Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2017) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

20th EURO Working Group on Transportation Meeting, EWGT 2017, 4-6 September 2017, Budapest, Hungary

Generating macroscopic, purpose-dependent trips through Monte Carlo sampling techniques

Ariane Scheffer [a*], Guido Cantelmo [a], Prof. Francesco Viti [a,]

[a] *Faculty of Science, Technology and Communication, University of Luxembourg, 6, rue Coudenhove-Kalergi, L-1359 Luxembourg-Kirchberg, Luxembourg*

Abstract

While estimating origin-destination (OD) demand flows usually requires a large amount of data, nowadays a key issue in traffic engineering is to estimate the trip purpose while protecting user privacy. The aim of this work is to derive from macroscopic and aggregate information the temporal distribution for the production of each traffic zone of a system, with a trip-purpose specification. We suggest different procedures for estimating the production factors, which are based on the precision level of the available information. If time-dependent demand data is available, the production factor can be estimated through a simple Monte Carlo simulation model. Otherwise, a Markov Chain Monte Carlo (MCMC) approach is proposed to approximate a set of functions that describe the production of purpose-specific trips with regard to one specific zone along the day. This algorithm requires a lower level of input information and computes the likelihood with regard to the number of generated and attracted trips. Application of the models is shown using available real data collected through a one-week travel diary within the area of Ghent, Belgium.

Keywords: OD matrix, Activity based, Demand, Markov Chain Monte Carlo, Estimation

1. Introduction

The last decades have witnessed an immense effort in bringing travel demand models to a new level of comprehension. While, since the early 80s, person travel has been modelled with a trip-oriented rather than activity-

[*] Corresponding author. Tel.: +352 46 66 44 5593; fax: +352 46 66 44 35593.
E-mail address: ariane.scheff@gmail.com

oriented specification, this approach has been universally criticized for being unrealistic (McNally, 2007). Activity-Based (AB) models have been also developed, but their application is limited due to the high cost of collecting input data (Bowman and Ben-Akiva, 2001). The advent of new technologies allowed researchers to gather immense volume of information, making possible to implement AB models on large urban and regional scenarios. Despite this technological advancement, calibrating a reliable demand model is still a difficult challenge (Toole et al., 2015). One of the main reasons is that different assumptions within the demand model lead to different errors. For instance, a disaggregate model can be extremely powerful for long-term prediction of the demand or for evaluating phenomena such as activity relocation. However, when the congestion during the rush hour is the main concern of the modeller, this level of detail is not required and may even lead to biased analysis.

By limiting our focus on the problem of linking input data and demand models, research can be divided in two main branches: including information on the trip purpose within standard trip-based models (*macroscopic* representation of the demand) and generating a fully disaggregate population of agents to evaluate tour of activities (*microscopic* representation of the demand).

Concerning the first group of models, the most widely adopted methodology is the so called Four-Step Model. Because its application is nearly universal, on the one hand various methods are still developed to tend improving the efficiency of this model (Lim and Kim, 2016), on the other hand other approaches use new data sources as input, in order to avoid collecting expensive travel diaries (Toole et al., 2015). However, all these models provide a limited insight in terms of temporal distribution of the demand and its activity specification. In this sense, the most established approach to reconstruct a realistic temporal profile for the travel demand is the dynamic OD estimation problem (Antoniou et al., 2016; Cascetta et al., 1993). While many works focus on improving the consistency between OD matrix and traffic performances (Antoniou et al., 2015), there are only a few works dealing with the critical point of explicitly capturing the behavioural component of the demand (Flötteröd, 2009). Many of these involve a utility-based approach (Flötteröd et al., 2009). Although incorporated on a broader scale, the utility evaluation is still based on an individual point of view. Because they are built on the relationship between the benefits to perform an activity at a certain time and the scheduling, models based on the utility theory do fully consider the behavioural aspect of the demand. (Ettema and Timmermans, 2003) offer an extensive review of these, within the framework of departure time choice. (Bhat, 2002) describes application of random utility in discrete choice models for travel demand analysis.

