Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 151 (2019) 870–877

Procedia Computer Science

www.elsevier.com/locate/procedia

The 8th International Workshop on Agent-based Mobility, Traffic and Transportation Models, April 29 - May 2, 2019, Leuven, Belgium

The MATSim Open Berlin Scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data

Dominik Ziemke[a,*], Ihab Kaddoura[a], Kai Nagel[a]

[a]*Technische Universität Berlin, Transport Systems Planning and Transport Telematics, Salzufer 17-19, 10587 Berlin, Germany*

Abstract

With more diverse transport policies being proposed and the advent of novel transport services and technologies, the transport system is becoming more individualized in many aspects. Transport models, the most important tool to assess policies and schemes, need to be sufficiently expressive to address these developments. Agent-based transport models, where travelers with individual properties and the ability to act and decide autonomously are resolved individually, allow to appropriately model and analyze such policies. This paper describes the *MATSim Open Berlin Scenario*, a transport simulation scenario for the Berlin metropolitan area implemented in the agent-based transport simulation framework MATSim. The scenario is solely based on open data and the demand for transport is created based on a fully synthetic procedure. Contrary to most transport simulation scenarios, no information from a travel diary survey is required as input. As such, the scenario generation procedure described in this study is spatially transferable and facilitates the creation of agent-based transport simulation scenarios for arbitrary regions.

Keywords: transport modeling; agent-based simulation; synthetic demand generation; open data

1. Introduction

Transport models are the most important tool to assess transport policies and schemes, to forecast their outcomes and to assess their effectiveness. In contrast to earlier decades where transport policies mainly consisted in infrastructure investment schemes, today more diverse solutions, largely based on transport demand management, are sought to balance the needs imposed on urban systems like high levels of mobility, improved sustainability, and livability. Policies intended to meet these ends include, for instance, improvements of the public transport system, vehicle shar-

* Corresponding author. Tel.: +49-30-314-21383 ; fax: +49-30-314-26269.
E-mail address: ziemke@vsp.tu-berlin.de

10.1016/j.procs.2019.04.120

ing systems, ride-pooling services, transit-oriented development, congestion or peak-hour pricing, emission pricing, parking management, and environmental zones.

To be able to assess such policies, models that allow to analyze travelers' reactions to these policies are required. Agent-based transport models, where individual people with individual properties and the ability to act and decide autonomously are individually resolved, allow to appropriately model and analyze such policies. The present paper describes the *MATSim Open Berlin Scenario*, a transport simulation scenario for the Berlin metropolitan area implemented in the agent-based transport simulation framework MATSim. The scenario contains full day plans of agents who represent all adult people living in the states of Berlin and Brandenburg, and is set-up such that it represents the traffic in the state of Berlin realistically. It encompasses all relevant modes of transport, i.e. private car as driver, private car as passenger, bicycle, walking, and public transport, the latter being simulated based on the actual public transport schedule. It relies on openly available and unrestrictedly usable data ('open data'), mostly in standardized formats, to ensure spatial transferability of the scenario generation procedure. It is set-up to be policy-sensitive in terms of route, mode, and departure time choice.

The development of the transport simulation scenario is based on a microscopic (i.e. person-specific), econometric activity scheduling model and a traffic-count-based automated calibration procedure. The underlying development concept has been described by Ziemke et al. [19], however in the context of a scenario that only covers car traffic, while this study describes a thoroughly updated version containing all relevant modes of transport.

The described scenario generation approach based on openly available input data is related to the approach by Kickhöfer et al. [12] for Santiago de Chile. The Santiago scenario relies on a region-specific travel survey, a type of dataset that is not available in many regions and even if, mostly hard to obtain and procure, e.g. for data privacy reasons. In contrast, the scenario generation approach described in the present paper, is solely based on data which are almost universally regarded as open data and that may be obtained comparatively easily in most places of the world and be procured without significant restrictions. As such, the development procedure of the *MATSim Open Berlin scenario* is readily transferable to other regions, as exemplified by a scenario for the Ruhr region in Germany, which is based on this scenario's development procedure [20].

