

---

---

# **Location Choice Modeling for Shopping and Leisure Activities with MATSim: Utility Function Extension and Validation Results**

**Andreas Horni**

**Darren M. Scott**

**Michael Balmer**

**Kay.W. Axhausen**

**Transport and Spatial Planning**

**2009**

**STRC**

**9<sup>th</sup> Swiss Transport Research Conference**

Monte Verità / Ascona, September 9. - 11. 2009

Transport and Spatial Planning

## Location Choice Modeling for Shopping and Leisure Activities with MATSim: Utility Function Extension and Validation Results

Andreas Horni  
Institute for Transport Planning and Systems (IVT)  
ETH Zürich  
CH-8093 Zürich  
phone: +41-44-633 31 51  
fax: +41-44-633 10 57  
horni@ivt.baug.ethz.ch

Darren M. Scott  
Center for Spatial Analysis,  
School of Geography &  
Earth Sciences  
McMaster University  
Hamilton, Ontario L8S 4K1  
phone: +1-905-525-9140,  
ext. 24953  
fax: +1-905-546-0463  
scottdm@mcmaster.ca

Michael Balmer  
Institute for Transport Planning and Systems (IVT)  
ETH Zürich  
CH-8093 Zürich  
phone: +41-44-633 27 80  
fax: +41-44-633 10 57  
balmer@ivt.baug.ethz.ch

Kay.W. Axhausen  
Institute for Transport Planning and Systems (IVT)  
ETH Zürich  
CH-8093 Zürich  
phone: +41-44-633 39 43  
fax: +41-44-633 10 57  
axhausen@ivt.baug.ethz.ch

2009

### Abstract

This paper presents validation results for the activity-based multi-agent transport simulation MATSim (<http://www.matsim.org>), where the main focus lies on the location choice module for shopping and leisure activities. Validation results are produced by simulating a 10 % sample of the Swiss motorized individual traffic.

For Switzerland detailed information about home, working and education locations together with the associated trip matrices are provided by the census. Naturally this level of detail is not available for shopping and leisure trips. In MATSim so far—to create feasible activity chains—location choice for these activities was done in a preprocessing step based on a simple nearest neighbor search, which clearly leads to a systematic underestimation of the traffic volume. In this paper a two-fold shopping and leisure location choice model is presented that produces substantially better results.

First and foremost, the to-date exclusively *time-based* utility function for shopping activities is extended to take into account further determinants of shopping location choice, such as the *store size* and the *stores density* in a given neighborhood. In activity-based models, shopping lo-

cation choice is influenced by leisure location choice, which means that a meaningful shopping location choice model requires a sound leisure location choice model. The long-term goal of MATSim is to model leisure location choice by utility maximization and by including models of social interaction. But these models are far from being productive in agent-based transportation models in general. Hence, we introduce *hollow* space-time prisms that are derived from empirical data. This approach is—to our knowledge—a novel extension of Hägerstrand's time geography that by construction produces statistically correct leisure location choice and improves the simulation results in general.

Furthermore, the potential of MATSim to also serve as a hypothesis testing tool—besides being a planning tool—is highlighted in this paper. It is shown that MATSim provides the possibility to test models, generated by utility maximizing approaches such as e. g., discrete choice models, in large scenarios, whereby use can be made of data (e. g., count data) that is potentially qualitatively distinct from the data that were used for estimating and validating the models in question in earlier stages.

## **Keywords**

shopping location choice

# 1 Introduction

The activity-based multi-agent simulation toolkit MATSim (MATSim-T, 2008; Balmer, 2007; Rieser *et al.*, 2007) adopts a co-evolutionary approach to capturing the patterns of people's activity scheduling and participation behavior at a high level of detail.

In ?Horni *et al.* (2008) the MATSim location choice module (for the time being shopping and leisure activities only) is introduced. The authors demonstrate that the huge search space for real-world problems that is spanned by location choice has to be handled by *local* search methods, well-known from numerical optimization. It is furthermore shown that a local search can be incorporated easily and consistently into MATSim using Hagerstrand's time-geographic approach (Hagerstrand, 1989). This first part which can be seen as the verification part is continued by the validation part that is presented in this paper. For that purpose the motorized individual traffic of the Swiss population for a working day is simulated.

