An operational alternative to origin/destination matrices

Gunnar Flötteröd[1]

[1]Department of Communication and Transport Systems Linköping University, Sweden
& The Swedish National Road and Transport Research Institute
2025-01-19

SHORT SUMMARY

A round-trip based alternative to origin/destination matrices is presented. The formal approach treats the round-trips of a population as a distributed quantity and adopts statistical techniques to the evaluation of this distribution. An operational Bayesian calibration approach is presented as a round-trip-based counterpart to origin/destination matrix estimation. Several concrete applications of the framework are presented. We believe that the vastly increased representative power of a round-trip-based model when compared to origin/destination matrices, in combination with its compatibility with modern agent-based simulation packages, outweighs its somewhat more involved technical development.
Keywords: Operations research applications, activity-based modeling, transport network modeling

1 INTRODUCTION

Origin/destination (OD) matrices are ubiquitous in both person and freight network assignment modeling. The rows of a static OD matrix represent origins, its columns represent destinations, and its entries represent (freight or person) transport between origins and destinations. Dynamic OD matrices add a (usually discrete) time index, meaning that they map the triplet (origin, destination, time index) onto a transport demand. OD matrices are covered in any transport modeling textbook (e.g., Ortuzar and Willumsen, 2004; Cascetta, 2001).

The widespread use of OD matrices may be traced back to three ingredients: their simple and intuitive structure, their compatibility with mainstream network assignment models, and the availability of a comprehensive mathematical machinery for estimating OD matrices. Still, their appealing simplicity comes with limitations: An OD matrix does not encode any relationship between the individual movements it represents, neither does a dynamic OD matrix encode any temporal relationships. The resulting independence assumption across movements may violate mass conservation (e.g., more travelers leave a shopping mall than enter it) and causality (e.g., even if the time-sum of entries and exits to a shopping mall is zero, the entries may occur after the exits). In addition, the cell entries of a single OD matrix represent a homogeneous number of anonymous movements, and capturing heterogeneity by introducing group-specific OD matrices quickly reaches the computational limits of representing a correspondingly large number of OD matrices with correspondingly small entries. For similar reasons, the number of time steps resp. origins and destinations (usually traffic analysis zones) is limited, introducing a temporal resp. spatial aggregation bias.

These observations are not new and were one of the drivers for developing agent-based models (Horni et al., 2016; Nagel and Flötteröd, 2012). An agent-based transport demand is represented based on one trip-list per agent, with each trip being (minimally) annotated by origin, destination, and departure time. This allows, at the agent level, to ensure both spatial consistency (origin of one trip must be destination of the previous trip) and temporal consistency (departure time of one trip must not be earlier than arrival time of the previous trip). Agent-based models for strategic planning (where a long-term, stationary state of the transport system of interest is considered) may consider circular trip lists: Person agents return to their homes on a daily basis, freight vehicle agents return to their depots at possibly larger time intervals (consider, for example, weekly train schedules).

This document describes an operational simulation/estimation framework for trip-list-based transport representations. It makes concrete and develops further the ideas sketched in Flötteröd (2024), going beyond that reference by (i) specifying in detail a round-trip sampling approach, (ii) considering not just one but arbitrarily many round-trips simultaneously, and (iii) presenting a Bayesian calibration framework. We believe that the vastly increased representative power of this approach when compared to OD matrices, in combination with its compatibility with modern agent-based simulation packages, outweighs the somewhat more involved technical developments.