2. Methodology

The generation of the *MATSim Open Berlin scenario* consists of the creation of a synthetic population of the study area and the prepartion of other input information for the activity-scheduling model CEMDAP, the application of CEMDAP (multiple times) to create a set of initial daily activity-travel patterns for each agent, a calibration-simulation run of MATSim in conjunction with the CaDyTS calibration procedure to obtain suitable (mainly in terms of activity locations) plans from the initial plan set created by CEMDAP, and a subsequent calibration of transport mode choice.

The following data sources are used: 'Zensus 2011', a nation-wide census of Germany [16], 'Pendlerstatistik 2009' (commuter statistics) [5], OpenStreetMap [15], local GTFS [17] data, local traffic counts, BASt counts on freight traffic[6], shapefiles that describe the municipality geometries in Brandenburg and LOR[1] geometries in Berlin, and CORINE land cover [7] data. For validation of various properties of travel, the Berlin SrV 2008 [1] and Berlin MiD 2008 [11] travel surveys are used.

2.1. Generation of a synthetic population and other input information for the activity-scheduling model CEMDAP

The population of the scenario consists of all persons aged 18 and above who reside in the German states of Berlin and Brandenburg according to the 'Zensus 2011' [16]. For each municipality, the census file contains the total population, differentiated by gender and eleven age classes, the number of employees, differentiated by gender, the number of students, and additional information. Workplaces of employees are informed by the 'Pendlerstatistik 2009' commuter statistics [5]. Its files contain the number of socially-secured employees, differentiated by gender, for each residential-municipality/employment-municipality pair. As Berlin is only one, albeit large and populous, municipality, this approach requires further refinement for agents residing and/or working in Berlin, as will be discussed later. For

[1] LOR stands for "Lebensweltlich orientierte Räume" and is a neighborhood-oriented zone system.

the surrounding federal state of Brandenburg, with mostly smaller municipalities, no further refinement is necessary. While the commuter statistics only inform about the numbers of socially-secured employees, the census contains all employees. To match these informations, a scaling procedure is applied, which scales municipality-to-municipality commuter numbers such that the number of all commuters (as scaled up from the socially-secured commuters of the commuter statistics) residing in a zone matches the number of employed residents of that zone (as by the census). This procedure is applied for all age-and-gender-specific subgroups of the population separately to ensure a sufficient level of accuracy.

Based on this, the procedure iterates over all municipalities and creates in each municipality persons, assigning them age and gender according to the census. For the municipality of Berlin, one of Berlin's 447 LORs (neighborhood-oriented districts) is chosen randomly as their home location. Since LORs are defined in such a way that the population of a LOR does not fall below or exceed a certain minimum or maximum [4], this procedure approximates population densities in Berlin. Agents are assigned to be employed or to be a student in accordance with the relevant numbers from the census. Work or school location zones are assigned according to the commuter information. If the zone is Berlin, a random LOR is chosen for this location.

This information is written out into a persons file in the format required by the activity-scheduling model CEMDAP [3]. In the same step, one household is created for each agent, i.e. household structures are not taken into account and the household file is written out. Children (people under the age of 18 years) are not included.

Furthermore, CEMDAP requires information on transport levels of service and land use. Travel times and travel costs between any two zones, contained in the level-of-service file, are computed based on Euclidean distances between zone centroids, using shapefiles and speed assumptions. CEMDAP's zones file can contain information regarding numbers of different types of employees in the zone and some accessibility measures for each zone. In the setup of this study, no such information is used and the values of specific properties of the zones are set to dummy values, rendering each zone (apart from their location) equally attractive. The reason for this approach is that location choice is not finalized in CEMDAP, but in a later stage taking advantage of MATSim's co-evolutionary algorithm in interplay with a calibration procedure (cf. Sec. 2.3).

2.2. Application of CEMDAP to create multiple initial daily activity-travel pattern for each agent

Next, CEMDAP (Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns) [2], a software implementation of a system of random-utility-based models that represent the decision-making behavior of individuals, is used to assign daily activity-travel patterns for each member of the synthetic population[2]. CEMDAP's output consists of the complete daily activity-travel patterns of each individual of the synthetic population and describes the sequence of activities and intervening trips that a person undertakes during the day.