Alternative demand generation methods emerged, which integrate inside the traditional Four-Step-Model a microscopic alternative (Vrtic et al., 2007). However, these models still rely on highly specific probability functions, which are very hard to calibrate without a large volume of data.

Concerning the second group, on the contrary, the available models are by nature able to integrate this behavioural component originally lacking in trip-based models. Concretely, procedures are mostly based on the creation of a virtual set of agents and households, provided with specific attributes and subject to activity scheduling (Arentze and Timmermans, 2004; Bowman and Ben-Akiva, 2001). This founding step of disaggregate demand modelling is well known in literature as Synthetic Population Generation. The starting point is usually composed of aggregate socio-economic characteristics together with disaggregate information of a sample of the population. Merging aggregate data from different sources means incorporating strong assumptions on their distributions (Farooq et al., 2013). The two main options to generate this data are synthetic reconstruction, and in particular Iterative Proportional Fitting (IPF), and reweighting methods like Combinatorial Optimization (CO) (Mueller and Axhausen, 2011). More recently, Farooq et al. (2013) introduced a third category: Markov process-based methods. In the IPF algorithm, a contingency table is evaluated iteratively, based on the correlation of attributes in the sample; a population is created by replicating the sample accordingly, it has been used since a long time and is still recently (Arentze et al., 2007; Duguay et al., 1976; Ye et al., 2009). In the CO, a weight is linked to the sample to select a combination of households from the dataset (Voas and Williamson, 2000).

The fact remains that, if these complex models can yield good results, major drawbacks appear when evaluating their usage. On one hand, the aggregate data required needs to be very consistent and extremely timely and accurate. On the other hand, data sources need to be representative of the entire population. To get viable output of such models, a large amount of information is needed and the performance usually increases with the quantity, quality and precision of inputs (Barthelemy and Toint, 2013; Farooq et al., 2013). However, in many countries, such as Belgium, privacy restriction are so tight to make almost impossible to implement AB models without conventional (expensive) travel surveys (Barthelemy and Toint, 2013). In order to overcome this issue, sample-free synthetic reconstruction methods have recently appeared (Barthelemy and Toint, 2013; Gargiulo et al., 2010). They overcome the restriction of micro-samples or travel-survey necessity but are still based on very specific probability distributions. Besides, experiments concluded their lower performance in comparison to sample-based approaches (Lenormand and Deffuant, 2012).

We believe that explicitly capturing the correlation between an aggregate representation of the demand and its behavioural component is needed to reduce uncertainty in transport demand models. In this work, the focus is on reducing as much as possible the required information at the individual level. This paper, presents the possibility to use Monte Carlo (MC) simulation models to characterize demand flows with respect to traffic zones. In this approach the demand is modelled as a convolution of different activity patterns. The problem of estimating activity patterns from the OD matrix turns in this way to the problem of finding the best fitting for our static/dynamic demand sample.

Since the proposed approach tries to estimate the number of (purpose-dependent) components within an OD matrix, it can remind of mixture models. However, MC models have the advantage of including additional information in the estimation and so reduce the number of unrealistic solutions with respect to the Gaussian mixture model for example. MC methods are more and more used in the travel behaviour modelling and in the creation of synthetic population (Beckman et al., 1996) since they allow to draw samples and to estimate discrete outcomes from known probabilities for the different variables used to qualify agents of the disaggregate models.

The rest of this paper is structured as follows. The next section introduces the Monte Carlo based methodology, which will be used to estimate the probability functions. Then, Section 2 introduces the database we use to test the reliability of the proposed approach and the case study. Lastly, in Section 3 conclusions are discussed.

