments. Full-time employees with one work tour have the longest activity durations at the workplace. The duration is shorter for agents working part-time as well as for agents going to work twice per day.

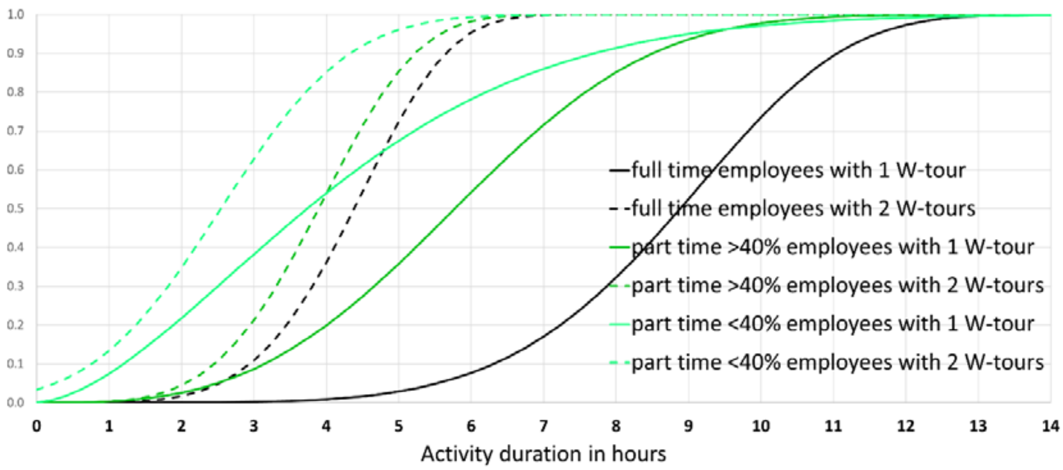

Figure 2. Probability distributions of desired activity duration (e.g. work)

2.5 Rule-based plan-building in MOBi.plans

The plan-building step is a novel method developed for SIMBA MOBi. It is the crucial link between the activity-based demand model and the agent-based network simulation. Up to this step, each agent has determined preferences including all daily choices. These preferences are now transformed into a full day plan, which is consistent with natural time and space constraints: All activities and travel must be performed within 24 hours in an order where none of the episodes must overlap. The plan-building procedure is subdivided into three steps:

1. Generation of alternative preferences for destinations and activity durations.
2. A random combination of secondary activity destinations and activity durations is chosen while considering time budgets. If no satisfying combination has been found, agents redo all daily choices to be consistent with time budgets.
3. Finally, the starting time for each activity is chosen. In this final step, travel times and activity duration are no longer changed.

Generation of destination and activity duration alternatives

The plan-building procedure tries to find a combination of destinations and activity duration based on time budgets. Given a fixed set of activities A for each individual agent, alternatives for destinations and durations are generated. A contains information about number and type of activities each agent has chosen based on the daily preferences. We call the set of destination alternatives D and the set of duration alternatives P. Each element $d \in D$ represents an array containing information about the destinations of each activity $a \in A$. Each element $p \in P$ represents an array containing information about the duration of each activity $a \in A$.

Choosing destination and duration alternatives based on time budgets

This step targets to find a reasonable combination of destinations and activity durations based on travel time (bud_{travel}), activity performing (bud_{perf}) and total out-of-home ($bud_{outofhome}$) budgets.

The algorithm is defined as follows:

1. Calculate the sum over all activity durations for each array of durations $p \in P$. Each element in P_{perf} is a duration-related number and $|P| = |P_{perf}|$:

$$P_{perf} = \forall p \in P, \sum_{activity \in p} duration_{activity} \tag{6}$$

2. Reduce set P based on bud_{perf}. Only alternatives are retained which fulfill the budget constraint. Otherwise, the element with the minimal total duration is retained in P_{red}:

$$P_{red} = \begin{cases} \{dur \mid dur \in P_{perf}, dur < bud_{perf}\}, & \min(P_{perf}) < bud_{perf} \\ \{\min(P_{perf})\}, & otherwise \end{cases} \tag{7}$$

3. Calculate the sum over all trip travel times towards each destination in the array of destinations $d \in D$. Each element in D_{travel} is a travel-time-related number and $|D| = |D_{travel}|$:

$$D_{travel} = \forall d \in D, \sum_{destination \in d} traveltime_{trip\_to\_destination} \tag{8}$$

4. Reduce set D based on bud_{travel}:

$$D_{red} = \begin{cases} \{trav \mid trav \in D_{travel}, trav < bud_{travel}\}, & \min(D_{travel}) < bud_{travel} \\ \{\min(D_{travel})\}, & otherwise \end{cases} \tag{9}$$

5. Repeatedly (max. N times) try a combination of a random destination plan $d_{rand} \in D_{red}$ and a random duration plan $p_{rand} \in P_{red}$ until out-of-home time is lower than $bud_{outofhome}$. Out-of-home time is defined as:

$$t_{outofhome}(d_{rand}, p_{rand}) = tot\_travel\_time_{d_{rand}} + tot\_act\_duration_{p_{rand}} \tag{10}$$

   Note that this is not minimization problem. Using pure random combinations, original preferences are maintained as much as possible

6. The final plan is the combination of d_{rand} and p_{rand} if the total out-of-home time of the chosen plan is lower than $bud_{outofhome}$ after max. N repetitions. Otherwise, the agents must repeat their full daily mobility preferences. That means these agents must go back to the tour generation step and redo all their daily choices while not changing their preferences. Long-term locations are not adjusted.