The software is applied using a model parameter set ('model specification file') that is taken from CEMDAP's latest implementation for the Los Angeles metropolitan areas in the United States. On the one hand, this procedure is associated with the assumption that persons of the model estimation context (Los Angeles) with certain demographic attributes do not have highly different preferences as persons with the same demographic attributes in the model application context (Berlin). On the other hand, this does *not* imply that the approach would neglect differences between the two regions: First, the obvious fact that the synthetic population and the geography of the two regions are different has already been accounted for as described in section 2.1. Second, in contrast to a standard use of CEMDAP (i.e. when applied in a regional context for which an explicit model estimation has been carried out), the modeling results are not directly and unalteredly used for analysis. Instead, CEMDAP is run multiple times (specifically, five times) to generate a *selection* of initial activity-travel patterns for each agent. All these patterns are fed into the MATSim transport simulation [10], which, in correspondence with a calibration procedure (cf. Sec. 2.3), sorts out those activity-travel patterns that do not contribute sufficiently well to a reproduction of real-world traffic patterns as they are given on the basis of traffic counts.

In a preceding study [19], different methods of model transfer have been described and it has been explained that an update of model parameters is generally advisable and found to lead to better results than approaches without

[2] For technical details regarding the application of CEMDAP, the reader is referred to Bhat et al. [3]. For technical information concerning the transfer of information between CEMDAP and MATSim, see Ziemke and Nagel [18].

updating. While in other model transfer studies updating has mostly been done with regard to parts of the model parameters, in the present approach, model updating operates on initial full daily activity plans, while the CEMDAP model parameters themselves remain unchanged. This updating on full daily activity plans is carried out by MATSim's co-evolutionary algorithm (cf. sections 2.3 and 2.4), while the activity-travel patterns generated by CEMDAP are solely used as start solutions for agents' plans.

2.3. Calibration of activity locations using MATSim in conjunction with the CaDyTS calibration procedure

As a result of the multifold application of CEMDAP (cf. Sec. 2.2), multiple initial daily activity-travel patterns are obtained for each member of the synthetic population. The task of this modeling step is to choose those patterns that best represent real-world travel observations. Since the main difference among the different initial activity-travel patterns consists in different activity locations, this constitutes a model of *location choice*. For computational resource efficiency, a 10% sample of the full population is created.

In MATSim, microscopic home and work locations are required. They are chosen by a random draw of a coordinate within the geographies of the home or work zones, respectively. A coordinate is only kept if it falls within an area of land use where the corresponding activity seems likely (e.g. home activities only in areas where residential buildings are allowed). This is informed by CORINE ('Coordination of information on the environment') land cover data [7], a standardized data collection on land cover in Europe. The home location is kept constant over the various plans of a given agent. The network for the transport simulation is based on OpenStreetMap [15] data and accounts for all road categories within the Berlin city boundaries and all main roads in the surrounding state of Brandenburg (cf. Fig. 6a). The network consists of 73,689 nodes and 159,039 single-direction car-only links.

Real-world traffic observations are given by 8,304 hourly count values for 346 count stations. 250 of these count stations are operated by the Berlin Traffic Management Center (Verkehrsmanagementzentrale), while the remaining 96 stations belong to the motorway administration. In the latter, there is no distinction by vehicle type such that it was decided to calibrate against the undistinguished counts of all vehicles for all count stations. To ensure consistency with travel demand segments, freight traffic, therefore, needs to be represented in the simulation as well. Therefore, a simplified representation of freight agents was added to the scenario based on BASt [6] freight counts for all major arterials leading to and from Berlin. The freight agents' trip origins and destinations are fixed and excluded from the location choice calibration. Note that counts data constitute the only component of all input data which are not openly available, but have to be acquired. As such data are, however, available in the majority of regions, even in less-developed regions where such data can be provided by means of manual counts, the inclusion of this dataset is not regarded as a major hurdle to model transferability.