2. Methodology

To separate the demand into purpose-dependent flows, different Monte-Carlo simulation methods have been used. This class of algorithms is employed in many domains including transportation engineering. In this case, we study standard Monte Carlo and Markov Chain Monte Carlo (MCMC) approaches. For static cases, the first approach uses a disaggregation of flows into individual trips and draws individual and adequate trip purposes from a probability distribution. Preferably this simple method, named crude Monte Carlo Sampling (MCS) is used due to its simplicity and establishment in demand modelling. In the dynamic case, the distribution along the day is approximated through MCMC, which explicitly accounts for the temporal correlation between activities.

2.1. Static case: crude Monte Carlo Sampling

The starting point of the crude Monte Carlo sampling methodology is an OD matrix. To be put into practice, a distribution of the property of interest can be used or approximated to calculate attribute-specific generation factors. If a dynamic OD matrix is available this methodology allows calculating these factors for different periods of the day, even with approximated and static probability distribution.

The procedure is based on a disaggregation of the available OD matrices, in a way that the total number of trips corresponds to a list of journeys. This list is then provided with attributes. In contrast to activity-based demand models, the disaggregated unit of interest remains a trip and not an individual user. Random draws are used to predict a single choice and the goal is to obtain a valid distribution of trip purposes. Numbers of different criteria can be used, and all the conditioning properties can be inputs of the simulation, as long as the information is known with sufficient precision. Explanatory variables are easily introduced, and without increasing too much the complexity of the simulation. According to the number of travellers and complexity of the inputs, different sampling methods (variance reduction techniques for example) can be implemented, in order to obtain more accurate results.

Starting from the OD matrix, the disaggregation process returns a matrix of trips containing their origin and destination. Attributes, which in this study are the trip purposes, are then allocated to individual trips resulting from this disaggregation. In this conception, we use a list of coupled attributes of interest to directly observe and infer conditional probabilities and estimate generation factors of zones by time of the day. The selection of the attribute is performed by selecting the one that appears the most in the number of iterations (number of times the Monte Carlo simulation is performed), for example for the activity type ATj, in zone zi and for the reference time-period t the following equation is used to select activity M:

$$M^{zi}{}_{ATj}(t) = \frac{(Total\ Number\ of\ Activity\ ATj)}{(Total\ number\ of\ iteration)} \tag{1}$$

Once the activities have been allocated to each of the trips, this information is used to infer the activity type as function of zones, for a specific time interval (be it period of the day or hours for instance) and so be able to estimate a distribution. In order to take advantage of the number of iterations, the sum of trips having a specific attribute considers all different MCS performed. Results are observed to calculate specific (in time t and for trip-purpose *ATj*) generation factors *F* according to the following equation, summing “informed” trips *T* (i.e. trips for which a purpose has been determined):

$$F^{zi}_{ATj}(t) = \frac{\sum_{\mathrm{N}\ iterations} \sum_t T^{zi}_{ATj}}{\sum_{\mathrm{N}\ iterations} \sum_{\#\ zones} \sum_t T^{z}_{ATj}} \tag{2}$$

Considering several iterations does not affect the calculation of the factor but allows a better coverage of all the options, making good use of the MCS.

2.2. Dynamic case: Markov Chain Monte Carlo

The goal of the Markov Chain Monte Carlo (MCMC) model is to detect the attribute-specific primitives of the daily demand. In order to perform this estimation, the first element to consider is the number of options for the attribute of interest and their initial distribution. Then decisions need to be made about the a-priori knowledge of the parameter of interest: functional form and unknown parameters. For each of the unknown parameters, assumptions can also be integrated in the procedure, which will allow controlling the result and obtain more realistic estimations. This is known as the prior of parameters. This distribution can be informative or not, and can differ for each of the parameters.

In each iteration, parameters are renewed and their combination is evaluated according to the a-priori information. This given score is combined to another one which takes into consideration how the set of parameters explains the available data. This likelihood of the distribution is calculated with respect to “evidences”. These evidences are in the form of a signal corresponding to the number of trips arriving to or departing from one specific zone, along the day. The ideal configuration is to have that signal in the form of a dynamic OD matrix, nevertheless the point of view can be adapted to handle GSM data or loop detector, for example.

