EUROPEAN COMMISSION
European Research Area

SEVENTH FRAMEWORK PROGRAMME

Funded under Socio-economic Sciences & Humanities

Investigating the MATSim warm and hot start capability

Thomas W. Nicolai

Berlin Institute of Technology (TU Berlin)

FP7-244557

SustainCity

EUROPEAN COMMISSION
European Research Area

SEVENTH FRAMEWORK PROGRAMME

Funded under Socio-economic Sciences & Humanities

Contents

Investigating the MATSim warm and hot start capability

Thomas W. Nicolai
TU Berlin
Salzufer 17-19
10587 Berlin Germany
phone: +49 30 314-29 592
fax: +49 30 314-26 269
nicolai@vsp.tu-berlin.de

Abstract

Modern disaggregated, agent-based travel models like MATSim typically require a large amount of computing time to generate the traffic assignment. This challenge is addressed by a so-called warm and hot start capability of MATSim that will be explained in this paper. In both cases information from previous MATSim runs are recycled in order to reduce the computing times for the traffic flow simulation.

1 Introduction

In this paper MATSim is used as a travel model plug-in within the micro-simulation land use model UrbanSim (Waddell, 2000, 2002; Waddell and Ulfarsson, 2004; Miller *et al.*, 2005; OPUS User Guide, 2011). In this configuration UrbanSim calls the travel model in regular intervals and over several years to update the traffic conditions from the current land use. Readers who are interested in a detailed description of the simulation and integration approach of MATSim with UrbanSim are referred to (e.g. Nicolai and Nagel, 2010, 2012a).

This paper will focus on MATSim. For the present study the land use model is used to provide the input for the MATSim runs such as home and work locations of individual agents. MATSim uses this information to generate the initial travel demand, which is optimized over many iteration cycles. In this situation it seems reasonable to reuse the travel schedules and routing decision of agents that have not changed compared to a previous MATSim run, e.g. agents that still have the same residence and work location, in order to reduce the number of iterations of the optimization process and thus the computing times for the traffic simulation. This methodology is used by the so-called *warm* and *hot start*. Their workings will be explained in the following section. The proposed approaches are applied to a real-world scenario. The scenario is described in Sec. 1.2, Sec. 2 illustrates the outcomes of this application. This closes with a discussion and conclusion.

At this point a brief overview about relevant MATSim notations, the general simulation process inside MATSim and in particular the iterative demand optimization process is provided to establish a common communication base.

The simulation process inside MATSim consists of the following parts (e.g. Balmer *et al.*, 2009):

- **Initial demand**: MATSim requires the physical infrastructure, determined by the road network and facilities (i.e. activity locations like home, work, shopping or leisure) and the population including the demand of each individual agent. In MATSim the demand of an agent is called plan (Balmer *et al.*, 2009). A plan encodes the daily routine of an agent. It contains the agents travel schedule including its intended activities and routing decisions between the activity locations (Balmer *et al.*, 2005). Moreover, a plan captures the order, type, location, duration as well as other time constraints for every activity and the selected mode, route and expected departure and travel times of each leg (Balmer *et al.*, 2005). A leg describes a part of a trip that uses exactly one transport mode (Balmer *et al.*, 2005).

- **Iterative demand optimization**: In an iterative demand optimization process the demand for each individual agent is improved. It takes into account physical constraints, e.g. the road network, and the interaction between the agents. The optimization process consists of an iteration cycle with three main steps, “Execution”, “Scoring” and “Re-planning”, which are explained below in more detail.

- **Analysis**: Finally, the simulation results such as the resulting population and demand and the traffic conditions on the network can be used for post-process analysis.

As mentioned before the iterative demand optimization process consists of an iterative loop with three steps, see Fig. 1. The main steps are summarized in the following, more in-depth information is given in (Balmer *et al.*, 2005; Raney and Nagel, 2006; Balmer *et al.*, 2009) :

1. **Execution**: The mobility simulation executes all agents with their selected plan simultaneously on the road network (physical layer). At this stage agents are interacting with the physical environment and with other agents.

2. **Scoring**: All executed plans are scored by a utility function that determines the performance of each plan. For this task various scoring functions can be used. By default MATSim uses the so-called Charypar-Nagel scoring function, see (Charypar and Nagel, 2005)

3. **Re-planning**: In this step some agents choose between existing plans, some re-evaluate plans with bad scores and some obtain new plans by modifying existing ones.

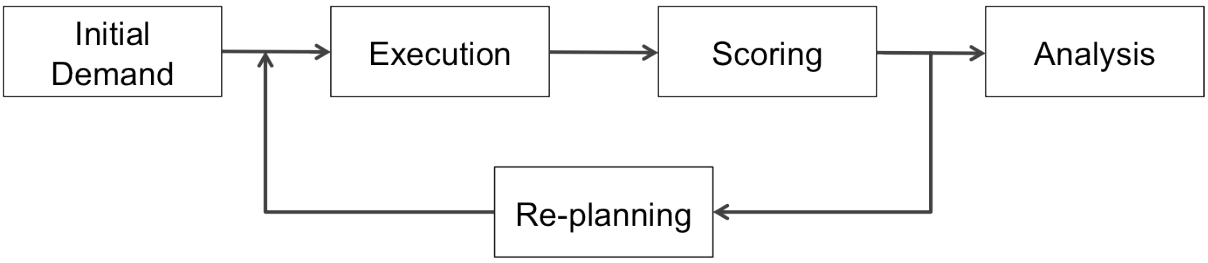


Figure 1: This shows the process structure of MATSim. Once the initial demand is generated, agent’s plans are optimized in an iterative process until a relaxed state of the system (usually a user equilibrium) has been reached. Finally, MATSim obtains detailed output from the traffic simulation, which can be used for further analysis.