To calibrate an agent's plan choice set against traffic counts, MATSim is run in conjunction with CaDyTS (Calibration of dynamic traffic simulations) [8]. This procedure has been described in more detail by Ziemke et al. [19]. In contrast to that work, however, travel demand is *not* scaled down to the share of car travelers since the goal is to create a scenario that considers travelers of *all* modes. Instead, transport network properties are scaled up by the inverse of the Berlin-specific modal share for cars (i.e. to a value of 0.31 instead of 0.10, which would be the usual for a scenario with 10% sample size). Therefore, all travelers travel on the car network in this modeling step. Because of the up-scaling of the network, the traffic dynamics of all tripmakers traveling on the car network are still approximately the same as if only the car travelers moved on the original network.

This allows the calibration procedure to exert influence on the choice of plans of all agents, and not just car travelers as it was done previously [19]. It involves the implicit assumption that spatial patterns of travelers in different modes of transport are similar, which is obviously not the case. This initial inconsistency is, however, remedied in the subsequent modeling step (cf. Sec. 2.4), where agents are allowed to change their modes of transport.

The calibration effect of CaDyTS is incorporated into MATSim via the scoring of plans [13, 14] in addition to the default behavioral scoring of activity participation and travel. CaDyTS's calibration effect acts as an additional scoring component, which evaluates how well a particular plan of an agent matches with real-world traffic observations. If a plan is conducive to reproducing real-world traffic patterns, CaDyTS will reward this plan with a positive score offset, while a plan that produces trips that are not in line with real-world observations will receive a negative score offset. This offset should be interpreted as an alternative-specific constant of the choice model, where CaDyTS is the method to determine its value and the different plans constitute the alternatives. As this procedure is integrated into MATSim's iterative transport demand adaptation process, travel demand becomes more realistic over the course of iterations.

In order to obtain realistic values of travel time savings (VTTS), activities are differentiated based on their typical durations, e.g. 'shopping' is split into different activity types for pre-defined bins of typical durations [13], e.g. 'shopping_600" for a shopping activity of up to 600 s, or 'shopping_1200' for a shopping activity of 600 to 1200 s, and so on.

2.4. Calibration of the choice of transport modes

A subsequent manual scenario calibration is carried out using the selected output plans of the previous step (cf. Sec. 2.3) as input plans, together with the original network (without up-scaling). An openly available GTFS dataset for the Berlin-Brandenburg region [17] is used to generate MATSim public transport schedule and vehicles files and to add public transport links to the network. In contrast to the previous calibration step, this calibration is done manually and without any influence of CaDyTS. That is, the agents' scores solely depend on the generalized time- and distance-related travel cost as well as the time spent performing activities. Person agents are enabled to change their modes of transport, departure times, and routes. Besides 'car' (as driver), the available **modes of transport** include 'ride' (i.e. car as passenger), 'public transport', 'bicycle', and 'walk'. In the current scenario, all modes besides car and public transport are so-called teleported modes, i.e. trips made by these modes are handled by putting the agent to the location of the next activity with a time lag. For walk and bicycle, the time lag corresponds to the beeline distance and a mode-specific speed (cf. Tab. 2). For the ride mode, the agents are routed on the network and the teleportation time is obtained from the (congested) travel time of the car mode along that route. During the calibration process, the ride mode share was fixed and agents are not allowed to switch from or to the 'ride' mode. For trip chains starting and ending at the same activity location, the transport mode may be changed to car or bicycle (i.e. chain-based modes) for the whole tour or a combination of public transit and walk. In addition, agents are enabled to randomly shift their **departure times** within a range of 2 hours. For activities with durations larger than 2 hours, the end time is adjusted and for activities with durations smaller than 2 hours, the activity duration is changed. Furthermore, agents are enabled to adjust their **routes**. Freight agents are only enabled to change their route. To obtain more realistic free-speed travel times, the free speeds of inner-city street from OpenStreetMap are reduced by its half in order to account for traffic lights, acceleration and deceleration at intersections etc. An explicit modeling of traffic signals is possible via a corresponding extension [9] if data on signals of the study are accessible, which is currently not the case in Berlin. Simulation settings are stated in Tab. 1, where innovative strategies include transport mode choice, departure time choice, and route choice. The mode-specific parameters, which have been adjusted in this calibration step and constitute the result of this step, as well as other predefined mode-specific settings are given in Tab. 2. For all scoring function parameters not shown in Tab. 1, MATSim's corresponding default values are used [13].