The set of parameters is accepted or rejected as a whole and each parameter serves as a starting point for the following proposed parameter. For this reason, a high number of iterations must be carried out, and the required number increases rapidly with the number of parameters handled. Nevertheless, the methodology is applied for each zone in parallel, independently. For this reason, the complexity is not raising and the rapidity is stable with and increasing number of areas.

In the case of activity type estimation, the method takes as prior the assumptions on the generation of activities, along the day. The priors can be of different types, and specific to different zones, according to the available information and assumptions.

For each zone, we consider that the complete demand *D* is the combination of the demand for *n* activity types, being each characterized by a probability of appearance *p* and an individual probability distribution *W*.

$$D_Z = \sum_{i=1}^{n} p_{iZ}\ W_{iZ} \tag{3}$$

The distribution of the demand along the day is the sum of the *n* distributions, each one corresponding to an activity, multiplied by a factor corresponding to the probability of performing this activity rather than another. In the study presented here, no variations of the activity type’s spread are taken into consideration. Nevertheless, the flexibility of Markov Chain Monte Carlo would allow to take this parameter as an additional criterion.

The next choice to be made is the functional form of the distribution. Criteria for the choice of the distribution shaping the demand are to obtain an asymmetric distribution, without increasing the complexity of parameter estimation. Furthermore, the parameters should include a parameter which determines the "location" or shift of the distribution for estimating a typical departure time by activity. In the testing example, all the activity types are based on the extreme value distribution. Studying different functional forms is beyond the scope of this paper and will be the focus of a future paper.

The role of the simulation is then to evaluate the plausible parameters that explain the most the data for each activity component. All of these parameters are subject to the output of the simulation in the form of a posterior i.e. a probability distribution. The prior distribution is a uniform distribution, between zero and two. The location parameter, on the contrary has a more meaningful interpretation. Because it relates directly to the typical departure time of a specific activity type, this is where the knowledge and assumptions are the most easily introduced.

3. Case Study

3.1. Database

The dataset used for validating the proposed model is the *BMW* dataset (Castaigne et al., 2009), which was collected by the University of Namur in 2008. 717 valid travel diaries were collected, which describe a one-week period for the city of Ghent, Belgium. Other analysis of this database are available in (Raux et al., 2016).

In order to apply the proposed methodology, the first necessary step is to create a dynamic OD matrix. After ensuring the consistency of the database, we artificially generated an OD matrix by aggregating all trips occurring during weekdays of the study period. No difference is made between working days. In the BMW study, the area was composed of 17 zones. Even though these zones are difficult to be identically recreated, postal codes have been used to cluster the respondents in 17 artificial zones. Five supplementary geographical units were created for the city centre of Ghent, inside which trips represent 61% of the complete demand.

Activities performed were also clustered: starting from a detailed description (12 possible trip purposes) we deducted five groups:

- AT1 - Activities usually located in residential areas i.e. “$Home$” and “$Visit\ to\ family\ or\ friends$”;
- AT2 - Activity “$Work$”;
- AT3 - Leisure activities, such as “$walking/\ riding$”, “$leisure/\ sport/culture$" and “$other\ activities$”;
- AT4 - Regular and unavoidable activities, such as “$drop\ off/\ pick\ up$" and “eat”;
- AT5 - Activities often located in town centres, such as “$shopping$”, “$School$" and “$Personal\ business$”

Other groups of activities can be identified in accordance to the same dataset, which can be used in the following for estimating activities within zones. Due to geographical interpretation of the clusters, this is not the case here and tests are carried out in order to identify the above activity types.