In SIMBA MOBi, we use $|D| = 3$ and $|P| = 10$. It is a trade-off between gaining information and increasing computational cost for an additional alternative in one of the sets. In this version of the model, time budgets are the same for all agents but depend on how many times the agents went through the loop of choosing daily mobility preferences and trying to build a plan within the defined budget. The following parameters are derived from calibration and resulted in the best model quality:

Table 3. Time budgets

Budget [h]	**Iteration 1**	**Iteration 2**	**Iteration 3**	**Iteration 4**
bud_{travel}	12.0	5.0	4.0	3.0
bud_{perf}	14.0	12.0	11.0	10.0
$bud_{outofhome}$	13.5	14.0	15.0	16.5

Using the time budgets as given in Table 3, over 99% of all agents found a valid plan after their choosing their daily preferences for the first time ($t_{outofhome} < bud_{outofhome}$). The <1% remaining agents will repeat the choice of their daily preferences. After iteration 4, only few agents (<0.1 ‰) have not found a valid plan considering the defined time budgets.

Activity start-time choice

Input for the activity start-time choice are activity chains with complete information about activity and trip durations. Travel times between activities are looked up for each individual trip depending on its mode. Hence, each tour has a fixed duration at this stage. Another input are distributions representing activity start-time preferences. A selection of these distributions for some market segments is shown in Figure 3. Work and education show a strong morning peak while secondary activities like shopping and leisure are likely to start later in the morning without significant peaks.

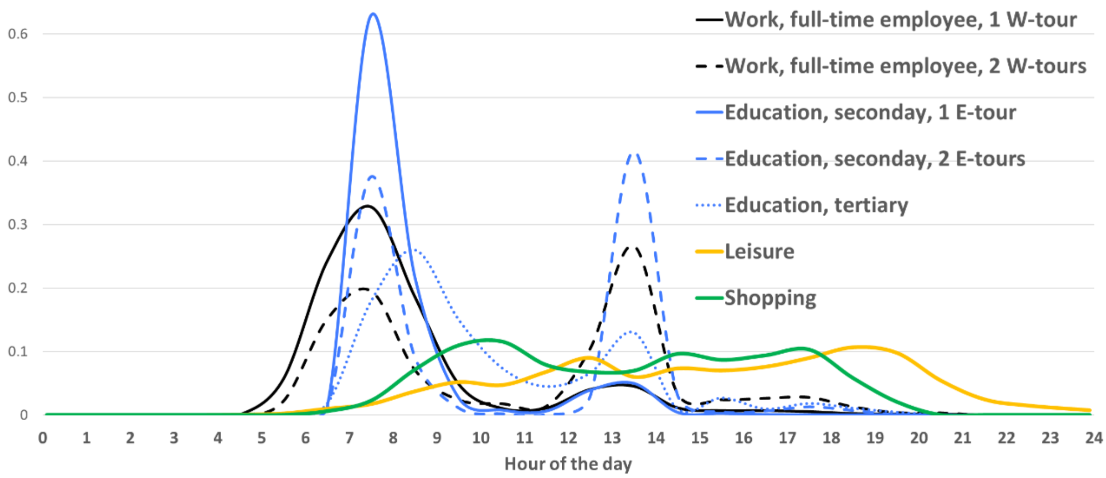

Figure 3. Probability distributions of activity start-times for selected activities

With these inputs given, the scheduling procedure then follows the principle of the "outward" approach described in Castiglione et al. (2015), with the rationale that priority is given to the primary activities and their durations, and that secondary activities and travel episodes are added before and/or after the primary activities. As the start-times for secondary activities are only chosen implicitly, we implemented an iterative algorithm that scores the start-times of the secondary activities. The score is based on the probability curves as shown in Figure 3. In the case of a badly scored choice, the agents make an alternative start-time choice for their primary activities. If they achieve a higher score, they might then take the new plan.

The above algorithm discards tours from the schedule if they do not fit into the 24 hours of the day after a certain number of iterations. This is the last resort to assure integrity of the plans. But these cases are extremely rare (0.06‰ of all trips), the step of plan-building is effective in assuring that the timing constraints are met by the scheduling procedure. At the end of the MOBi.plans procedure, each agent has a synthetic day plan that meets the strong plan integrity requirement of MATSim. This strong integrity distinguishes the model presented in this article from most other activity-based models, which feed only individual trips into aggregated assignment models and often ignore this requirement.