Table 1: General simulation parameters in the mode-choice calibration step.

Parameter	Value
Sample size	10 %
Flow / Storage capacity factors	0.10
Weight of innovative strategies	0.05

Table 2: Mode parameters.

Mode	Alt.-spec. constant	Monetary dist. rate	Teleportation speed
Car	-1.5	-0.0002 EUR/m	-
Public transport	-0.6	-	-
Bicycle	-1.85	-	12.0 km/h
Walk	0.0	-	4.0 km/h

3. Results and discussion

Fig. 1 and Fig. 2 depicts the relation of simulated and real-world traffic volumes over all count stations. Fig. 1 shows the outcome after the first calibration step (location choice, cf. Sec. 2.3), while Fig. 2 shows the outcome after completing the second calibration step (mode choice, cf. Sec. 2.4). Since in the second calibration step, no CaDyTS calibration is active, this step also serves as a stability test (i.e. a simulation run that confirms that choices made initially under the influence of CaDyTS stay largely the same after this influence is removed) as it has been discussed previously [19].

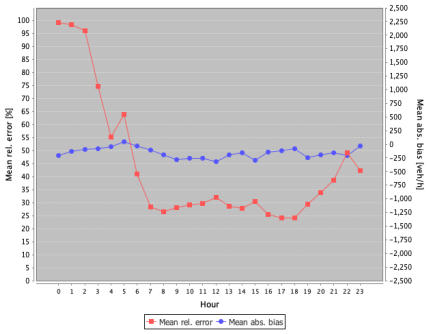


Fig. 1: Traffic counts match after the calibration of activity locations (cf. Sec. 2.3).

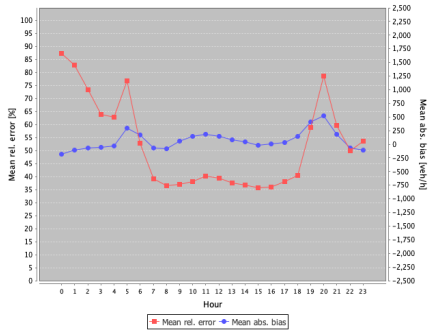


Fig. 2: Traffic counts match after the calibration of transport modes (cf. Sec. 2.4).

Tab. 3 provides the simulated mode shares after completing the mode choice calibration as well as the mode shares given in the Berlin travel surveys for the year 2008. For calibration step 2 (cf. Sec. 2.4), the aggregated modal split was compared with the the MiD 2008 [11] data; numbers obtained from the SrV 2008 [1] survey differ by some extent.

Table 3: Comparison of modes shares between simulation and surveys.

Mode of Transportation	**Mode Share**		
	Simulation (berlin-v5.2-10pct)	MiD 2008 survey [11]	SrV 2008 survey [1]
Car	29.6%	30.0%	24.8%
Public transport	22.4%	21.0%	26.1%
Ride	10.0%	10.0%	7.68%
Bicycle	10.6%	11.0%	12.82%
Walk	27.5%	28.0%	28.58%

Fig. 3 depicts properties of trips made by agents traveling by car compared to corresponding values of the SrV 2008 travel survey [1]. On the simulation side, only those agents who reside in Berlin are taken into account so that results are comparable to the survey which only covers this area. Both for simulation and survey, only trips shorter than 100 km are considered. Travel distances and durations match very well. Also the share of activity types at trip ends is