Usually, the iteration cycle is allowed to continue until a *relaxed* state of the system has been reached (Balmer *et al.*, 2005, p.98). However, there is no quantitative measure of when this state is reached (Balmer *et al.*, 2005, p.98). According to (Balmer *et al.*, 2009, p.70) a relaxed state is reached when “the utility for each agent does not noticeably change through variation of the day plans” and “the trajectory of average utility per iteration represents a stationary process”.

In the subsequent section, the relaxed state of the system will be used as an important indicator to assess the performance of the proposed *warm* and *hot start* approach.

1.1 Implementation

At this point three notations are introduced: cold, warm and hot start.

Cold start This was in principle described in the previous section. It means that MATSim generates the initial travel demand based on the provided UrbanSim population and current land use. This is illustrated in Fig. 2.

Warm and hot start Warm and hot start are describing the capability of MATSim to start a simulation from a pre-existing, relaxed plans file to recycle information, e.g. route, transport mode and departure time choices, from previous MATSim runs. In other words, MATSim "remembers" the travel schedule (daily plan) of each traveller. Consequently, fewer iterations are and thus less computing time is required to reach a relaxed state of the system.

The difference between warm and hot start is given by the use of distinct plans files. In warm start, MATSim will always use the same relaxed *initial plans* file as shown in Fig. 2. However, when running UrbanSim over many years with a changing population, e.g. due to relocation of households and firms, the initial plans file will become less and less correct and requires more and more iterations to get MATSim back into a relaxed state. This issue is addressed by hot start. Here, MATSim stores an *updated plans* file after each run that incorporates all changes of the current UrbanSim population. As a result differences between the updated plans file and the UrbanSim population of a subsequent year (UrbanSim iteration) are kept small. As opposed to warm start, MATSim uses the updated instead of the initial plans file with hot start as shown in Fig. 2.

Implementation Technically, when MATSim performs a warm or hot start it reads a plans file together with the current UrbanSim population and compares them. When converting the current UrbanSim population into MATSim agents the person id's form UrbanSim are taken. Thus, UrbanSim persons can be individually identified among the MATSim agents via their person id.

MATSim keeps all plans or persons from the initial plans file that have not changed. In order to determine this, the following decision tree is expired for each person:

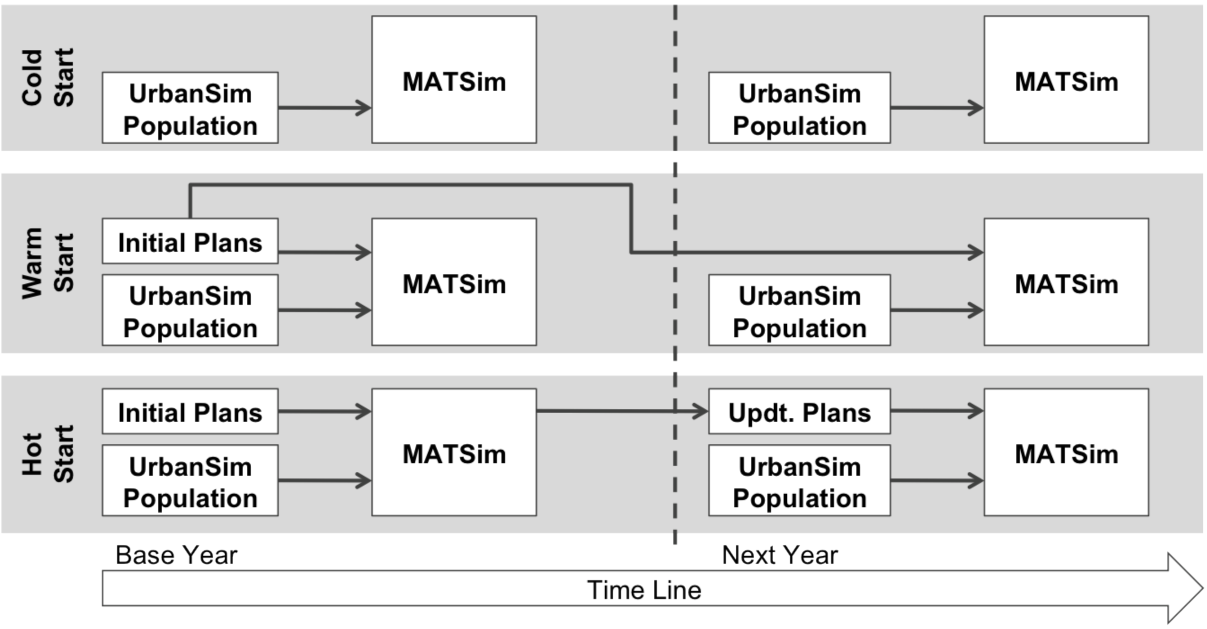


Figure 2: This illustrates the working of the (i) cold, (ii) warm and (iii) hot start in MATSim. With (i) cold start MATSim generates the initial demand from the current UrbanSim population. With warm and hot start MATSim recycles information from previous runs; as a result less iterations are required to reach a relaxed state of the system. (ii) Warm start will always use the same plans file, which becomes less and less correct when running UrbanSim over a long time span. Opposed to warm start, (iii) in hot start after each run an updated plans file is stored, incorporates changes from the current the UrbanSim population. In this case, MATSim start starts from the updated plans file instead of the initial plans file.

1. A person from the current UrbanSim population exists in the plans file, i.e. an agent plan with the same person id exists in the plans file.
2. The person has the same employment status.
3. The person has the same home location.
4. And if applicable, the person has the same work location.

If one of this points did not apply a new plan is generated.

1.2 Scenario: Zurich, Switzerland

The cold, warm and hot start above approach are now applied to a real-world scenario, the city of Zurich (Switzerland), a parcel-based UrbanSim application. A comprehensive description about the Zurich application is given by (Schirmer *et al.*, 2011; Schirmer, 2010); the scenario is already described in many previous papers (e.g. Nicolai and Nagel, 2012a,b). At this point a brief overview is provided.

The Zurich case study uses the year 2000 as the UrbanSim base year. It stores the initial state of the study area. The data that is needed to create the base year such as a population census, mobility census, enterprise census, etc., comes from several sources that can be divided into two main categories: governmental and private data. The sources for governmental data includes various Swiss federal, cantonal,[1] and municipal offices. The acquisition of private data includes private institutions, web-sites, and self created data.

The Zurich application consists of 40'407 parcels, 234 zones, 336'291 inhabitants, and 316'703 jobs. In the following the UrbanSim base year, 2000, is used to create the input for the MATSim runs. After that, UrbanSim is no longer needed for the present setting.