3.2. Results

3.2.1. Crude Monte Carlo

To determine to what extent the methodology can be used in a dynamic application, Figure (1a) shows the scatterplot of proportions of trip-purposes for each hour of the day for two simulations as compared to the survey data. In the first case, the period-specific distribution among trip-purposes is known with precision and specifically for twenty-four hours and in the second case, the same daily distribution, regardless of the time, is applied to all trips. This results in every hour of the day having the exact same distribution (each of the five horizontal lines on Figure (1a) corresponding to a specific trip purpose). Hence, this methodology can give extremely precise results (R^2=1) in case of static analysis or of detailed input information, whereas it becomes poor for short reference periods.

For the following results, the distribution among five activity types is given with an hourly precision and the activity-specific generation factor of zones is calculated for the complete day. Figure (1b) shows the generated production factors with respect to each activity type. It can be seen that the combination of varying distributions and number of trips departing from a zone at a certain time can lead to a observable activity-specification of zones. The reliability of these results compared to the sample data is confirmed by the scatterplot of Figure (1c). The simulation can well approximate the distribution of the purpose-specific trips for each zone. For example, the identified residential area of Ghent proves to be a major generating zone for the trip-purpose "*work*" while it has a very much lower weight in trips destined to return home.

One step further in the analysis of the results is the ability of the Monte Carlo simulation to detect variations in the generation for each activity type, during three time periods. In light of the survey data, this value is presented for the activity type "*Work*" as an illustrative example in Fig. 2.

a. *b.* *c.*

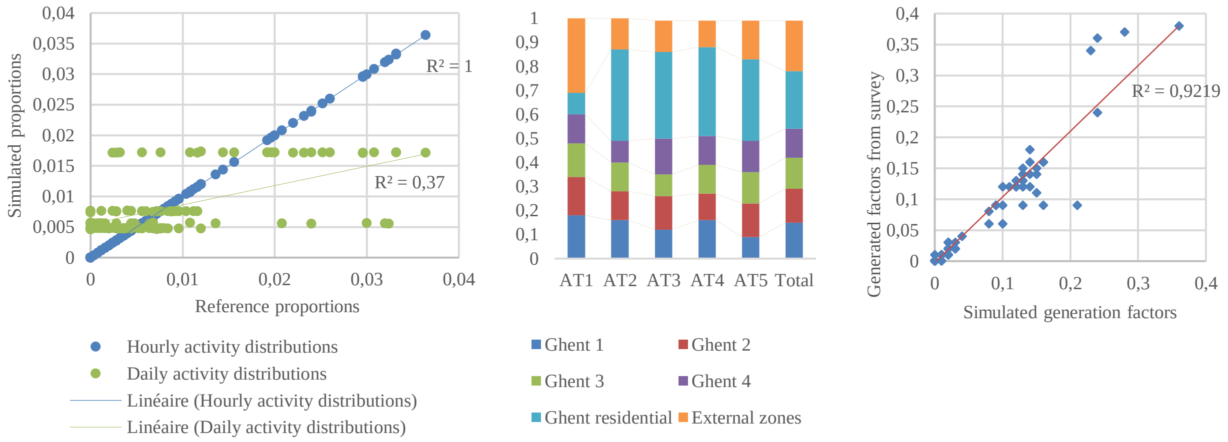

Fig. 1. (a) Scatter proportion of activities by hour of the day for two precision levels; (b) Simulated generation factor for five activity types; (c) Scatterplot of simulated generation factors vs. survey data.

Table 1. Activity proportions and number of trips by departure zone.

Zone	AT1	AT2	AT3	AT4	AT5	Total Number of trips
Ghent residential	37%	19%	12%	14%	19%	2815
Ghent 4	44%	11%	12%	11%	21%	1413
Ghent 3	45%	12%	13%	12%	19%	1477
Ghent 2	44%	12%	11%	12%	21%	1644
Ghent 1	42%	14%	11%	12%	21%	1728
External	51%	8%	13%	10%	18%	2441

Again, a first observation is that the model can evaluate various generation factors for different times of the day. In comparison with exact data, we observe that the global aspect is similar even though some underestimation of small zones is observed. Another approximate estimation is the distribution in the evening, when only 2% of the trips are performed to go to work, in opposition to distribution in the morning, prominent period for going to work, which is precisely transposed in the simulations results.