2 METHOD

Round-trip-lists

We discretize time into intervals (time bins) $k \in \mathcal{K} = \{1, \ldots, K\}$ and consider a population of agents $n = 1, \ldots, N$. A location set $\mathcal{L}_n$ of size L_n is available to agent n. We define agent n's location list $l_n = [l_n^{(1)}, \ldots, l_n^{(J_n)}]$ and departure time bin list $d_n = [d_n^{(1)}, \ldots, d_n^{(J_n)}]$ where $J_n \in \{1, \ldots, \min\{J^{\max}, K\}\}$ is the finite length of both lists, $l_n^{(j)}$ is the origin location of the jth trip, and $d_n^{(j)}$ is its departure time bin. The location and departure time list of agent n constitute its trip list $x_n = (l_n, d_n)$. To treat trip lists independently of an agent, we introduce the function $J(\cdot)$ that maps a trip list onto its length; obviously, $J(x_n) = J_n$. Causality is established by requiring

$$d_n^{(j)} < d_n^{(j+1)} \quad \text{for all } j \in \{1, \ldots, J_n - 1\}. \tag{1}$$

(Note that this implies the above requirement $J_n \leq K$.) Repeated trips with the same departure location are allowed for to enable the representation of intra-zonal travel when locations are traffic analysis zones.

We subsequently consider *round-trips* where the destination location of the last trip is the departure location of the first trip.[1] For instance, given an hourly time discretization, the round-trip

$$([\text{home, office, shopping mall}], [6,\ 16,\ 17])$$

means that the considered agent leaves home at 6 am for office work, departs at 4 pm from the office, makes a short stop at the shopping mall that is planned to end at 5 pm, and then returns back home. The departure time bins represent a *desired* time structure that may or may not be compatible with a given physical reality of finite travel speeds; more on this further below.

We consider *distributed* population round-trips $X = [X_1, \ldots, X_N]$ where $\Pr(X = x) = \Pr(X_1 = x_1, \ldots, X_N = x_N)$ is a discrete probability with (for finite population and maximum trip-list length) finite support. To evaluate this distribution, a method to sample round-trips from a given target distribution $p(x)$ is required. For generality, we do not yet make assumptions about where this target distribution comes from (ample examples further below) but merely assume it to be given. We rely here on the Metropolis-Hastings (MH) algorithm (Hastings, 1970). This algorithm has well-known advantages (generality) and disadvantages (possibly long run-times), the latter having been addressed in a substantial body of literature. Ross (2012) offers a detailed

[1] To capture trip-lists that are not round-trips within the same framework, one may add a (J_n+1)th trip where $l_n^{(J_n+1)}$ defines the destination of the J_nth trip and $d_n^{(J_n+1)}$ is arbitrary.

introduction. We focus subsequently on a basic round-trip-specific instance of the MH algorithm, omitting general-purpose discussions around its practicalities (convergence tests, extraction of statistics, etc.).

The state space of the considered MH algorithm is composed of population round-trips $x = [x_1, \ldots, x_N]$. The algorithm requires an irreducible proposal distribution $p(x, y)$ from any state x to any state y, which is subsequently developed.

MH proposal distribution for a single round-trip ($N = 1$)

Since a single round-trip is considered, the agent index n is suppressed in this section. Four operations on a single round-trip are subsequently defined; their combination will yield the desired proposal distribution.

INS(ERT) Given a round-trip x of length $J < J^{\max}$, an insertion index i is uniformly drawn from $\{1, \ldots, J+1\}$. A new location is drawn uniformly from the location set $\mathcal{L}$, and a new departure time bin is drawn uniformly from the set of not yet used departure time bins $\mathcal{K} \setminus \cup_{j=1}^{J} d^{(j)}$. For $i \leq J$, these values are inserted into the location and departure time list at index i. For $i = J+1$, they are appended to the end of the lists. To comply with (1), the departure time list is subsequently ordered by increasing magnitude *without changing the ordering of the location list.* Letting the round-trip y be the result of an insert operation at index i in round-trip x, the probability of obtaining y from x is

$$q_{\mathrm{INS}}(x, y) = \frac{1}{J+1} \cdot \frac{1}{L} \cdot \frac{1}{K-J}. \tag{2}$$

Since the length of the round-trip increases by one, the new round-trip differs from the old one.

REM(OVE) Given a round-trip x of length $J > 0$, an index i is uniformly drawn from $\{1, \ldots, J\}$ and the ith element is removed from the location and departure time list. Letting y be the result of a remove operation from round-trip x, the probability of obtaining y from x is

$$q_{\mathrm{REM}}(x, y) = \frac{1}{J}. \tag{3}$$

Since the length of the round-trip decreases by one, the new round-trip differs from the old one.