Population and travel demand In order to speed up computation times, MATSim considers a 10% random sample of the synthetic UrbanSim population, consisting of 33'629 agents. All MATSim agents have complete day plans with "home-to-work-to-home" activity chains. Work activities can be started between 7 and 9 o'clock, and have a typical duration of 8 hours. The home activity has a typical duration of 12 hours and no temporal restriction.

Network and adjustments A revised Swiss regional planning network (Vrtic *et al.*, 2003; Chen *et al.*, 2008) is used that includes major European transit corridors as depicted in Fig. **??**. The network consists of 24'180 nodes and 60'492 links, where each link is defined by an origin and a destination node, a length, a free speed car travel time, a flow capacity and a number of lanes. In addition each link obtains congested car travel times once the traffic flow simulation in MATSim is completed (see Nicolai and Nagel, 2012a).

The following summarizes modifications to improve link capacities especially at the urban scale; a detailed description is given in (Chen *et al.*, 2008, pp.7). All links within a radius of 4 kilometers around the Zurich city center were modified as follows:

- Links that correspond to so-called primary[2] roads in OpenStreetMap[3] (OSM) get a capacity of at least 2000 vehicles per hour. Links with higher capacities remain unchanged.
- Links that correspond to secondary roads in OSM keep their initial capacity (usually between 1000 and 2000 vehicles per hour).

[1] A Swiss Canton corresponds to a federal state
[2] an open street map road classification is given at `http://wiki.openstreetmap.org/wiki/Highway_tag_usage`
[3] see `http://www.openstreetmap.org`

- The remaining links get a capacity with a maximum of 600 vehicles per hour. If the original capacity is lower, it is not changed.
- Finally, a few individual links are adjusted manually based on local knowledge.

In addition, the network flow and storage capacities are adjusted automatically based on the selected population sampling rate. More in-depth information about this is given in (Nicolai, 2012; Nicolai and Nagel, 2012a).

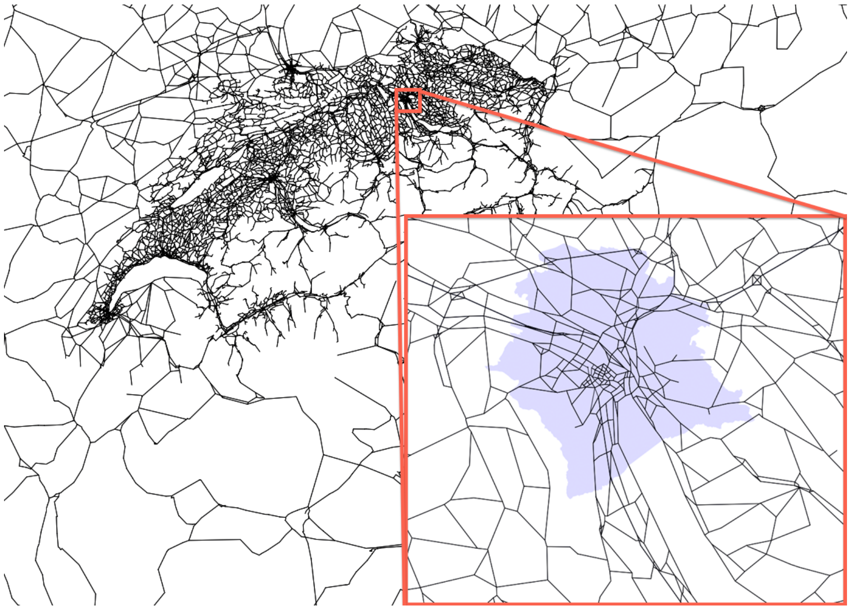


Figure 3: The Zurich case study network, area of Zurich (light blue) enlarged.