MATSim is primarily designed to provide a planning tool. In this paper the potential of MATSim to also serve as a hypothesis testing tool is highlighted. MATSim is based on utility maximization. Testing models that were generated by utility maximizing approaches (e. g., discrete choice models) in a large scenario is thus possible in a direct manner.

## 2 Method

### 2.1 MATSim

MATSim is an activity-based, easily extendable, open source multi-agent simulation toolkit implemented in JAVA and constructed to handle large-scale scenarios. MATSim is designed as a co-evolutionary system model. This means that while being in a competition for space-time slots on the transportation infrastructure with all the other agents, every agent iteratively optimizes its daily activity chain by *trial and error*. To do this, every agent possesses a memory of a fixed number of day plans, where each plan contains a daily activity chain and an associated utility value (in MATSim called *plan score*). The computation of the plan utility is compatible with micro-economic foundations and is described in more detail later. In every iteration prior to the execution of the micro-simulation (in MATSim called *Mobsim*, e. g., Cetin, 2005), every agent selects a plan from its memory probabilistically dependent on the plan utility for execution on the infrastructure and possibly preceding modification. A certain share of the agents (usually 10 %) is allowed to modify their plans, where time, route and location choices for shopping and leisure activities can be made. Prior to the selection of a plan for modification and execution, the plan with the lowest score is removed from the memory of these agents. The

plan which has been selected for modification is duplicated prior to modification. Time choice is based on local random mutation (e. g., Balmer *et al.*, 2005), route choice uses a best-response module, namely the A-star algorithm (Lefebvre and Balmer, 2007) and location choice uses a local search based on space-time constraints. An iteration is completed by evaluating the utility of every agent's performed daily routine, that represents a realization of the respective day plan.

### 2.1.1 Uniqueness of the MATSim Planning Equilibrium and Termination Condition

In MATSim an equilibrium with respect to the agents' day plans is searched. We call this a *scheduling* equilibrium in contrast to a *network* i. e. *Wardrop* equilibrium. It is subject of future research to specify this equilibrium in more detail, such as e. g., its relation to a Nash equilibrium (Nash, 1951). Likewise a task for the future is answering the crucial question about the uniqueness and the existence of such an equilibrium in our simulation as well as in reality. Nevertheless, for real-world scenarios having the two choice dimensions time and route choice, the procedure described above has so far shown convergence towards a unique planning equilibrium.

No explicit termination condition is established for MATSim yet, thus runs are stopped by the operator when a relaxed state with respect to the average utility and the average travel distances and times can be observed.

## 2.2 Location Choice for Shopping and Leisure

The long-term goal of MATSim is to model both shopping and leisure location choice by utility maximization and by including models of social interaction as e. g., according to the Swiss microcensus 20 % of leisure travel is generated by visiting friends. But these models are far from being productive in agent-based transportation models in general as leisure behavior is tremendously variable. To reach the above mentioned goal we follow a two-stage approach, where the first stage is presented in this paper.

In the first stage, the probably easier to handle shopping behavior is modeled strictly based on utility maximization. In activity-based models, shopping location choice is influenced by leisure location choice, which means that a meaningful shopping location choice model requires a sound leisure location choice model. Leisure location choice is thus in the first stage modeled by *hollow* space-time prisms, that are derived from leisure trip length statistics of the Swiss microcensus. This leads to more realistic leisure traffic in our model at least on an aggregated level. In the forthcoming second stage this approach will be extended or maybe replaced according to the above mentioned long-term design goal of MATSim.

### 2.2.1 Shopping

For shopping location choice the approach described in ? is applied in a slightly modified manner. The location choice algorithm described therein is derived from the potential-path-algorithm of Scott (2006) and works as follows: A new location for an activity to be relocated is chosen based on space-time prisms. For efficiency reasons the prisms are without loss of quality not calculated explicitly but approximated in a first step. In an iterative second step a tentatively and randomly chosen location of the approximate prisms is accepted as the new location or rejected based on the actual access of the location. That algorithm relocates per MATSim iteration all flexible activities in an activity chain in a recursive manner.

As MATSim is explicitly dedicated to large-scale scenarios efficient algorithms are crucial. Thus, that first algorithm is tailored to relocate at most one shopping (and afterwards one leisure) activity per plan and per iteration. Additionally, in our follow-up algorithm the location choice set is now defined by the travel speed as given in the last iteration for the respective plan. This makes the checking of the access to a certain location by routing much less important, such that it can be omitted. Nevertheless, the idea of local search is unaffected by these recent modifications.