FLIP_LOC(ATION) Given a round-trip x of length J, a location flip index i is uniformly drawn from $\{1, \ldots, J\}$. A new location is uniformly drawn from $\mathcal{L} \setminus l^{(i)}$ to replace $l^{(i)}$. Letting y be the result of flipping a location in round-trip x, the probability of obtaining y from x is

$$q_{\mathrm{FLIP\_LOC}}(x, y) = \frac{1}{J} \cdot \frac{1}{L-1}. \tag{4}$$

Since the newly drawn location must differ from the original value, the new round-trip differs from the old one.

FLIP_DEP(ARTURE) Given a round-trip x of length J, a departure time flip index i is uniformly drawn from $\{1, \ldots, J\}$. A random departure time is drawn from $\mathcal{K} \setminus \cup_{j=1}^{J} d^{(j)}$ to replace $d^{(i)}$. The resulting departure time list is sorted by increasing magnitude without changing the ordering of the location list. Letting y be the result of flipping a departure time in round-trip x, the probability of obtaining y from x is

$$q_{\mathrm{FLIP\_DEP}}(x, y) = \frac{1}{J} \cdot \frac{1}{K-J}. \tag{5}$$

Since the newly drawn time must differ from the original value, the new round-trip differs from the old one.

Given a round-trip y that results from applying any of the above operations to a given round-trip x, the applied operation is uniquely given: $J(y) > J(x)$ can only result from INS; $J(y) < J(x)$ can only result from REM; $J(x) = J(y)$ and $l_x \neq l_y$ can only result from FLIP_LOC; $J(x) = J(y)$ and $d_x \neq d_y$ can only result from FLIP_DEP; $J(x) = J(y)$ and $l_x = l_y$ and $d_x = d_y$ is impossible. Let $\phi_{\text{INS}}, \phi_{\text{REM}}, \phi_{\text{FLIP\_LOC}}, \phi_{\text{FLIP\_DEP}}$ be the selection probabilities of the respective operations; they are strictly positive and sum up to one.[2] Drawing an operation, applying it to round-trip x and receiving round-trip y hence occurs with the following probability:

$$q_{[1]}(x, y) \quad = \quad \begin{cases} \phi_{\text{INS}} \cdot q_{\text{INS}}(x, y) & \text{if } J(y) > J(x) \\ \phi_{\text{REM}} \cdot q_{\text{REM}}(x, y) & \text{if } J(y) < J(x) \\ \phi_{\text{FLIP\_LOC}} \cdot q_{\text{FLIP\_LOC}}(x, y) & \text{if } J(y) = J(x) \wedge l_x \neq l_y \\ \phi_{\text{FLIP\_DEP}} \cdot q_{\text{FLIP\_DEP}}(x, y) & \text{if } J(y) = J(x) \wedge d_x \neq d_y \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

This distribution is irreducible because any state (round-trip) b can be reached from any other state a with positive probability: (i) If $J(a) = J(b)$, a sequence of FLIP_LOC and FLIP_DEP operations that turn a into b is possible. (ii) If $J(a) < J(b)$ resp. $J(a) > J(b)$, a sequence of INS resp. REM operations is possible that yields an intermediate state a' with $J(a') = J(b)$, which then can be turned into b according to (i).