Preparatory MATSim run A preparatory MATSim run is performed based on the UrbanSim base year by running the traffic simulation for 100 iterations with a 10% random sample of the UrbanSim population. 10% of the agents perform “time adaptation”, which changes the departure times of an agent, and 10% adapt their routes. The remaining agents switch between their plans. After 80% of the iterations time and route adaptations are switched off; thus, agents only switch between existing plans.

The resulting relaxed plans file is used as input for warm and hot start runs as described in the subsequent section.

2 Simulation runs and results

In the present setup UrbanSim is executed for the period from 2000 (base year) to 2010. In this period MATSim is called for the years 2001, 2003, 2005, 2007 and 2009. MATSim uses the same re-planning configuration as for the preparatory run described in the previous section. This means, 10% of the agents use time and another 10% use route adaptation, the remaining agentes are allowed to switch between their plans, where time and route adaptation are switched off after 80% of the iterations.

To analyze the outcomes of the MATSim cold, warm and hot start three simulation runs are performed:

1. **Cold start**: No input plans file is provided. MATSim will run for 100 iterations.
2. **Warm start**: The plans file from the preparatory MATSim run is provided. The number of MATSim iterations are reduced to 60.
3. **Hot start**: In contrast to warm start MATSim takes the plans file from the preparatory MATSim run only for 2001, the first time MATSim is called. In subsequent calls MATSim uses the updated plan file from the previous run as explained above in Sec. 1.1. At this point MATSim iteration are limited to 30.

The outcomes of these simulation runs for the selected years are presented in Fig. 4, 5 and 6. They show the average score (utility) over all best (blue line), worst (red line) and executed (yellow line) plans. In addition, the total average over all plans of each agent is indicated by the green line. For more in-depth information about the relaxation process in MATSim is given in (Balmer *et al.*, 2009).

The evaluation will concentrate on the performance of the executed plans (yellow line) that are actually performed in the traffic flow simulation.

In Fig. 4 the results for cold start are presented. It can be noticed that the average score of executed plans initially starts at a very low level. In each iteration agents try to optimize their plans by applying the re-planning strategies as described above. Especially during the first 10 - 20 iterations the performance of the plans improves significantly. After 80 iterations the average score of executed plans converges into a relaxed, steady state and approaches the average best score (blue line). At this point time and route adaptation were switched off.

Warm start, see Fig. 5, shows a more differentiated picture. First, in 2001 (Fig. 4(a)) the score of the executed plans in iteration 0 is almost as high as the in iteration 60, the last iteration. In each subsequent MATSim run the score starts at a lower level. This is due to the initial plans file, which becomes less accurate over the years. Compared to cold start, less iterations are required to reach a relaxed state.

In Fig. 6 the results for hot start are shown. In contrast to warm start the score of the executed plans in iteration 0 can be considered as more stable, i.e. they do not start at a lower level in each subsequent year. This may be due to the updated plan files that are input. Moreover, even less iterations than in warm start are necessary to reach a relaxed state of the score. These few iterations can bee seen as the refinement procedure, where mainly agents that are changed or new try to optimize their plans.

At this point, the computing times for each scenario are presented in Tab. 1. For the present study the number of iterations are reduced by about one third for warm start (60 iterations) and about two third for hot start (30 iterations) compared to cold start (100 iterations). This reduction is also reflected in the computing times.

Mode	**2001**	**2003**	**2005**	**2007**	**2009**	**Total time**
Cold start	121 min	121 min	122 min	121 min	120 min	605 min
Warm start	73 min	74 min	75 min	76 min	76 min	374 min
Hot start	41 min	41 min	41 min	40 min	40 min	203 min

Table 1: This table lists the computation times for cold, warm and hot start. All measurements are performed on a Mac Book Pro with an Intel Core 2 Duo 2.5GHz processor and 4 GB of memory.

2.1 Discussion and conclusion

As the results show in Tab. 1, the warm and hot start requires less iterations, and thus less computing times, in order to reach a relaxed state of the system. This is achieved by recycling agent plans or travel schedules from previous MATSim runs.

As already mentioned above, there is no quantitative measure of when the system reached a relaxed state. For the present study the number of iterations for each simulation run, cold, warm and hot start, are selected for illustration purposes. The required number of iterations will vary depending on the respective scenario.