### 2.2.2 Leisure

As mentioned earlier location choice for leisure activities is likewise done on the basis of time geography, namely by hollow space-time prisms, or in other words, ring-shaped potential path areas, whose generation is described in the next section.

## 2.3 Leisure Location Choice Constrained by Hollow Space-Time Prisms Derived from Empirical Trip Lengths Data

### 2.3.1 Initial Demand

In a preprocessing step, i. e., when the initial demand is generated, a fixed trip distance  $d_i$  drawn from the Swiss microcensus leisure trip lengths distribution is assigned to each leisure  $trip_i$  for all plans. An initial location for the associated activity is assigned accordingly. Distances are given in terms of crow-fly distances as route lengths are not reported by the microcensus.

In disaggregated activity-based models this assignment process is crucial and non-trivial. In our model the distances  $d_i$  are assigned consistently with the agents' activity chain, as derived from the Swiss microcensus. We assign the leisure trip distances dependent on the two follow-

ing factors. First, in our probabilistic algorithm the assigned leisure trip length is negatively influenced by the existence of a work activity posterior to a leisure activity in the chain, as assuming shorter leisure trips that precede working activities is natural. Second, for leisure activities shorter than 7 hours the trip length is positively influenced by the duration associated with the respective leisure activity. For activities with preferred durations above 7 hours uniform assignment is executed.

This second step is based on the observed relation in the microcensus between trip distance and activity duration at the destination. For durations up to 7 hours a linear dependency can be seen which fades to noise caused by a small sample size (see Figure 1). The literature reports a positive and significant linear relation between trip distance and activity duration (Kitamura *et al.*, 1998) for home-based trips having the purposes recreation/shopping. For non-home-based recreation/shopping trips a positive but insignificant correlation is reported therein. This insignificance can presumably be explained with the non-linearity of the relation, as in fact, a positive and significant non-linear relation between the two variables is reported by Iragael (2007). However, further research clearly is necessary.