MH proposal distribution for multiple round-trips ($N > 1$)

We return to indexing the different round-trips in x by $n = 1, \ldots, N$ and compose the proposal distribution for population round-trips from that for single round-trips. We sweep once over all round-trips $n = 1, \ldots, N$ and modify round-trip x_n with probability $\phi_{\text{MOD}} > 0$ into $y_n \neq x_n$ according to $q(x_n, y_n)$ in (6), otherwise we let $y_n = x_n$.[3] If this results in $x = y$ (meaning that no modification has taken place), we repeat the process until $x \neq y$. This implements the proposal distribution

$$q_{[1:N]}(x, y) \quad = \quad \begin{cases} \frac{1}{1-\phi_{\text{MOD}}^N} \prod_{n=1}^{N} \begin{cases} \phi_{\text{MOD}} \cdot q_{[1]}(x_n, y_n) & \text{if } x_n \neq y_n \\ 1 & \text{otherwise} \end{cases} & \text{if } x \neq y \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

We have already established that any single round-trip can be turned into any other single round-trip by a suitable operation sequence. Consider now the problem of turning any population round-trip $a = (a_1, \ldots, a_N)$ into any other population round-trip $b = (b_1, \ldots, b_N)$. This can be achieved by only selecting $n = 1$ for modification until a_1 has been turned into b_1, then doing the same for $n = 2$, etc. This sequence of selecting round-trips for modification arises with positive probability, and so does every single modification from a_n into b_n. This establishes the irreducibility of the population proposal distribution $q_{[1:N]}$.

MH target weights

A major advantage of the MH algorithm in the given context is that it only requires an unnormalized version $t(x)$ of a given target distribution $p(x)$, meaning that any $t(x)$ with $p(x) =$

[2]Our experimentation so far suggests that a uniform selection distribution performs well.
[3]Letting $\phi_{\text{MOD}} = 1/N$ has performed well in our experimentation so far.

$t(x)/\sum_{x'} t(x')$ for all x is sufficient to use the algorithm to draw from $p(x)$. This avoids a normalization of the target distribution over the possibly gigantic state space.

The time structure of a round-trip represents an ambition; its physically feasibility depends on the movement duration between its locations. When there is a need to evaluate the physical realization of a round-trip, a deterministic mapping g from the population round-trips x onto some data structure representing the realized movement experiences (a movement simulation) can be used. The corresponding target distribution would still only depend on x but internally also evaluate the movement simulation g. We subsequently suppress the possible use of g when writing out a target distribution.

The concrete form of the target distribution is application-specific. We discuss below the problem of calibrating a population round-trip distribution against some given data-set. This constitutes the round-trip-based counterpart of OD matrix estimation.

Round-trip calibration and uninformed prior distribution

We adopt a Bayesian approach. Given the data y, a likelihood function $p(y \mid x)$ is formulated that expresses the probability of observing the data y given the population round-trips x. A Bayesian approach to conditioning the round trip distribution onto the data amounts to sampling from the (un-normalized) target distribution

$$t(x \mid y) \quad \sim \quad p(y \mid x) p^{\text{prior}}(x) \tag{8}$$

where the prior distribution $p^{\text{prior}}(x)$ represents available knowledge about the population round-trips before having seen the data y; if available, this may comprise travel behavioral assumptions (Flötteröd et al., 2011).

The construction of an uniformed prior distribution requires some care. Considering any agent n, the number of round-trip configurations of length J available to that agent is

$$\#_n(J) \quad = \quad L_n^J \binom{K}{J} \tag{9}$$

where the first factor accounts for all possible location configurations and the second factor represents the number of possible departure time configurations, given their ordering by increasing magnitude. This means that the number of available round-trip configurations of a given length grows combinatorically with that length. A naive uniform prior distribution over all possible round-trips of an agent would hence be biased towards longer round-trips.

A maximum entropy (ME) approach is hence adopted to represent a maximally uninformed round-trip distribution. The prior distribution $p_n^{\text{ME}}(x_n)$ over all possible round-trips of agent n is chosen to maximize entropy subject to the constraint that the expected round-trip length equals some exogenously given parameter $\overline{J}_n$.[4] The corresponding Lagrangian reads

$$L(p_n^{\text{ME}}, \lambda, \gamma) \quad = \quad -\sum_{x_n} p_n^{\text{ME}}(x_n) \ln p_n^{\text{ME}}(x_n) \ldots \tag{10}$$

$$+\lambda \cdot \left(\sum_{x_n} p_n^{\text{ME}}(x_n) - 1 \right) + \gamma \cdot \left(\sum_{x_n} J(x_n) p_n^{\text{ME}}(x_n) - \overline{J}_n \right) \tag{11}$$

where the first constraint with multiplier λ requires probabilities to sum to one and the second constraint with multiplier γ enforces the expected length. Evaluating first-order conditions on p_n^{ME}:

$$\frac{dL}{dp_n^{\text{ME}}(x_n)} = 0 \quad \Rightarrow \quad e^{\lambda - 1} e^{\gamma J(x_n)}. \tag{12}$$