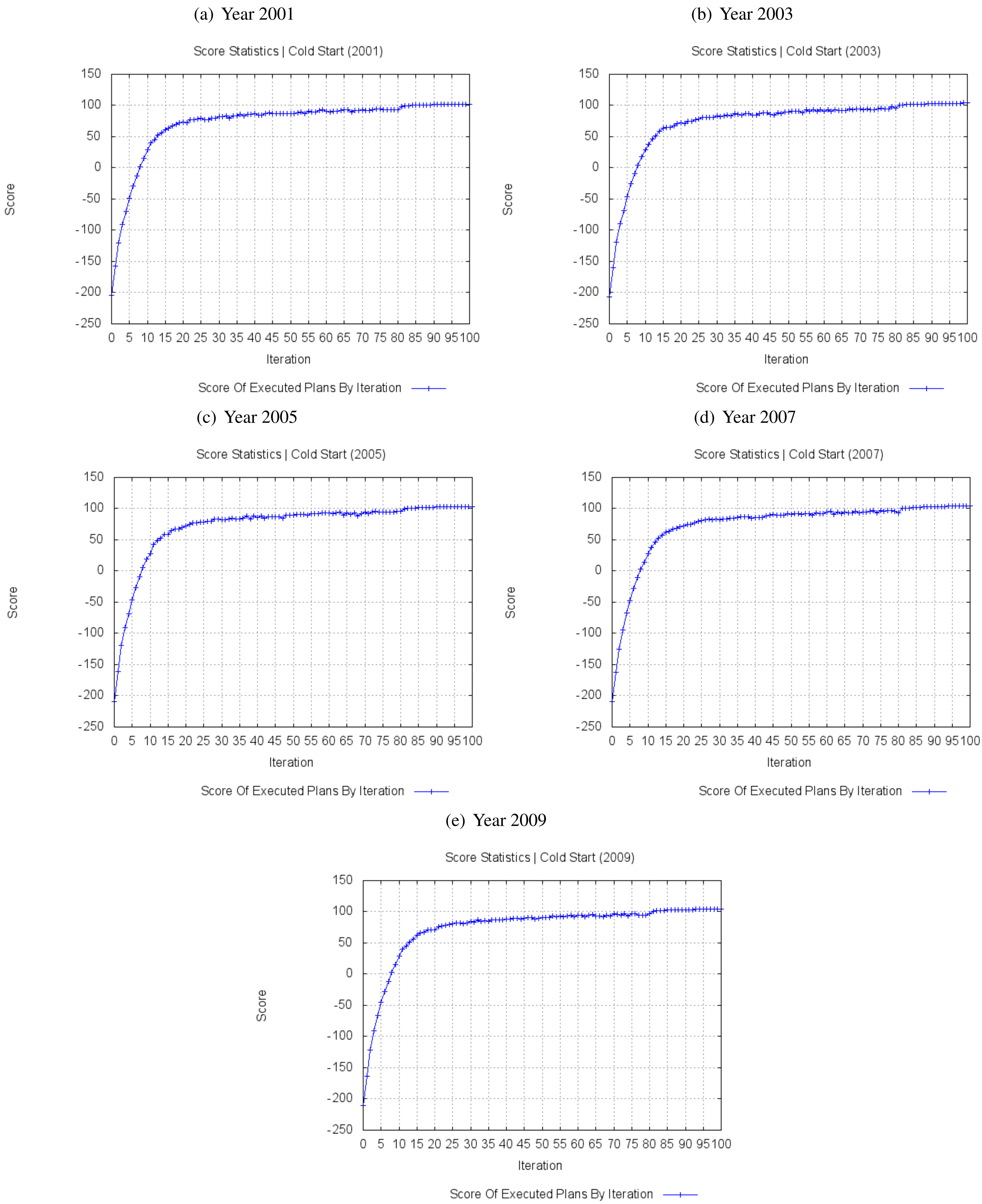
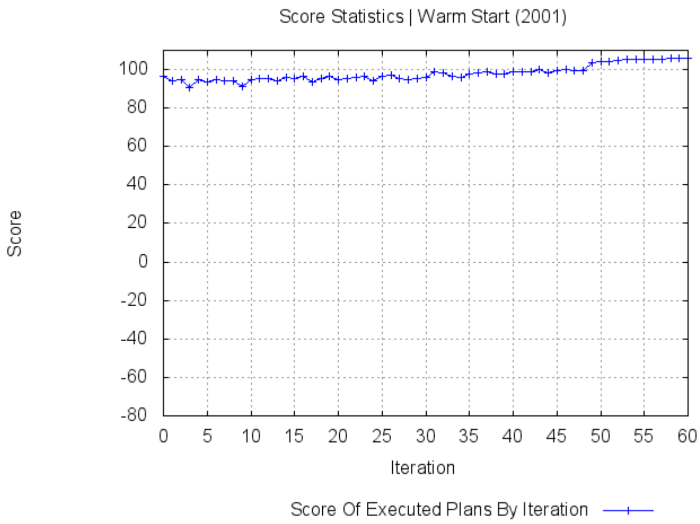
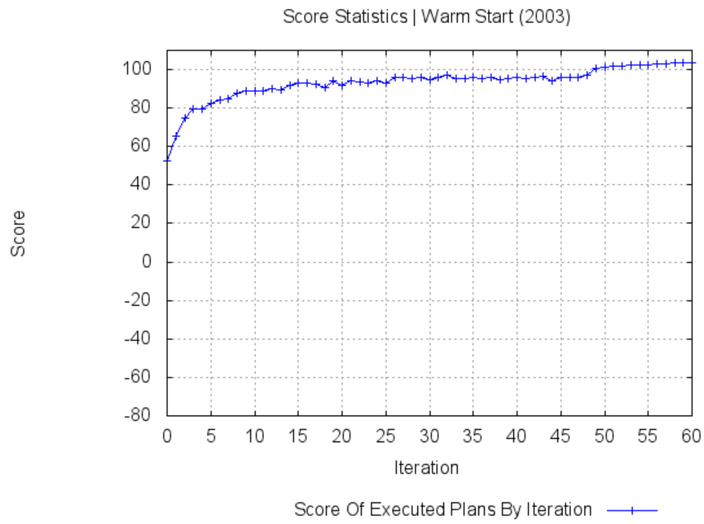