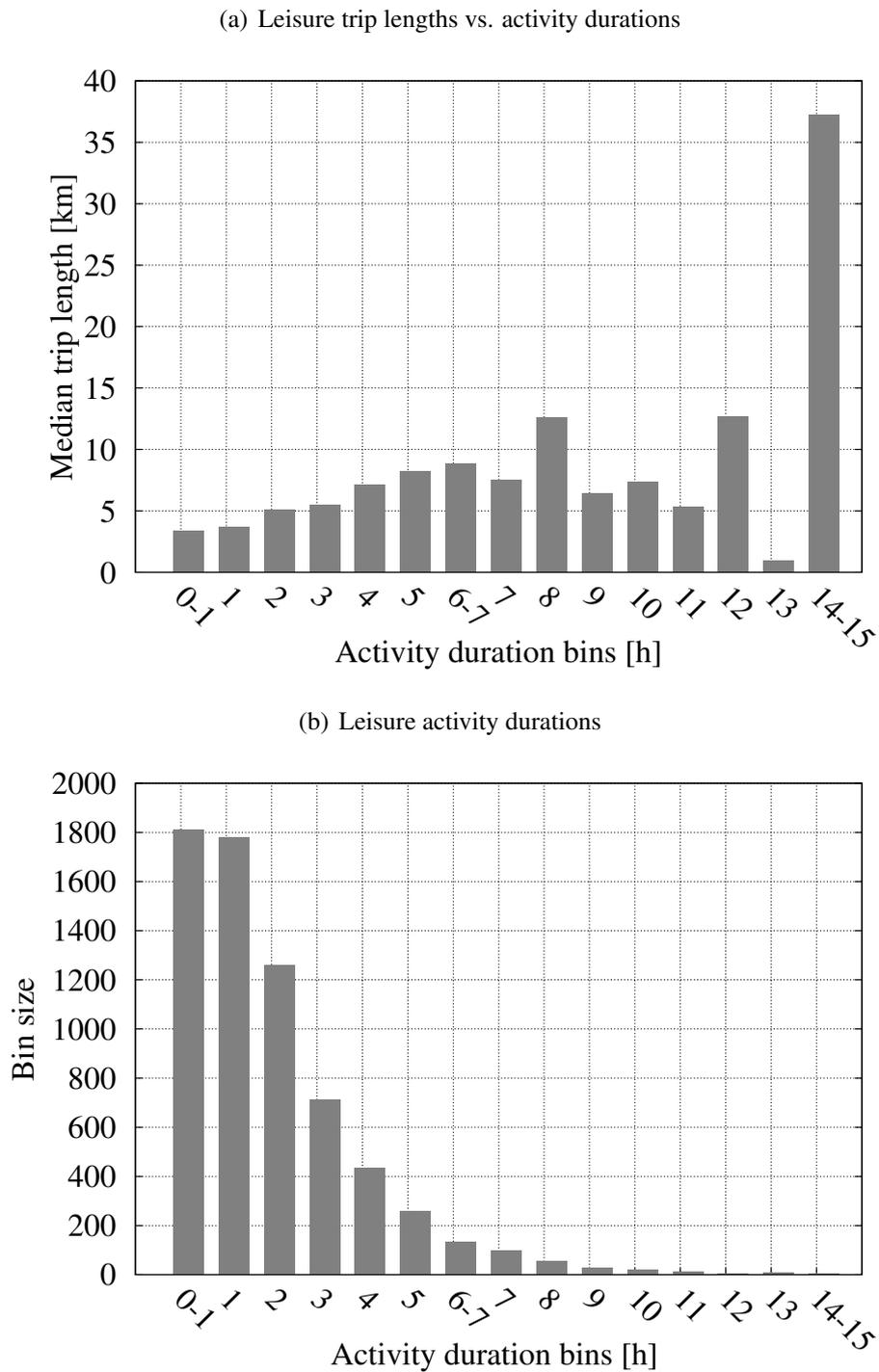
### **2.3.2 Replanning Step**

In the replanning step the leisure location choice set is given by a ring area with a relatively small difference between outer radius and inner radius ( $\epsilon = \text{MAX}(500m, 0.1 \times d_i)$ ), where the average of the two radii equals  $d_i$ . For efficiency reasons without loss of quality the ring choice set is not generated explicitly. Instead a location is randomly drawn in the circle with radius  $\epsilon$  around the point  $\vec{x}$ , where  $\vec{x}$  is given as the location of the preceding activity plus a random vector of length  $d_i$ . This allows to only searching a small region around a given point, which can be implemented by means of very efficient spatial data structures such as for example quad-trees. The complete location choice set is only constructed if after a number of trials no location is found.

### **2.3.3 Leisure Activity Locations**

In contrast to e. g., shopping activities, leisure activities can be performed essentially everywhere (e. g., walking, hiking etc.). Hence, the list of buildings dedicated to leisure activities given by the Federal Enterprise Census 2001 (Swiss Federal Statistical Office, 2001) (e. g., restaurants) is an incomplete set of all potential leisure locations. For our hollow space-time prisms approach it is thus most advantageous to assume a potential leisure location on every network link.

Figure 1: Swiss microcensus: Leisure activity durations and trip distances



## **2.4 Simulation Scenario**

### **2.4.1 Demand**

The initial demand of the simulated scenario is derived from the Swiss Census of Population 2000 (Swiss Federal Statistical Office, 2000) and the National Travel Survey for the years 2000 and 2005 (Swiss Federal Statistical Office, 2006) (Swiss microcensus). For our scenario, we drew a 10 % sample of the motorized individual traffic of the Swiss population, which results in 896'457 trips or 228'401 agents to be simulated. An average week day is modeled (Monday to Friday excluding national holidays). Activities are classified as home, working, education, shopping and leisure activities.

A detailed description of the MATSim population generation is given in Meister *et al.* (2009). First, a synthetic population is generated from the Swiss Census, which contains detailed demographic data of *all* Swiss inhabitants and furthermore exact home locations and corresponding work locations at municipality level. In a second step, these locations at municipality level are assigned to the persons. In the third step, the microcensus is used to generate activity chains that are assigned to the population based on the previous step and dependent on socio-demographic attributes, such as age, work status or driving license ownership.

Assuming a unique planning equilibrium the initial assignment of locations for shopping activities is irrelevant and can thus be done randomly. But to start with reasonable distances we use a simple neighborhood search to find initial shopping locations. For the hollow space-time prisms approach initial leisure locations are assigned according to the procedure described in the previous section.