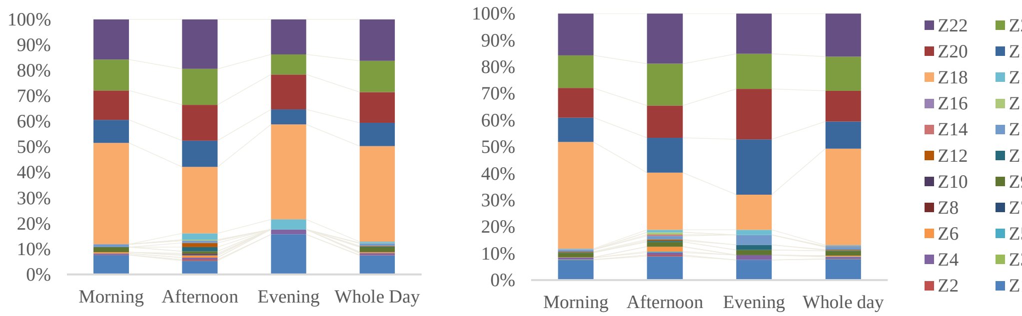

Fig. 2. Generation factor for activity *work* for three time periods (a) From the simulation; (b) From the survey.

In general, the simulation appears to give particularly good results in the case of large number of trips, and less detailed categories e.g. three time periods rather than hours, internal and external zones rather than individual zones. Yet for these results an important amount of information is needed as input of the simulation. Still if the information is more disaggregated or additional consideration is added to the attribute allocation, a better approximation can be expected.

3.2.2. Markov Chain Monte Carlo

In a second phase, the Markov Chain Monte Carlo has been used to determine the primitives of the signal of trips departing from the largest zone outside Ghent, i.e. zone one. 1120 trips were used as input, and departure times were used to determine two parameters of five functions characterising the activity types. The total distribution between activities is assumed fixed, based on the complete survey proportions. Analysis of the quality of the results is then based on the comparison of appearance between the real demand and the estimated demand for the traffic zone.

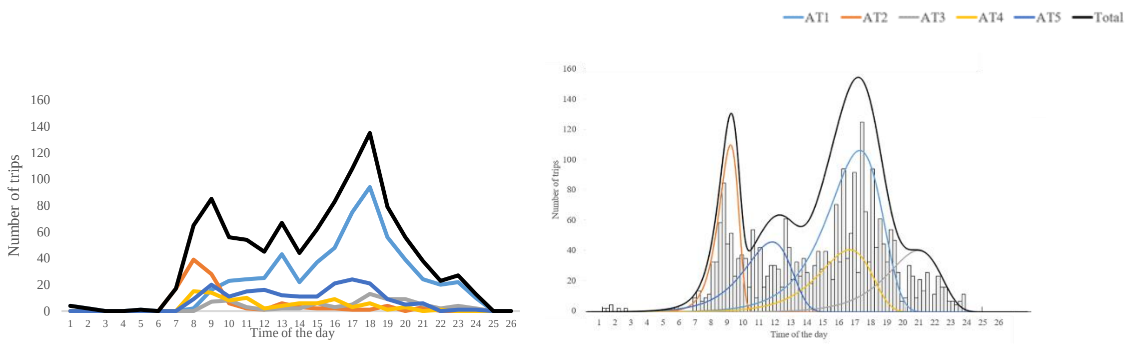

Fig. 3. Departure profile by activity for zone 1 (a) From the survey; (b) From the simulation.