[4]The arguably simplest setting of this parameter is $J_n = J$ for all n, with J the estimated total number of trips made in a study region divided by its population.

Inserting this into the probability sum constraint yields

$$p_n^{\mathrm{ME}}(x_n) = \frac{e^{\gamma J(x_n)}}{\sum_{J=0}^{\min\{K,J^{\max}\}} \#_n(J)e^{\gamma J}}, \tag{13}$$

which has, unsurprisingly (Anas, 1983), the form of a multinomial logit round-trip choice model with scale parameter γ. This parameter is obtained by inserting (13) into the mean length constraint and numerically solving the one-dimensional problem

$$\frac{\sum_{J=0}^{\min\{K,J^{\max}\}} J \cdot \#(J)e^{\gamma J}}{\sum_{J=0}^{\min\{K,J^{\max}\}} \#(J)e^{\gamma J}} = \overline{J}_n \tag{14}$$

for γ.

Overall, the ME population round-trip distribution becomes

$$p^{\mathrm{ME}}(x) = \prod_{n=1}^{N} p_n^{\mathrm{ME}}(x_n) \tag{15}$$

in combination with (13) and (14). The symmetry of this development to the long-standing ME-based estimation of OD matrices may be noteworthy (Van Zuylen and Willumsen, 1980).

3 RESULTS AND DISCUSSION

A primary result of this work is the above described modeling/estimation framework, which is accompanied by a freely available software implementation at `https://github.com/vtisweden/matsim-projects/tree/master/roundtrips`. The remainder of this section summarizes existing and ongoing applications of the framework, illustrating its versatility and applicability.

An early version of the framework was applied by Flötteröd (2024) to study daily electrical vehicle charging patterns in the Swedish municipality Skellefteå. To accommodate location-specific charging decisions, the physical network was extended into one where each location was duplicated into one (location, charging) and one (location, no-charging) node; sampling paths on the extended network hence also induced temporal charging patterns. The target distribution aimed to represent basic time- and land use assumptions. Figure 1(left) shows part of the study region, and Figure 1(right) exemplifies the estimated home- and time-of-day-dependent en-route charging patterns.

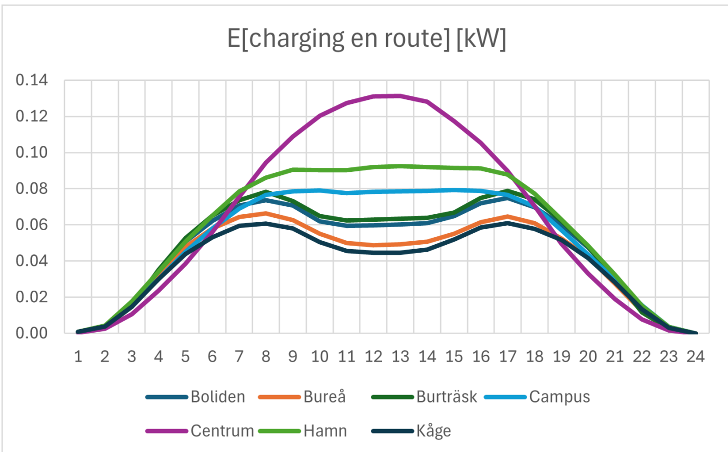

Figure 1: Electric vehicle charging analysis

A more recent application is that of Charalampidou et al. (2025) where implications of shopping location placement and opening times in a fifteen-minute-city setting are investigated. Here, the round trip encodes weekly activity/travel patterns. To encode activity participation, the physical network is extended into one where each location is duplicated into several (location, activity) nodes, representing participation in a particular activity at that location. The target distribution is set such that round trips respecting weekly activity-specific time use assumptions receive high probabilities. Figure 2(left) shows the study region (the Liesing district in the city of Vienna, Austria) and Figure 2(right) exemplifies the resulting shopping activity participation time structure.