2.6 *Agent-based network simulation in MOBi.sim*

The aim of MOBi.sim is to connect supply and demand in an agent-based network simulation using the software MATSim (Horni et al., 2016). Starting with pre-defined day plans (as generated by MOBi.plans), individual travelers (agents) iteratively update routes through the network and adapt choices of travel mode and time of day. These decisions are based on individual travel conditions during an iteration which are computed with high resolution in both time and space (fully disaggregate multimodal assignment). The decisions are compared to a memory of several plans from previous iterations. While discarding plans with bad decisions and keeping plans with good scores, the simulation converges to a network equilibrium.

The following strategies are used in MOBi.sim to update the plans during the simulation:

- Subtour-based mode choice: Random choice of a mode for each subtour in the plan. Includes all five modes of travel that are defined in MOBi (Table 2).
- Activity start-times: Choice of new start- and end times for each activity. This choice is limited to a range of ±30min compared to the original decision made in MOBi.plans.
- Route choice: Choice of a route in the network depending on the network condition. By updating their routes, agents can bypass heavily congested areas.

MOBi.sim was built based on the experience of earlier MATSim models of Switzerland (Meister et a., 2008), then calibrated in-house by SBB (Scherr et al., 2018) and improved with several new software features:

- Routing and simulation algorithms for public transport (Rieser et al., 2018).
- A parking cost and access time model for cars, which is included in the algorithms of car routing as well as in traffic flow simulation.
- A mode and route choice model ("scoring model" in MATSim terms) that is differentiated by multiple socio-economic groups.

To obtain comprehensive traffic on the networks, we added exogenous demand to the network simulation in MOBi.sim as MOBi.plans produces the travel demand of the resident population of Switzerland only. The exogenous demand includes international rail travel, border crossing road traffic, airport travel by non-residents, travel by tourists and visitors (both road and rail). The exogenous demand is derived from either internal SBB data sources (e.g. rail booking systems), from other travel models, and from specific surveys, e.g. the national border-crossing OD-survey. The exogenous travel demand consists of single trips only. MATSim allows for inclusion of such isolated trips to the traffic simulation.

From the agent-based traffic flow model we also derive level of service indicators for all modes. These indicators are computed from/to discrete geo-codes, then aggregated as OD matrices from/to zones for MOBi.plans. This feedback loop from MOBi.sim to MOBi.plans is important as it informs the agents about congestion and guarantees that capacity constraints are considered by all agents in the choices taken in the MOBi.plans module.

2.7 *Calibration methodology*

This model includes a great number of behavioral parameters. Our approach to determine these parameters can be summarized as follows:

- Statistical estimation was used whenever possible. Especially in MOBi.plans, which consists mainly of multinomial LOGIT models (MNL), the betas were either estimated by us or in other cases – for instance mode choice – existing parameters estimated for Switzerland were used (Weis et al., 2016).
- Shadow prices for the destination choice models were determined using iterative fitting methods that are common practice in aggregated travel demand modeling.
- In the case of MATSim, no estimation routines exist for the utility functions. Here, we used betas which originally were estimated for MNL models as the basis for calibration. These betas were then manually adjusted during calibration and differentiated by groups of agents. All manual adjustments were validated by verifying substitution rates, general reasonableness, and by watching the resulting elasticities in sensitivity tests.

3. Model validation

The model was validated in comparison to comprehensive travel statistics that include the national travel diary survey (Federal Statistical Office, 2017), commuter matrices, the national rail OD-survey, rail counts, other public transportation counts, road traffic counts, and SBB corporate data. Significant effort went into comparing, understanding and preparing all available empirical data sources. This section shows a small selection of validation statistics and different model quality measurements of the SIMBA MOBi 2.0 release. Much more validation results are available for this model but would exceed the limits of this publication.

3.1 System-wide statistics of travel demand

Table 4 shows system-wide validation of travel demand. The mean number of trips generated (3.76 trips) is slightly lower than the empirical number reported in the national travel diary survey (3.87 trips). This small difference is caused by differences in the socio-economic composition between the synthetic population and the survey sample. Time budgets play an important role in the plan-building (Section 2.5) the get consistent travel times and activity duration times, resulting in the total out-of-home time. Table 4 shows that our rule-based approach works reasonably with the mean out-of-home time being slightly shorter than reported in the survey. It should be noted that we intentionally calibrated the durations of some secondary activity types shorter to ease the constraints for the plan-building procedure. The mean trip distances (overall, per activity type, per mode) were calibrated against survey data in combination with count data. In the case of public transport, we relied more on count data and hence allowed an error to the travel diary survey.