Comparing these two figures, it appears that the proposed methodology identifies activity patterns of the zone in their general shapes. In particular, peaks are correlated to the appropriate trip-purposes: the most evident are indeed "*work*" in the morning and "*going home*" in the afternoon. However, the model recreates the peak at lunch time and in the evening, but without being able to accurately determine the other activity components. Two reasons explain this issue; not only fewer observations exist for these activities but also the choice of the functional form of the function is not as suited as for the peaks. Another test for the simulation quality is the comparison between the estimated location parameter of each evaluated distribution and the typical departure time of each activity calculated from the database. Again, the two main activity types (Home and Work) are correctly characterized, whereas the three secondary purposes compose peaks at regular intervals during the day, without specific relationship to the real meaning of trip-purpose. It is important to note that the procedure treats without much difference these activity-types: same prior, almost same percentage of users, which can explain the inaccuracy to distinguish them. In sum, this experiment shows that a MCMC-based methodology is promising and allows determining activity-based priors, giving acceptable results for the most characteristic activity-types. Nevertheless, more precision in the selection of likelihood form and activity clustering would be necessary for a better recognition of the complete demand.

4. Conclusions

In this paper, we presented a novel procedure to estimate time-dependent purpose-specific flows with respect to the zone of departure of a journey and without information at the user-level. Two tools based on Monte Carlo techniques are proposed for evaluating static activity-specific generation factors dynamic classification of flows, with respect to trip-purpose. For these simulations, flows are disaggregated and used as evidences for the calibration of an unknown distribution, with all the available information they can contain. Assumptions might be added to the models in order to improve their reliability. The validity of the models gets subject to particular care in case of a large number of parameters and requires a high amount of iterations of the simulation to get a stable result. It is still generally able to recognize typical characteristics, such as morning commute, but the link to specific activity can be concluded over a second phase. The components representing activities without hallmarks and less occurrences are indeed less easily identifiable. A drawback of both models is that the quality of representation and evaluation of less common cases. In particular, small zones, off-peak hours as well as activities with fewer observations. Nevertheless, by choosing more constraints and selecting each parameter with a great precision a priori, the model itself can also become more reliable. However, this means the handling of one more parameter and so, an increase in the complexity. The proper distributions are function of the possible knowledge and assumption that can be made on the behavioural components, each activity can possibly be represented by a different kind of distribution. Because of the flexibility of these methods, the models can be used for more attributes, more assumptions to force the distributions having the expected properties, and using various evidences. In the worst-case scenario, the current model would not be applicable if no distribution can provide a good fit of the data.

References

Antoniou, C., et al.., 2016. Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation. Transp. Res. Part C l 66, 79–98. doi:10.1016/j.trc.2015.08.009

Antoniou, et al.., 2015. W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models. Transp. Res. Part C, Special Issue on International Symposium on Transportation and Traffic Theory 59, 129–146. doi:10.1016/j.trc.2015.04.030

Arentze, T., Timmermans, H., Hofman, F., 2007. Creating Synthetic Household Populations: Problems and Approach. Transp. Res. Rec. J. Transp. Res. Board 2014, 85–91. doi:10.3141/2014-11

Arentze, T.A., Timmermans, H.J.P., 2004. A learning-based transportation oriented simulation system. Transp. Res. Part B Methodol. 38, 613–633. doi:10.1016/j.trb.2002.10.001

Barthelemy, J., Toint, P.L., 2013. Synthetic Population Generation Without a Sample. Transp. Sci. 47, 266–279. doi:10.1287/trsc.1120.0408

Beckman, R.J., Baggerly, K.A., McKay, M.D., 1996. Creating synthetic baseline populations. Transp. Res. Part Policy Pract. 30, 415–429. doi:10.1016/0965-8564(96)00004-3

Bhat, C., 2002. Random Utility-Based Discrete Choice Models for Travel Demand Analysis, in: Transportation Systems Planning, New Directions in Civil Engineering. CRC Press. doi:10.1201/9781420042283.ch10

Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. Part Policy Pract. 35, 1–28. doi:10.1016/S0965-8564(99)00043-9

Cascetta, E., Inaudi, D., Marquis, G., 1993. Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts. Transp. Sci. 27, 363–373.