Figure 4: Cold start: These plots visualize the score (blue line) of the executed agent plans by iteration. It can be noticed that the score (performance of the plans) converges in to a relaxed, steady state after 80 iterations.

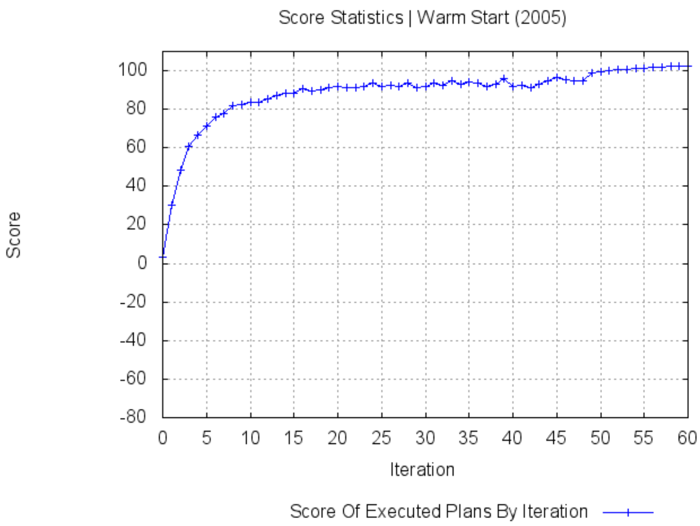
(a) Year 2001



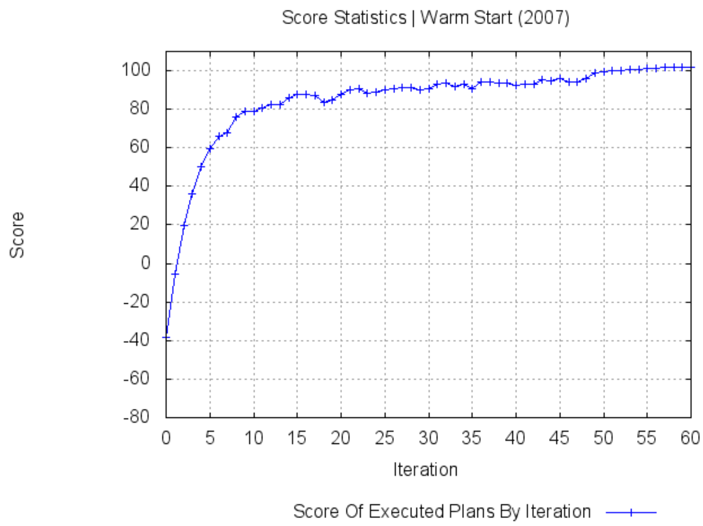
(b) Year 2003



(c) Year 2005



(d) Year 2007



(e) Year 2009

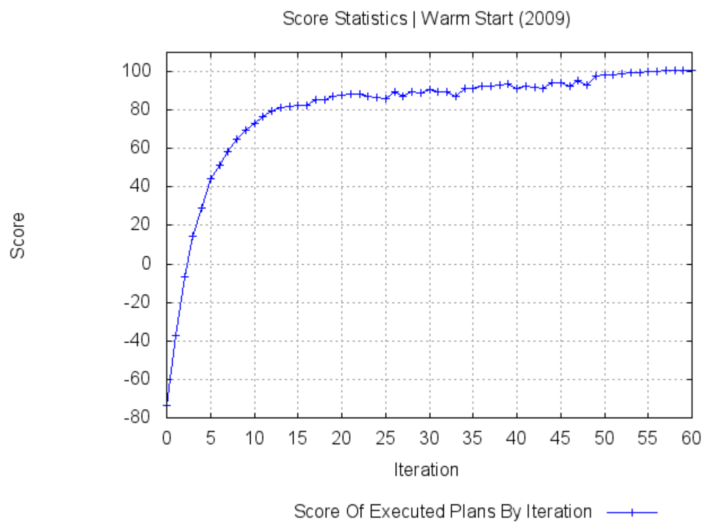
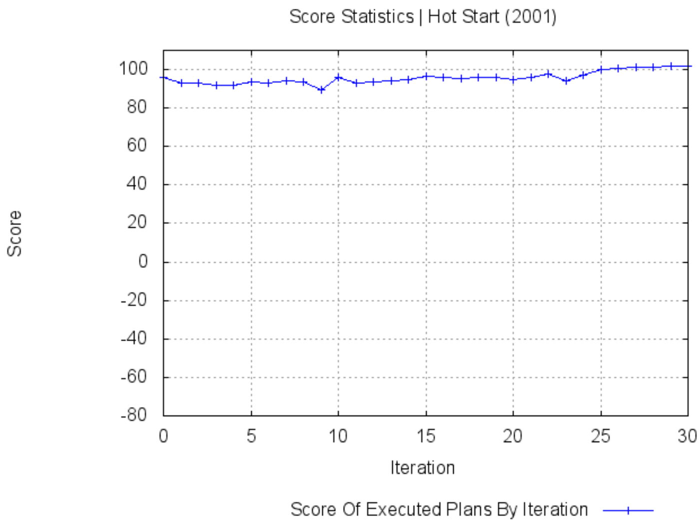
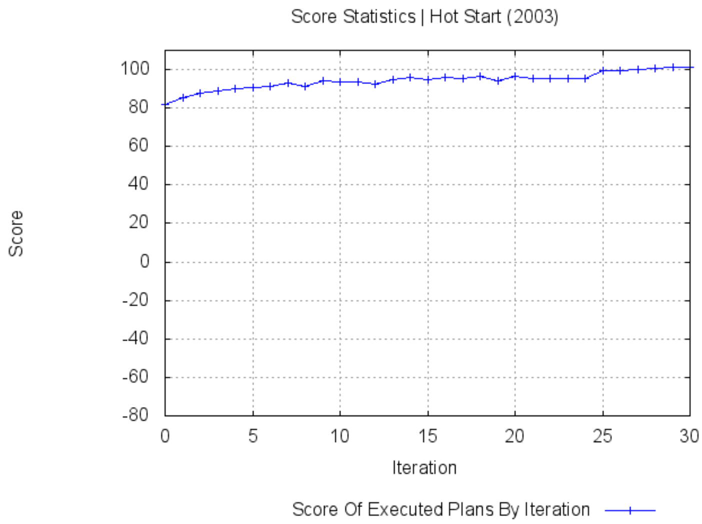


Figure 5: Warm start: These plots visualize the score (blue line) of the executed agent plans by iteration. In each subsequent MATSim run the score (performance of the plans) starts at a lower level, which indicates that the initial plans file becomes less and less correct. Fewer iterations (50 iterations) compared to cold start is required to reach a relaxed, steady state of the system.