Table 4. Key statistics of travel demand

	SIMBA MOBi	Travel diary survey
Number of persons (observations)	7’979’430	39’075
Mean number of tours	1.49	1.56
Mean number of trips	3.76	3.87
Mean number of work trips	0.51	0.53
Mean number of leisure trips	0.72	0.74
Mean travel time [h]	1.52	1.59
Mean activity duration time [h]	5.92	6.38
Mean out-of-home time [h]	7.44	7.97
Mean trip distance [km]	36.63	38.65
Mean work trip distance [km]	12.33	12.21
Mean leisure trip distance [km]	6.79	8.50
Share of public transport PKM[1] [%]	29.8	31.4
Share of car PKM[1] [%]	51.8	51.3

1) person distance (km) traveled

3.2 Disaggregated mode choice results

In addition to aggregated statistics, many disaggregated statistics of travel demand were compared as well. This was done by time of day, socio-economic groups, land used type, sub-regions and individual cities.

The following Figure 4 shows mode shares depending on person groups according to their different public transport subscriptions. The mode shares shown are based on the total distance traveled. The results show that travelers without subscription show a very different mode choice behavior than persons who are subscribed. The most significant behavior is shown by persons owning a “General-Abo”, a yearly pass that entitles to use unlimited public transport nationwide, including rail and local services (bus, tram, boat, etc.). This group travels 80% of their daily distance with public transport. "Verbund-Abo" means owning a regional pass, "Halbtax-Abo" means owning a half-price pass for public transport. It is an advantage of microscopic modeling that these different market segments can be distinguished in the evaluation, but more importantly, their specific behavior can be replicated by the model. It should be noted that it is not sufficient to have specific mode choice parameters for each group, but that also their home locations need to be realistic (subscription owners tend to live close to railway stations) and finally, destination choice needs to be calibrated such that subscription owners prefer destinations with high public transport accessibility.

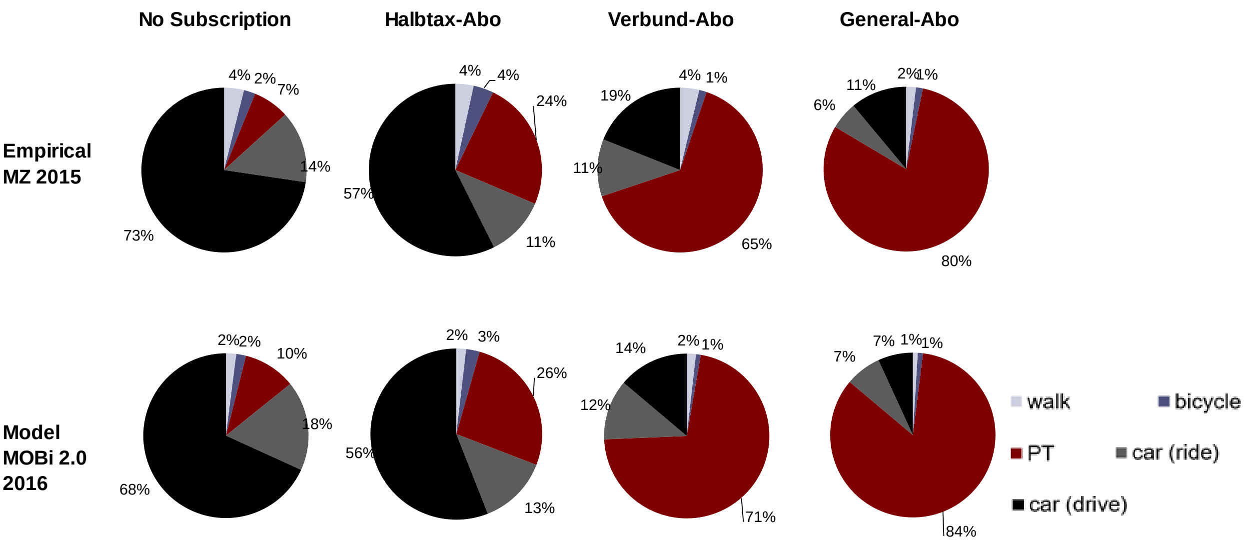

Figure 4. *Mode shares, different types of PT subscriptions (in person km traveled)*

3.3 Destination and location choice

As described in Section 2.3, location choice depends on network conditions (e.g. travel time), person specific attributes and the destination's attraction (e.g. number of jobs in a zone). Figure 5 shows the work commuter flows in SIMBA MOBi compared to empirical values as provided by the Federal Statistical Office.