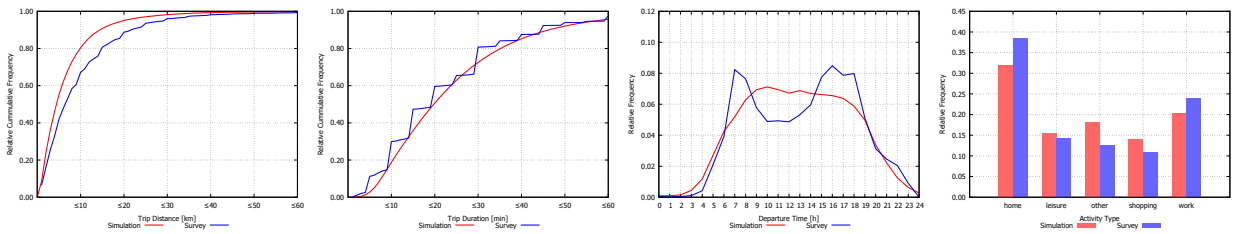
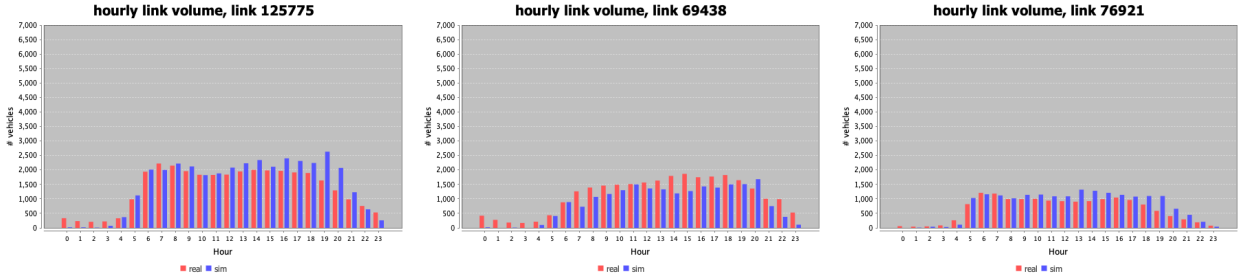


Fig. 3: Comparison of (beeline) distance, duration, and departure time distributions and shares of activity types at trip ends by *car* in simulation and survey.

reproduced sufficiently. Compared to the survey, the simulation produces more midday traffic. This has already been observed in a previous car-only version of this scenario [19]. As explained in Sec. 2.3, the simulation and the counts against which it has been calibrated (cf. Sec. 2.3) contain freight traffic, whereas the survey only contains personal traffic with the typical morning and afternoon peaks, which is likely the main reason for the visible deviation. Also, the counts contain service traffic (e.g. craftspeople, mobile nursing service etc.) that the survey does not cover.

Fig. 4 depicts a comparison between simulated and counted traffic volumes at three arbitrarily selected individual counts stations in different parts of Berlin. While the overall quality of the match for the count stations is good, it is



(a) Northern inner motorway ring (A100). (b) Leipziger Straße in Mitte. (c) Märkische Allee (L158) in Marzahn.

Fig. 4: Comparison of simulated and counted traffic volumes of three individual count stations.

also visible that there is no marked midday drop as found in the survey data (cf. Fig. 3), reaffirming the above remarks. Fig. 5 contains a comparison for trips made by public transport, showing good results for all compared properties. The steps in the survey graph of the trip duration diagram are a result of self-reported travel times in the survey where respondents tend to report 'round' numbers rather than exact values by the minute.