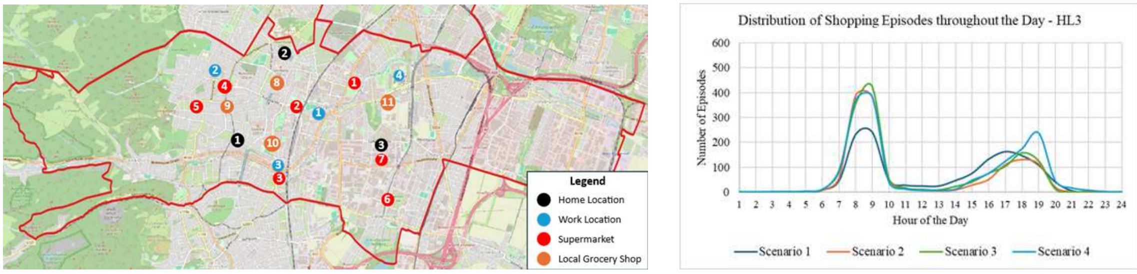

Figure 2: Time use and activity participation in a 15-minute-city

An instance of a multi-round trip calibration for the entire city of Vienna is presented by Rupprecht et al. (2025). The approach starts out from the ME prior (15) and combines it with two likelihood terms, one representing the all-day reproduction of an available static target OD matrix, and the other one representing spatio-temporal travel and activity participation summary statistics. The scatter-plot in Figure 3(left) displays the almost perfect reproduction of the static target OD matrix from a representative population of 50'000 round-tips, while Figure 3(right) illustrates the within-day time structure of the same 50'000 round-trips. The resulting population round-trips serve as initial travel plans file for an agent-based route/mode/departure time network assignment model.

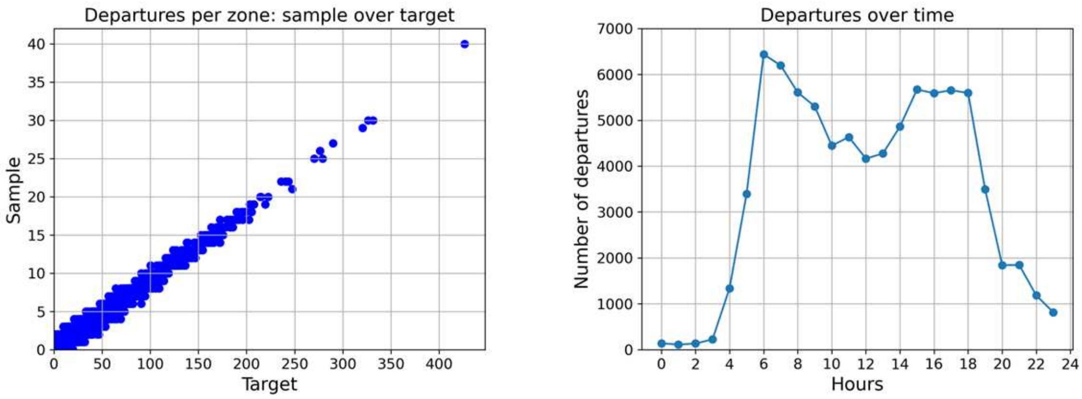

Figure 3: All-day travel pattern synthesis for agent-based traffic assignment

One ongoing but not yet documented application is the inclusion of travel survey data in the population round-trip estimation problem, aiming to complement the Swedish national person transport model with time-of-day dynamics. Another application is the generation of freight vehicle round-trips for Sweden and neighboring countries, aiming at an improved representation of freight consolidation in the Swedish national freight model.