### **2.4.2 Supply**

The activity location data set for working, education and shopping activities is derived from the Federal Enterprise Census. Home locations are given by the Swiss census. Leisure locations are defined as described above. In total more than 1.7 million activity locations are used in the simulation. Data for shopping store opening times are gathered from various sources. The network is an updated and corrected version of the Swiss National Transport Model (Vrtic *et al.*, 2003). Comparable data are available in most countries from official sources, such as censuses, national travel diary studies and commercial sources, such as navigation network providers, yellow pages publishers or business directories.

## 2.5 Activity Differentiation in Utility Maximization Models

The agents' base repertoire of activities as described above is a common activity classification in transportation planning (e. g., Swiss Federal Statistical Office, 2006). However, to improve simulation results a more detailed activity differentiation is assumed to be beneficial. But the definition of activities or more precisely activity types is complex as pointed out in the following.

Utility maximizing methods for activity planning are only meaningful if *substitutability* with respect to people's undertakings exists or, in other words, if these undertakings can be grouped into a finite number of activity classes.

As people live in a constrained world, the existence of substitutability is in general a valid assumption and thus, utility maximizing methods are applicable. The crucial question is where to draw the line between different activity types. We assume that substitutability of shopping activities is primarily influenced by substitutability of goods, which naturally leads to a first differentiation between *grocery* and *non-grocery* shopping. This partitioning is efficient, as the two new shopping activity types are both of substantial size. The Swiss microcensus reports a share of 66 % for grocery shopping trips.

Similar to the assignment of hollow space-time prisms to activities, the assignment of grocery and non-grocery shopping activities is non-trivial for disaggregated methods. As a first approach it is done randomly during preprocessing, but in the future a more sophisticated process has to be researched.

## 2.6 MATSim Utility Function

In this work the utility function of Charypar and Nagel (2005) is applied. This function is frequently used in MATSim. It is based on the *Vickrey* model for road congestion as described in Vickrey (1969) and Arnott *et al.* (1993). Originally, this utility function was constructed to model departure time choice. However, several studies accomplished by MATSim, such as e. g., Balmer *et al.* (2009), have given evidence that the function is also productive for modeling route choice. Consequently, it builds the starting point for our location choice utility function, where—beginning with the extension described in the next chapter—further determinants of location choice will be successively added in the future.

The utility of a plan  $U_{plan}$  (described in detail in Charypar and Nagel, 2005) is computed as the sum of all activity utilities  $U_{act,i}$  plus the sum of all travel (dis)utilities  $U_{trav,i}$

$$U_{plan} = \sum_{i=1}^n U_{act,i}(type_i, start_i, dur_i) + \sum_{i=2}^n U_{trav,i}(loc_{i-1}, loc_i)$$

where  $type_i$ ,  $start_i$  and  $dur_i$  are the type, start time and duration of the activity respectively. The utility of an activity is defined by:

$$U_{act,i} = (U_{dur,i} + U_{wait,i} + U_{late.ar,i} + U_{early.dp,i} + U_{short.dur,i})$$

$U_{dur,i}$  is the utility of performing an activity, where the opening times of activity locations are taken into account.  $U_{wait,i}$  denotes the disutility of waiting, and  $U_{late.ar,i}$  and  $U_{early.dp,i}$  give the disutility of late arrival and early departure respectively.  $U_{short.dur,i}$  is the penalty for too short an activity participation time.

## 2.7 Extension of the MATSim Utility Function

As mentioned earlier, the goal of this work is to extent the utility function for shopping activities to on the one hand improve the simulation results and on the other hand to show the potential of MATSim to serve as a hypothesis testing tool. The construction of a comprehensive parameter setting and systematic parameter calibration is a subject of future work.

In this work the above described utility function is extended by two store attributes that are determinants of shopping location choice as reported by the literature.

### 2.7.1 Store Size

A positive (but weak) relation between shopping store size and store attractiveness is reported for example in Lademann (2007). In our framework a higher attractiveness value can be transformed to an additional across-the-board benefit that an agent gets for visiting a large store. It is a priori expected that this new configuration leads to longer shopping trip distances as the number of large stores is relatively small and people thus probably need to drive farther.