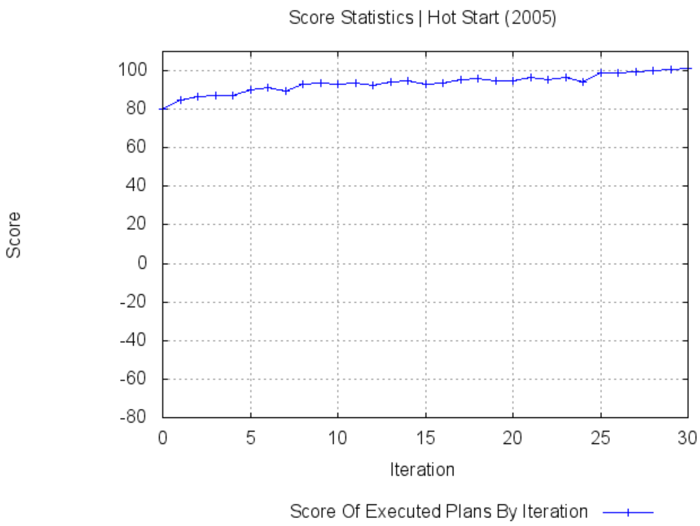
(a) Year 2001



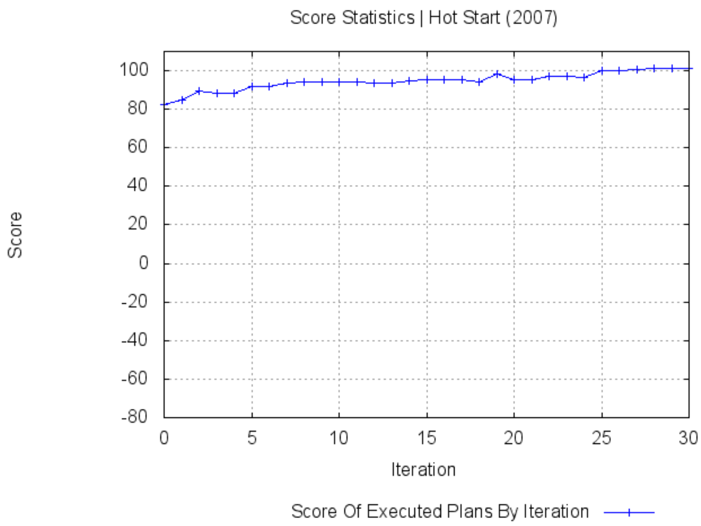
(b) Year 2003



(c) Year 2005



(d) Year 2007



(e) Year 2009

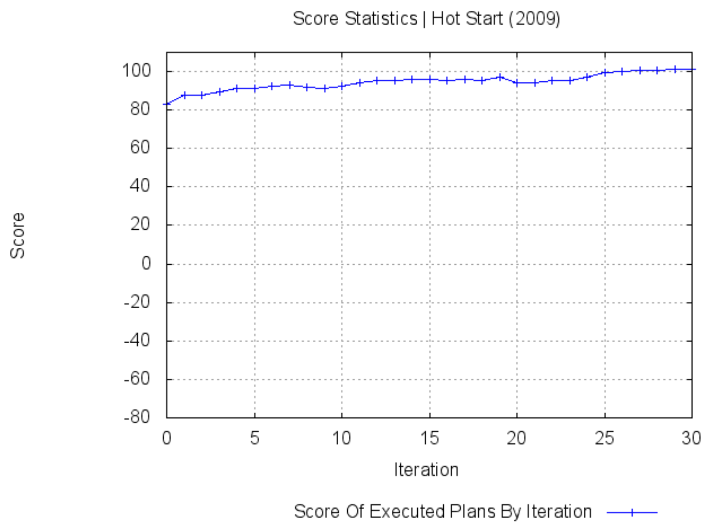


Figure 6: Hot start: These plots visualize the score (blue line) of the executed agent plans by iteration. The initial score (performance of the plans) always starts on a consistently high level. Less iterations compared to cold and warm start are required a relaxed, steady state of the system.

3 References

Balmer, M., B. Raney and K. Nagel (2005) Adjustment of activity timing and duration in an agent-based traffic flow simulation, in H. Timmermans (ed.) *Progress in activity-based analysis*, 91–114, Elsevier, Oxford, UK.

Balmer, M., M. Rieser, K. Meister, D. Charypar, N. Lefebvre, K. Nagel and K. Axhausen (2009) MATSim-T: Architecture and simulation times, in A. Bazzan and F. Klügl (eds.) *Multi-Agent Systems for Traffic and Transportation*, 57–78, IGI Global.

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

Chen, Y., M. Rieser, D. Grether and K. Nagel (2008) Improving a large-scale agent-based simulation scenario, *VSP Working Paper*, **08-15**, TU Berlin, Transport Systems Planning and Transport Telematics. See www.vsp.tu-berlin.de/publications.

Miller, E., K. Nagel, H. Ševčíková, D. Socha and P. Waddell (2005) OPUS: An open platform for urban simulation, paper presented at *9th Conference on Computers in Urban planning and urban management (CUPUM)*, University College London, UK, `http://128.40.111.250/cupum/searchpapers/detail.asp?pID=428`. See www.cupum.org.

Nicolai, T. and K. Nagel (2010) Coupling MATSim and UrbanSim: Software design issues, *VSP Working Paper*, **10-13**, TU Berlin, Transport Systems Planning and Transport Telematics. See www.vsp.tu-berlin.de/publications.

Nicolai, T. and K. Nagel (2012a) Coupling transport and land-use: Investigating accessibility indicators for feedback from a travel to a land use model, paper presented at *Latsis Symposium 2012 – 1st European Symposium on Quantitative Methods in Transportation Systems*, Lausanne, Switzerland. Also VSP WP 12-16, see www.vsp.tu-berlin.de/publications.

Nicolai, T. and K. Nagel (2012b) Sensitivity tests with high resolution accessibility computations, *VSP Working Paper*, **12-22**, TU Berlin, Transport Systems Planning and Transport Telematics. See www.vsp.tu-berlin.de/publications.

Nicolai, T. W. (2012) Using MATSim as a travel model plug-in to UrbanSim, *VSP Working Paper*, **12-29**, TU Berlin, Transport Systems Planning and Transport Telematics. See www.vsp.tu-berlin.de/publications.

OPUS User Guide (2011) *The Open Platform for Urban Simulation and UrbanSim Version 4.3*, University of California Berkley and University of Washington, January 2011, `http://www.urbansim.org`.

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

Schirmer, P. (2010) Options and constraints of a parcel based approach in 'UrbanSimE', paper presented at *Proceedings of the 10th Swiss Transport Research Conference (STRC)*. See www.strc.ch.

Schirmer, P., C. Zöllig, B. R. Bodenmann and K. W. Axhausen (2011) The Zurich case study of UrbanSim, paper presented at *European Regional Science Association Conference*, June 2011.

Vrtic, M., P. Fröhlich and K. Axhausen (2003) Schweizerische Netzmodelle für Strassen- und Schienenverkehr, in T. Bieger, C. Lässer and R. Maggi (eds.) *Jahrbuch 2002/03 Schweizerische Verkehrswirtschaft*, 119–140, Schweizer Verkehrswissenschaftliche Gesellschaft, St. Gallen.

Waddell, P. (2000) A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of urbansim, *Environment and Planning B: Planning and Design*, **27**, 247 – 263.

Waddell, P. (2002) Urbansim: Modeling urban development for land use, transportation, and environmental planning, *Journal of American planning Association*, **68** (3) 297 – 314.

Waddell, P. and G. F. Ulfarsson (2004) Introduction to urban simulation: Design and development of operational models, paper presented at *Handbook of transport geography and spatial systems*, vol. 5, 203–236.