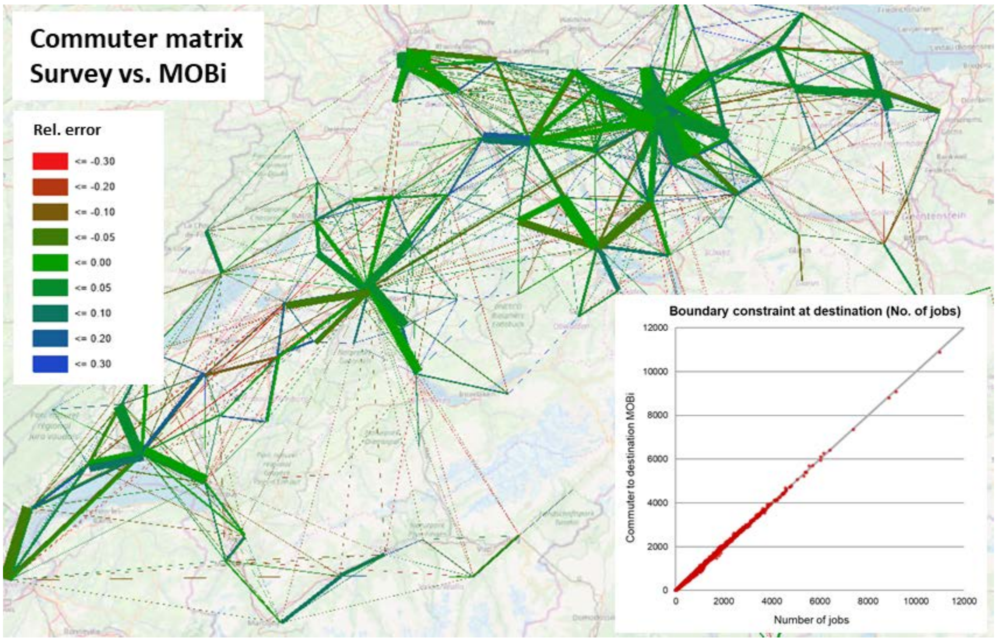

Figure 5. *Commuter flows in SIMBA MOBi*

To obtain a good calibration of location choice, shadow prices, both for OD pairs as well as for destinations, were added to the utility functions as described in Section 2.4. OD specific shadow prices help to match empirical commuter flows, and to represent the polycentric structure of Switzerland with regional characteristics such as locally spoken languages. Destination specific shadow-prices allow to meet boundary constraints (e.g. number of jobs at the destination). For secondary destinations (shopping, leisure, etc.), no empirical OD-data was available. The national travel diary survey (Federal Statistical Office, 2017) was used to calibrate trip length distributions for all secondary activities.

3.4 Public transport network loads and quality assessment

Figure 6 shows rail passenger volumes compared to counts. In most cases, the error is within a range of ± 5%. The main corridors of high demand match the empirical data very well. Calibration problems remain in touristic areas of the Alps. Even though the model includes touristic demand (as described in Section 2.6), we assume that non-resident demand is underestimated in the model, especially in cases where route choice is less dependent on travel time than on the natural beauty that can be watched while riding a train.

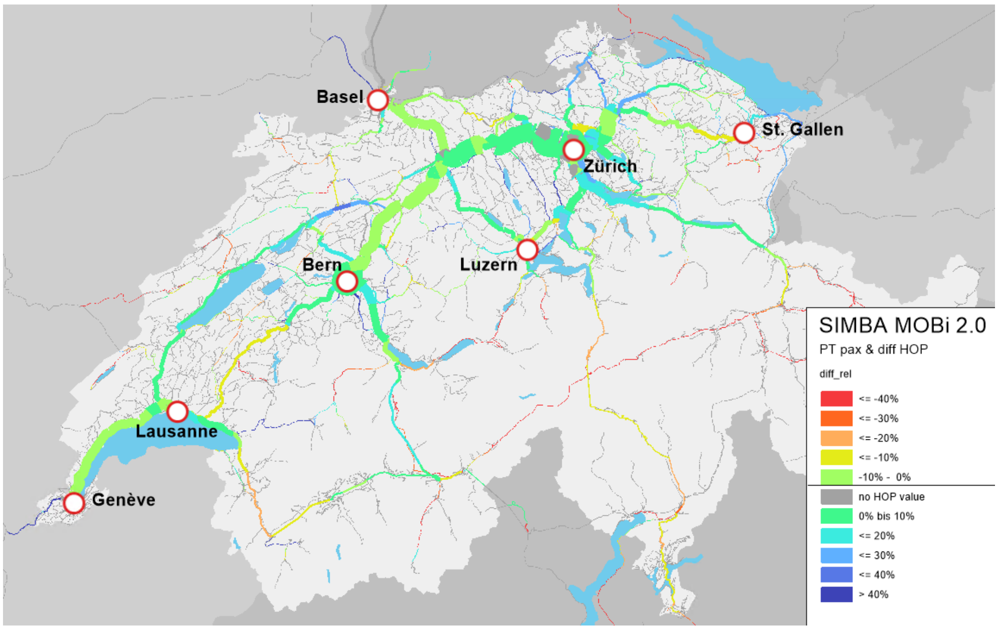

Figure 6. Rail passenger volumes – model versus counts ("HOP" represents the counts)

To assess the model quality for calibration criteria consisting of numerous count data, traditional error measures like *%mean absolute error* (%MAE) and *%root mean square error* (%RMSE) were used. Additionally, *scalable quality value* (short: SQV, see Friedrich et al., 2019) was also used. The SQV ranges between 0 (no match) and 1 (perfect match). It is expressed as:

$$g_{SQV} = \frac{1}{1 + \sqrt{\frac{(m-c)^2}{f \cdot c}}} \tag{11}$$

Where m is the model value and c the counted value. f is a scale factor and depends on the type of the counts. The calibration objective was to maximize the number of counts with a good SQV. Table 5 shows the error statistics for %MAE and %RMSE as well as the SQV for three model statistics (passenger volumes at railway stations, passenger link volumes and road counts). Clearly, the primary calibration objective was public transportation. In both boardings, alightings and passenger volumes on links, more than 50% of all counts had an SQV higher than 0.9. A small percentage (<4%) had an unsatisfactory quality.