Castaigne, M., Cornelis, E., Frederix, R., Tampere, C., Toint, P., Viti, F., Walle, F., 2009. Behaviour and mobility within the week BMW. Final Rep. Belg. Sci. Policy Res. Programme Sci. Sustain. Dev. Bruss. 120.

Duguay, G., Jung, W., McFadden, D., 1976. SYNSAM: A Methodology for Synthesizing Household Transportation Survey Data. Urban Travel Demand Forecasting Project, Institute of Transportation Studies.

Ettema, D., Timmermans, H., 2003. Modeling Departure Time Choice in the Context of Activity Scheduling Behavior. Transp. Res. Rec. J. Transp. Res. Board 1831, 39–46. doi:10.3141/1831-05

Farooq, B., Bierlaire, M., Hurtubia, R., Flötteröd, G., 2013. Simulation based population synthesis. Transp. Res. Part B Methodol. 58, 243–263.

Flötteröd, G., 2009. Cadyts - a free calibration tool for dynamic traffic simulations. Presented at the 9th Swiss Transport Research Conference.

Flötteröd, G., Chen, Y., Rieser, M., Nagel, K., 2009. Behavioral calibration of a large-scale travel behavior microsimulation, in: 12 International Conference on Travel Behaviour Research.

Gargiulo, F., Ternes, S., Huet, S., Deffuant, G., 2010. An Iterative Approach for Generating Statistically Realistic Populations of Households. PLOS ONE 5, e8828. doi:10.1371/journal.pone.0008828

Lenormand, M., Deffuant, G., 2012. Generating a synthetic population of individuals in households: Sample-free vs sample-based methods. ArXiv12086403 Math Stat.

Lim, Y., Kim, H., 2016. A combined model of trip distribution and route choice problem. Transp. Transp. Sci. 12, 721–735. doi:10.1080/23249935.2016.1166171

Manski, C.F., 1977. The structure of random utility models. Theory Decis. 8, 229–254. doi:10.1007/BF00133443

Michael G. McNally, 2007. The Four-Step Model, in: Handbook of Transport Modelling, Handbooks in Transport. Emerald Group Publishing Limited, pp. 35–53. doi:10.1108/9780857245670-003

Mueller, K., Axhausen, K.W., 2011. Hierarchical IPF: Generating a synthetic population for Switzerland. Louvain-la-Neuve: (ERSA).

Raux, C., Ma, T.-Y., Cornelis, E., 2016. Variability in daily activity-travel patterns: the case of a one-week travel diary. Eur. Transp. Res. Rev. 8, 26. doi:10.1007/s12544-016-0213-9

Toole, J.L.,et al., 2015. The path most traveled: Travel demand estimation using big data resources. Transp. Res. Part C Emerg. Technol., Big Data in Transportation and Traffic Engineering 58, Part B, 162–177. doi:10.1016/j.trc.2015.04.022

Voas, D., Williamson, P., 2000. An evaluation of the combinatorial optimisation approach to the creation of synthetic microdata.

Vrtic, M., Fröhlich, P., Schüssler, N., Axhausen, K.W., Lohse, D., Schiller, C., Teichert, H., 2007. Two-dimensionally constrained disaggregate trip generation, distribution and mode choice model: Theory and application for a Swiss national model. Transp. Res. Part Policy Pract., August 2005 41, 857–873. doi:10.1016/j.tra.2006.10.003

Ye, X.,et al.., 2009. Methodology to Match Distributions of Both Household and Person Attributes in Generation of Synthetic Populations. Presented at the Transportation Research Board 88th Annual MeetingTransportation Research Board.