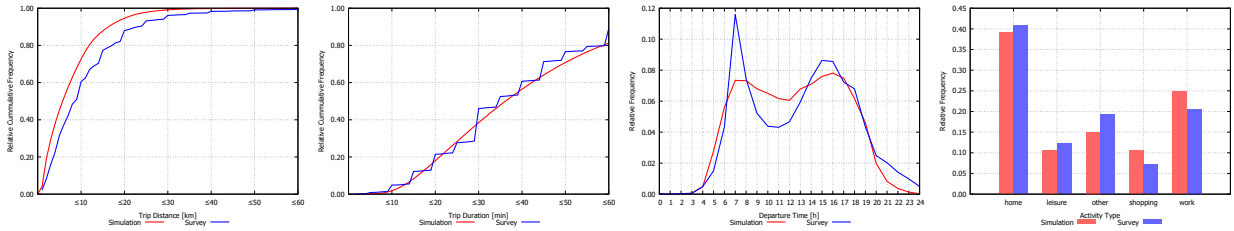
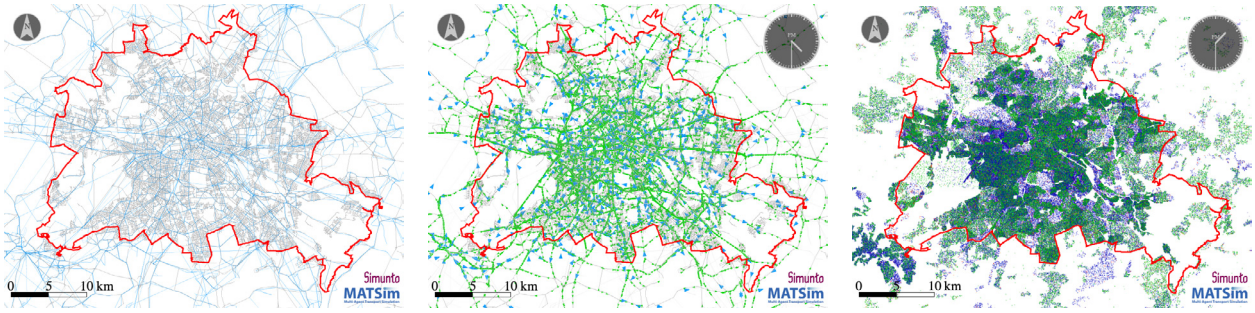


Fig. 5: Comparison of (beeline) distance, duration, and departure time distributions and shares of activity types at trip ends by *public transport* in simulation and survey.

Fig. 6 visualizes the network, the traffic situation in the afternoon, and agents' activities in the early afternoon of the scenario. Tab. 4 compares trip distances and durations by different modes between simulation and survey.



(a) Car network (grey) and public transit network (blue). (b) Cars (green) and public transit vehicles (blue). (c) Home (green) and work (blue) activities.

Fig. 6: Visualization of the network, vehicles, and activity locations.

With this study, it has been shown that it is possible to create a MATSim transport simulation scenario that represents real-world traffic of a large metropolitan region realistically and, at the same time, is solely based on open data and does not require information from a travel diary survey as input. As such, the scenario generation procedure described in this study is spatially transferable and facilitates the creation of agent-based transport simulation scenarios for arbitrary regions.

Table 4: Comparison between simulation and SrV 2008 survey.

Mode of Transport	Trip Distance		Trip Duration	
	Simulation	Survey	Simulation	Survey
Car (incl. freight)	9.2 km	9.6 km	26.9 min	22.3 min
Car	9.1 km	9.6 km	26.7 min	22.3 min
Public transport	9.1 km	11.7 km	45.1 min	40.2 min
Bicycle	2.6 km	3.6 km	17.3 min	17.6 min
Walk	0.7 km	1.0 km	8.9 min	14.0 min

4. Accessing the scenario and acknowledgement

Code and data to run the scenario and based on which own functionality can be added, is available online under https://github.com/matsim-vsp/matsim-berlin. The version documented and analyzed in this paper is the 10% scenario with release number 5.2 ('berlin-v5.2-10pct). All required input files and the ready scenario are accessible via https://svn.vsp.tu-berlin.de/repos/public-svn/matsim/scenarios/countries/de/berlin/berlin-v5.2-10pct/. The authors wish to thank Gabriel Thunig for technical work on analyses and Tilmann Schlenther who contributed automatizing the mapping between traffic counts and corresponding road segments. The rest of the VSP team, in particular Amit Agarwal, provided valuable feedback throughout the development process.

References

[1] Gerd-Axel Ahrens. Endbericht zur Verkehrserhebung Mobilität in Städten – SrV 2008 in Berlin. Technical report, Technische Universität Dresden, 2009. http://www.stadtentwicklung.berlin.de/verkehr/politik_planung/zahlen_fakten/download/2_SrV_endbericht_tudresden_2008_berlin.pdf.