Table 5. Error statistics and quality measures

		boarding+alightings at rail stations[1]	rail passenger volumes[2]	road counts[3]
N		1’143	3’848	2’455
%MAE		22.2	11.4	20.7
%RMSE		62.2	20.3	31.3
SQV	%excellent, SQV≥0.9	55.6	55.3	35.4
	%good, SQV≥0.85	17.2	22.3	16.8
	%satisfactory, SQV≥0.8	12.5	12.3	15.8
	%sufficient, SQV≥0.7	10.7	7.6	21.7
	%unsatisfactory, SQV<0.7	3.9	2.6	10.4

1) f = 15’000; 2) f = 10’000; 3) f = 10’000

3.5 Time-of-day dependent network volumes

An important feature of agent-based simulation is high resolution of time. As a result, traffic volumes for any time of day over 24 hours is produced as output. Figure 7 shows the time-of-day distribution of passengers entering a train at a selected (medium-sized) station. Simulation results are compared to actual rail counts. The comparison shows a good fit at this level with the peaks being slightly lower in MOBi than in the observed counts. Currently a calibration effort is underway to further improve time-of-day distributions but at the current calibration level, the model is already applied in public transport service planning.

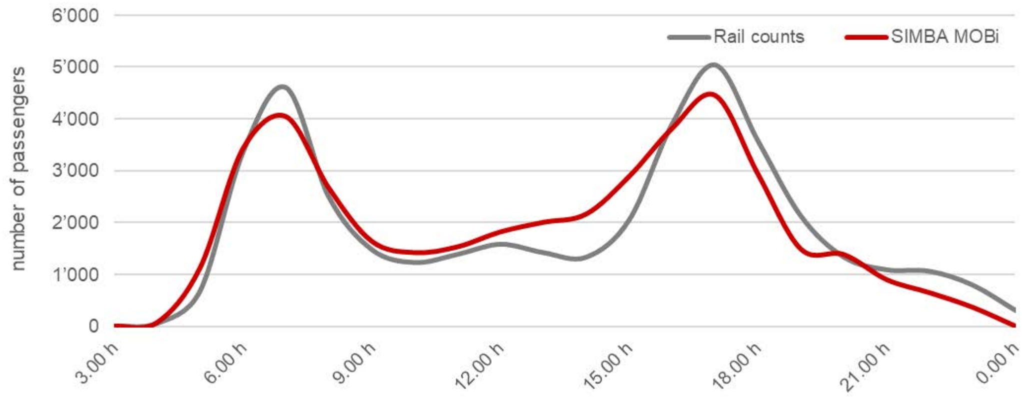

Figure 7. Time-of-day dependent travel demand on a selected train station

4. Summary and conclusions

A fully functional disaggregate multimodal travel model of Switzerland was developed. The model is microscopic through all model steps: from generation to network flow simulation. The travel demand module MOBi.plans uses an activity-based approach. It produces individual 24-hour day plans with exact geographic locations for all activities. Full consistency of time and space along the sequence of activities and travel within the 24-hour day plans is achieved for each agent based on a novel approach of plan-building and activity scheduling, that uses time budgets. The plan-building procedure respects the natural time and space constraints: each trip starts at the location of the previous activity; any travel and activity episode in the same day plan can only start after the previous episode is finished. This consistency or integrity is a necessary condition to feed the plans into the agent-based traffic flow simulation in the MOBi.sim module, which uses MATSim. Using this approach, activity-based demand has been successfully combined with agent-based traffic simulation and as a result, we have a model stream where each agent reacts to changes in transport supply through all model steps: from activity generation, destination and mode choice, over tour construction and day planning, through route choice and network flow simulation. The model enforces capacity constraints in the traffic flow simulation and capacity-constraint travel times are fed back into the agents' mode choice, destination choice and day planning. As shown with selected validation statistics, model calibration reached a level of quality that is expected of conventional macroscopic models at a national scale.

We identified the following research areas as important in advancing the methods of agent-based simulation:

- Increasing the speed of agent-based simulation using parallelization and high-performance computing.
- A better understanding of simulation noise, variation of model results within ensemble runs and convergence of MATSim simulation.
- The use of mathematical optimization to replace the plan-building heuristic which we apply in MOBi.plans.
- Empirical mobility research of travel demand peaks and the reasons how and when they are developing over the day, and methodological improvement of agent-based simulation to reproduce the peaks.
- Extending the idea of agents' learning based on an individual memory to all steps of individual travel demand, especially to the early steps (activity and tour generation, destination and mode choice, time choice and plan-building), where the activity-based approach involves individual preferences but no yet agent-based learning.
- Development of consistent mode choice models (trip- and tour-based) across travel demand and network simulation.