The Swiss Enterprise Census provides information about the store size in categories. For the time being the store size is regarded for grocery shopping only. The additional benefit for such an activity is arbitrarily assigned as

- 0 € for visiting a store of size up to  $400m^2$

- 3 € for visiting a store of size  $400 - 1000m^2$
- 9 € for visiting a store of size  $1000 - 2500m^2$
- 18 € for visiting a store larger than  $2500m^2$

### 2.7.2 Stores Density

Similar to (e. g., Robinson and Vickerman, 1976; Recker and Kostyniuk, 1978) we assume that the attractiveness of a store increases with the number of adjacent stores, that is with the stores density in the neighborhood of a that store. Thus, agents get an additional across-the-board benefit, which is positively related to the stores density in a neighborhood region with an arbitrarily chosen diameter of  $1500m$ . The additional utility is computed as  $u_{additional} = 1 \text{ €} \times \text{Min}(20, \text{number of stores in the neighborhood of the respective store})$ . Again, the additional utility term is a priori expected to increase the average trip travel distance as dense regions with respect to stores are relatively rare.

### 2.7.3 Future Work

The literature reports numerous additional influence factors of shopping location choice, such as availability of parking lots, price level, range of goods and others. These have to be tested—individually and combined—and possibly included in our utility function in the future. Another future topic is the further consideration of people’s diversity, which is straight forward for multi-agent simulations.

## 2.8 Calibration

The two main points to calibrate MATSim are the parameters of the utility function and the parameters of the *Mobsim*. The later are used to correct network approximations, where corrections of link capacities play a major role. No rules for a systematic calibration of MATSim are researched yet.

The parameters of the MATSim base utility function are well-founded being derived from the Vickrey (1969) scenario. The calibration of the two additional parameters for the recent utility function extension, as described above, is defined as a future task.

For calibrating the mobility simulation in principle the reported *work* trip duration statistics of the Swiss microcensus are most suitable as it is assumed that the trip durations reported therein are most consistent for trips with purpose work as they are highly repetitive. To minimize

the effects of trips having a different purpose (e. g., leisure) the morning peak time, which is dominated by work traffic should be used.

However, in MATSim intersection dynamics (e. g., traffic lights) and car-following dynamics are not yet included. Additionally the access time to the transportation system (e. g., leaving the parking lots) is systematically underestimated especially for coarse networks. This prevents calibration for the time being. However, calibration is a potential source for further significant improvements of the results and is hence an important future task.

## **2.9 Validation**

For validation traffic count data and Swiss microcensus data are used. A GPS-survey (Hackney *et al.*, 2007) providing link speed estimates for the complete network is readily available for future validation steps.

### **2.9.1 Count Data**

Traffic count data of the years 2004–2005 of automatic national, cantonal and municipal count data stations are used, that measure 655 network links in total (e. g. ASTRA, 2006). The hourly traffic volumes for a normal working day are computed and compared. The stations count all motorized traffic, whereas our simulation simulates the motorized individual traffic of the Swiss population only. This results in a certain natural underestimation of the traffic volumes.

Count data, is an aggregated information, and it is therefore not the prime validation means for multi-agent simulations. However, the generally limited data availability argues for using count data for validation of multi-agent simulations in conjunction with disaggregated data.

### **2.9.2 Swiss Microcensus**

For validation the shopping trip length and duration statistics of the Swiss microcensus are used. The Swiss microcensus 2005, reporting 33'000 daily activity chains, serves as a basis for the MATSim initial demand. I. e., it is model input data, and thus care about the applicability of these data for validation purposes is necessary in general. However, shopping location choice, although being influenced by leisure location choice and by the activity chain as a whole, is essentially not predetermined by these input data. Thus, it is a valid means for validating shopping location choice.

## 3 Results, Conclusions and Outlook

### 3.1 Scenario configurations

First and foremost the goal of this work is to evaluate shopping location choice. As mentioned earlier, time and route choice is so far done by means of a strictly time-based utility function. It is shown that this approach is not productive for location choice. Thus, this base case is successively refined. In detail, the following is analyzed:

- (i) the results quality of the case where neither hollow space-time prisms nor activity differentiation nor utility function extensions are applied,
- (ii) the effects of hollow space-time prisms,
- (iii) the influence of the shopping activity differentiation (grocery, non-grocery) and
- (iv) the effects of the utility function extensions for shopping activities.

For this purpose the following five different scenario configurations are applied:

#### Base case:

- **Configuration 0:** Location choice for shopping and leisure activities by means of the strictly time-based utility function, neither applying hollow space-time prisms nor activity differentiation nor utility function extensions.