[2] C.R. Bhat, J.Y. Guo, S. Srinivasan, and A. Sivakumar. A comprehensive econometric microsimulator for daily activity-travel patterns. *Transportation Research Record*, 1894:57–66, 2004. ISSN 0361-1981. doi: 10.3141/1894-07.

[3] C.R. Bhat, J. Guo, S. Srinivasan, and A. Sivakumar. *CEMDAP User' s Manual*. University of Texas at Austin, Austin, TX, USA, 2012.

[4] H. Bömermann, S. Jahn, and K. Nelius. Lebensweltlich orientierte Räume im Regionalen Bezugssystem (Teil 1). *Berliner Statistik*, 8:366–371, 2006. http://www.stadtentwicklung.berlin.de/planen/basisdaten_stadtentwicklung/lor/download/BerlinerStatistik0608.pdf.

[5] Bundesagentur für Arbeit. Pendlerstatistik 2010. CD-ROM, 2010.

[6] Bundesanstalt für Straßenwesen (BASt). Automatische Zählstellen 2016, accessed in February 2018. https://www.bast.de/.

[7] Copernicus Land Monitoring Service. CORINE Land Cover 2012. last accessed on 09 January 2019, 2012. https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012.

[8] G. Flötteröd. *Cadyts – Calibration of dynamic traffic simulations – Version 1.1.0 manual*. Transport and Mobility Laboratory, EPFL Lausanne, November 2010. URL http://home.abe.kth.se/~gunnarfl/files/cadyts/Cadyts_manual_1-1-0.pdf.

[9] D. Grether and T. Thunig. Traffic signals and lanes. In Horni et al. [10], chapter 12. doi: 10.5334/baw.

[10] A. Horni, K. Nagel, and K. W. Axhausen, editors. *The Multi-Agent Transport Simulation MATSim*. Ubiquity, London, 2016. doi: 10.5334/baw.

[11] Infas and DLR. Mobilität in Deutschland 2008 Ergebnisbericht, 2010.

[12] B. Kickhöfer, D. Hosse, K. Turner, and A. Tirachini. Creating an open MATSim scenario from open data: The case of Santiago de Chile. VSP Working Paper 16-02, TU Berlin, Transport Systems Planning and Transport Telematics, 2016. URL http://www.vsp.tu-berlin.de/publications.

[13] K. Nagel, B. Kickhöfer, A. Horni, and D. Charypar. A closer look at scoring. In Horni et al. [10], chapter 3. doi: 10.5334/baw.

[14] K. Nagel, M. Zilske, and G. Flötteröd. CaDyTS: Calibration of Dynamic Traffic Simulations. In Horni et al. [10], chapter 32. doi: 10.5334/baw.

[15] OpenStreetMap, last accessed on 25 June 2018. http://www.openstreetmap.org.

[16] Statistische Ämter des Bundes und der Länder. Zensus 2011, accessed 21 July 2017. www.zensus2011.de.

[17] Verkehrsverbund Berlin-Brandenburg. VBB-Fahrplandaten via GTFS, April 2017. https://www.vbb.de/unsere-themen/vbbdigital/api-entwicklerinfos/datensaetze.

[18] D. Ziemke and K. Nagel. Development of a fully synthetic and open scenario for agent-based transport simulations – the MATSim Open Berlin Scenario. VSP Working Paper 17-12, TU Berlin, Transport Systems Planning and Transport Telematics, 2017. URL http://www.vsp.tu-berlin.de/publications.

[19] D. Ziemke, K. Nagel, and C. Bhat. Integrating CEMDAP and MATSim to increase the transferability of transport demand models. *Transportation Research Record*, 2493:117–125, 2015. doi: 10.3141/2493-13.

[20] D. Ziemke, A. Agarwal, and I. Kadoura. Entwicklung eines regionalen, agentenbasierten Verkehrssimulationsmodells zur Analyse von Mmobilitätszenarien für die Region Ruhr. In *10. Wissenschaftsforum Mobilität*. VSP Working Paper 18-08, TU Berlin, Transport Systems Planning and Transport Telematics, 2017. URL http://www.vsp.tu-berlin.de/publications.