While we have initiated research projects addressing the research areas enumerated above (Bruno et al., 2019; Pougala et al., 2020), our internal model development is also working on improvement of usability and work-flow automation, and on computational efficiency. Work is also underway on including rail access with car and bicycle to the model. In addition, model elasticities and forecasting ability are being tested and analyzed.

SBB shares many software developments with the public in the open source environment. We also work with commercial software providers to advance the tools for agent-based modeling. Over the past three years, we have seen many improvements of software and methods and hope that these improvements will lower the threshold for other organizations to join us on the path towards putting agent-based simulation into real-world practice.

CRediT authorship contribution statement

Patrick Manser: conceptualization, methodology, software, validation, investigation, writing - original draft. **Wolfgang Scherr**: conceptualization, methodology, validation, writing - original draft, supervision, project administration. **Chetan Joshi**: methodology, formal analysis, writing - review & editing. **Nathalie Frischknecht**: validation, formal analysis, investigation, writing - original draft. **Denis Métrailler**: software, investigation, resources.

Acknowledgements

We thank Patrick Bützberger and David Henny for guidance; Kai Axhausen, Kai Nagel, Alex Erath and Michael Balmer for methodological advice; Andreas Justen for input data and validation; Michel Bierlaire for support with discrete choice models and the Biogeme software; Marcel Rieser and Sebastian Hörl for MATSim software extensions; Klaus Nökel and Martin Snethlage for reviewing the software architecture and VISUM support; Stephan Osterwald for support with research funding; the anonymous reviewers for helpful and constructive comments.

References

Arentze, T. A. (2013). Adaptive personalized travel information systems: A Bayesian method to learn users' personal preferences in multimodal transport networks. *IEEE Transactions on intelligent transportation systems*, 14(4), pp. 1957-1966.

Bhat, C. R., Guo, J. Y., Srinivasan, S., and Sivakumar, A. (2004). Comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record*, 1894(1), pp. 57-66.

Bierlaire, M. (2016). PythonBiogeme: a short introduction. *Report TRANSP-OR 160706*. Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Bodenmann B., P. Bürki, C. Philipp, N. Bernhard, K. Müller (2019). Synthetische Population 2017 – Modellierung mit dem Flächennutzungsmodell FaLC. *Report for Bundesamt für Raumentwicklung (ARE) and Swiss Federal Railways (SBB)*. Bern, Switzerland.

Bowman, J. L. (2009). Historical Development of Activity Based Model Theory and Practice. *Traffic Engineering and Control*, Vol. 50, no. 2: pp. 59-62 (part 1), no. 7: pp. 314-318 (part 2).

Bowman, J. L. and M. E. Ben-Akiva (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A*, 35(1), pp. 1-28.

Bösch, P., Müller, K., Ciari, F. (2016). The IVT baseline scenario. Paper presented at the *Swiss Transport Research Conference*. Ascona, Switzerland.

Boyce, D., H. Williams (2015). *Forecasting Urban Travel: Past, Present and Future.* Cheltenham, U.K. and Northampton, Mass.

Briem, L., Mallig, N., Vortisch, P. (2019). Creating an integrated agent-based travel demand model by combining mobiTopp and MATSim. *Procedia Computer Science,* 151, pp. 776–781.

Bruno, R., Mueller, M., Alonso, G., Hoefler, T. (2019). Towards High Performance Mobility Simulations. Paper presented at the *Swiss Transportation Research Conference*. Ascona, Switzerland.

Castiglione, J., Bradley, M., Gliebe, J. (2015). *Activity-Based Travel Demand Models: A Primer.* SHRP 2 Report S2-C46-RR-1. Transportation research Board (TRB). Washington, D.C.

Cetin, N., Nagel, K., Raney, B., Voellmy, A. (2002). Large-scale multi-agent transportation simulations. *Computer Physics Communications,* 147(1-2), pp. 559-564.

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

Danalet, A. and N. Mathys (2018). Mobility resources in Switzerland 2015. Paper presented at the *Swiss Transport Research Conference*. Ascona, Switzerland.

Esztergár-Kiss, D. (2017). System architecture of the activity chain optimization application. Proceedings of the *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems*, pp. 415-420. Napoli, Italy.

Federal Office for Spatial Development (2020). *Nationales Personenverkehrsmodell NPVM 2017.*

Federal Statistical Office (2017). *Population's transport behaviour 2015. Key results of the mobility and transport microcensus*. Neuchâtel, Switzerland.

Friedrich, M., Pestel, E., Schiller, C., Simon, R. (2019). Scalable GEH - a Quality Measure for Comparing Observed and Modelled Single Values in a Travel Demand Model Validation. *ransportation Research Record,* 2673(4), pp. 722-732.