4 SUMMARY

An operational round-trip based alternative to OD matrices has been presented, comprising a formal framework, an estimation method, and several application examples. The idea of replacing OD matrices by (round-)trip-lists has been around for a long time, and any agent-based transport simulation packages inevitably relies on some kind of traveler- or vehicle (round-)trip-list discretization of a possibly given OD matrix. This document contributes to this a formal and operational machinery to sampling population round trips from general target distributions. A concrete Bayesian calibration approach resting on this machinery is developed, offering a round-trip based alternative to OD matrix estimation.

Ample further developments of the presented method are possible; these comprise the specification of other than uniform MH proposal distributions, refinements of the basic MH approach (e.g. into Gibbs sampling where the round trip distribution of any single agent is conditioned on the round trips of all other agents), and the not obvious exploitation of parallel computing facilities to speeding up the inherently sequential MH process. A comparison to generative AI approaches such as that of Shone and Hillel (2024) may be attempted, even though much of the appeal of the presented method is its derivation from first principles, while generative AI appears to be more of a black-box approach.

ACKNOWLEDGMENTS

This work was supported by the Swedish Transport Administration, grants TRV 2023/33572 and TRV 2023/109327.

BIBLIOGRAPHY

Alex Anas. Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research Part B: Methodological*, 17(1):13–23, 1983. ISSN 0191-2615. doi: https://doi.org/10.1016/0191-2615(83)90023-1. 2

E. Cascetta. *Transportation Systems Engineering: Theory and Methods.* Kluwer Academic Publishers, 2001. 1

G. Charalampidou, G. Flötteröd, and Y. Susilo. Sampling weekly activity schedules to analyze shopping behavior under the 15-minute city concept. In *(submitted) 13th Symposium of the European Association for Research in Transportation*, Munich, Germany, 2025. 3

G. Flötteröd. Charging demand analysis in the face of overwhelming uncertainty. In *12th Symposium of the European Association for Research in Transportation*, Aalto, Finland, 2024. 1, 3

G. Flötteröd, M. Bierlaire, and K. Nagel. Bayesian demand calibration for dynamic traffic simulations. *Transportation Science*, 45(4):541–561, 2011. 2

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 04 1970. ISSN 0006-3444. doi: 10.1093/biomet/57.1.97. 2

Andreas Horni, Kai Nagel, and Kay Axhausen, editors. *Multi-Agent Transport Simulation MATSim.* Ubiquity Press, London, Aug 2016. ISBN 978-1-909188-75-4, 978-1-909188-76-1, 978-1-909188-77-8, 978-1-909188-78-5. doi: 10.5334/baw. 1

K. Nagel and G. Flötteröd. Agent-based traffic assignment: going from trips to behavioral travelers. In R.M. Pendyala and C. Bhat, editors, *Travel Behaviour Research in an Evolving World*, chapter 12, pages 261–293. Emerald Group Publishing, Bingley, United Kingdom, 2012. 1

J. Ortuzar and L.G. Willumsen. *Modelling Transport*. Wiley, 2004. 1

S.M. Ross. *Simulation*. Knovel Library. Elsevier Science, 2012. ISBN 9780124158252. 2

F.-X. Rupprecht, G. Flötteröd, and Y. Susilo. Metropolis-hasting based synthesis of all-day round-trips from od-matrices and limited population segmentation. In *(submitted) 13th Symposium of the European Association for Research in Transportation*, Munich, Germany, 2025. 3

F. Shone and T. Hillel. Activity sequence modelling with deep generative models. In *12th Symposium of the European Association for Research in Transportation*, Aalto, Finland, 2024. 4

Henk J. Van Zuylen and Luis G. Willumsen. The most likely trip matrix estimated from traffic counts. *Transportation Research Part B: Methodological*, 14(3):281–293, 1980. ISSN 0191-2615. doi: https://doi.org/10.1016/0191-2615(80)90008-9. 2