Hilgert, T., Kagerbauer, M., Schuster, T., Becker, C. (2016). Optimization of Individual Travel Behavior through Customized Mobility Services and their Effects on Travel Demand and Transportation Systems. *Transportation Research Procedia*, 19, pp. 58–69.

Hillel, T., Pougala, J., Manser, P., Luethi, R., Scherr, W., Bierlaire, M. (2020, forthcoming). Modelling mobility tool availability at a household and individual level: A case study of Switzerland. Paper presented at the *hEART conference*. Lyon, France.

Hörl, S., Axhausen, K. (2019). Relaxation-discretization algorithm for spatially constrained secondary location assignment. Paper presented at the *TRB Annual Meeting*. Washington, D.C.

Hörl, S., Balac, M. and Axhausen, K.W. (2018). A first look at bridging discrete choice modeling and agent-based microsimulation in MATSim. *Procedia computer science*, 130, pp. 900-907.

Hörl, S. (2017). A MATSim scenario for autonomous vehicles in La Défense and Île-de-France. *IVT Working Paper – ETH research Collection*. Zürich.

Horni, A., Nagel, K., Axhausen, K. (2016). *The Multi-Agent Transport Simulation MATSim*. London.

Horni, A., Nagel, K., Axhausen, K. (2011). High-resolution destination choice in agent-based demand models. *Working Paper – ETH research Collection*. Zürich.

Mallig, N., Kagerbauer, M., Vortisch, P. (2013). mobitopp – a modular agent-based travel demand modelling framework. *Procedia Computer Science,* 19, pp. 854–859.

Meister, K., Rieser, M., Ciari, F., Horni, A., Balmer, M., Axhausen, K. (2008). Anwendung eines agentenbasierten Modells der Verkehrsnachfrage auf die Schweiz. Paper presented at *Heureka.* Stuttgart, Germany.

Moeckel, R., Kuehnel, N., Llorca, C. and Moreno, A. (2019). Agent-Based Travel Demand Modeling: Agility of an Advanced Disaggregate Trip-Based Model. Paper presented at the *TRB Annual Meeting*. Washington, D.C.

Ordonez, S. A. M. (2016). *Simulating Work-Leisure Cycles in Large Scale Scenarios. Models and Implementation*. *Doctoral thesis*. ETH, Zürich.

Pougala, J., Hillel, T., Bierlaire, M. (2020). Scheduling of daily activities: an optimization approach', Paper presented at the *Swiss Transportation Research Conference*. Ascona, Switzerland.

PTV Visum [computer software] (2019). PTV Planung Transport Verkehr AG. Karlsruhe, Germany.

Rasouli, S. and Timmermans, H. (2014). Activity-based models of travel demand: promises, progress and prospects. *International Journal of Urban Sciences,* 18(1). Abington, U.K.

Rieser, M., Métrailler, D., Lieberherr, J. (2018). Adding Realism and Efficiency to Public Transportation in MATSim. Paper presented at the *Swiss Transport Research Conference*. Ascona, Switzerland.

Roorda, M.J. Miller, E.J., Habib, K.M.N. (2008). Validation of TASHA: A 24-hour activity scheduling microsimulation model, *Transportation Research Part A*, 42(2008), pp. 360–375.

Scherr, W., Bützberger, P., Manser, P. (2019). SIMBA MOBi: Microscopic Mobility Simu-lation for Corporate Planning. Paper presented at the *European Transport Conference*. Dublin, Ireland.

Scherr, W., Bützberger, P., Frischknecht, N. (2018). Micro Meets Macro: A Transport Model Architecture Aiming at Forecasting a Passenger Railway's Future. Paper presented at the *Swiss Transport Research Conference*. Ascona, Switzerland.

Scherr, W., Bützberger, P. (2016). Calibration of the Swiss Rail Model Using Longitudinal Demand Observation. Paper presented at the *Swiss Transport Research Conference*. Ascona, Switzerland.

Sierpiński G., Staniek M., Celiński I. (2016). Travel behavior profiling using a trip planner. *Transportation Research Procedia*, 14, pp. 1743-175.

Vovsha, P., Anderson, R, Giamo, G., Rousseau, G. (2016). Integrated Model of Travel Demand and Network Simulation. Paper presented at the *International Conference on Innovations in Travel Modeling (ITM).* Denver.

Vovsha, P., Bradley, M., Bowman, J. (2004). Activity-based travel forecasting models in the United States: Progress since 1995 and Prospects for the Future. Paper presented at *EIRASS Conf. on Progress in Activity-Based Analysis*, Maastricht, NL.

Weis, C., Vrtic, M., Axhausen, K. W., & Balać, M. (2016). SP-Befragung 2015 zum Verkehrsverhalten. *Report published by Bundesamt für Raumentwicklung (ARE).*

Ziemke, D., Nagel, K., Bhat, C. (2015). Integrating cemdap and matsim to increase the transferability of transport demand models. *Transportation Research Record: Journal of the Transportation Research Board*, 2493, pp. 117–125.Washington, D.C.