#### Hollow space-time prisms and activity differentiation:

- **Configuration 1:** Leisure location choice based on hollow space-time prisms.
- **Configuration 2:** Analog to configuration 1, but shopping activities are further differentiated as grocery and non-grocery activities based on random assignment.

#### Utility function extensions:

- **Configuration 3.1:** Analog to configuration 2, but the utility function is extended to take into account the *store size* for grocery shopping.
- **Configuration 3.2:** Analog to configuration 2, but the utility function is extended to take into account the *stores density* in the neighborhood of the store in question for both grocery and non-grocery shopping.

## 3.2 Simulation Results

Figure 2 and Figure 3 show the average trip length and duration and the shopping trip distribution respectively, for all configurations compared with the Swiss microcensus. Figure 4 and Figure 5 present a comparison with traffic count data. To evaluate the effects with respect to the count data, the relative error of the daily traffic volume is for configuration  $i$  calculated as follows:

$$\epsilon_{rel}(i) = \frac{\sum_{hours} \sum_{measuredlinks} (volumes_{simulated,i} - volumes_{counted})}{\sum_{hours} \sum_{measuredlinks} volumes_{counted}}$$

The effects with respect to count data generated by the extension of the utility function are small and thus in general their relevance can be questioned. But by reason of a general underestimation of the traffic volume in our model (missing border-crossing traffic and freight traffic), the important effects in terms of the spatial distribution of the traffic can not be evaluated systematically at the moment (e. g. traffic volume shifts between count stations). This means that at the moment only the effects generated by trip elongations are revealed systematically. For an accurate evaluation of these effects it is necessary to introduce weighting by the traffic volume share that is generated by shopping trips. For that purpose the following measure is calculated:

$$\epsilon_{rel,max}(i, j) = \frac{(avg(distance_{shopping}, j) - avg(distance_{shopping}, i)) \times \#shopping\ trips}{w_i},$$

where  $avg(distance_{shopping})$  is the average shopping trip distance and  $w_i$  is the total traffic volume ( $\sum_{trip\ purposes} \#trips \times avg.\ trip\ length$ ) for configuration  $i$ . For two configurations  $i$  and  $j$  this measure gives the maximal expected relative difference in terms of count data due to average trip length changes.

### 3.2.1 Configuration 0

In e. g., Brunner and Mason (1968) it is shown that access in terms of travel time is an important determinant of shopping location choice. Nevertheless, the base case shows that modeling shopping and leisure activity location choice on the basis of travel times only, results in a strong underestimation of the average trip length and duration for shopping and leisure trips and hence also in an underestimation of traffic volumes as shown by the comparison with count data.

### 3.2.2 Configuration 1

As mentioned earlier in activity-based models, shopping location choice is directly influenced by leisure location choice, which means that a meaningful shopping location choice model requires sound modeling of leisure location choice. The following indicators enforce the assumption that incorporating hollow-space time prisms for leisure activities actually leads to a better leisure location choice model in comparison with the base case.

A substantial improvement in matching the count data especially for the evening peak hour (compare Figure 4(a) and Figure 4(b) or Figure 5(a) and Figure 5(b)) is achieved by using hollow space-time prisms. The relative error  $\epsilon_{rel}$  is strongly reduced (configuration 0: -60.25 % configuration 1: -36,43 %).

The microcensus trip statistics are also much better met for shopping trips (Figure 2 and Figure 3) and by construction also for leisure trips. In addition, for work and education trips a slightly better match with the microcensus trip statistics can be observed (e. g., average working trip length: microcensus: 8.58 km; configuration 0: 6.65 km; configuration 1: 6.88 km and average education trip length: microcensus: 6.83 km; configuration 0: 3.75 km; configuration 1: 4.99 km). This is due to the fact that, working and education activities are—although not being subject of the location choice process—influenced to a certain extent by the location choices for shopping and leisure activities.

### 3.2.3 Configuration 2

The differentiation of shopping activities into grocery shopping and non-grocery shopping based on random assignment has to be stated as *not productive* because only very small effects with respect to the average shopping trip length and also with respect to count data can be observed.

### 3.2.4 Configuration 3.1 and Configuration 3.2

The extension of the utility function shows positive—but small—effects with respect to the shopping trip statistics in comparison with the Swiss microcensus (Figure 2 and Figure 3).

The spatial load patterns of shopping facilities (both grocery and non-grocery) are also influenced. Figure 6 shows the shopping facility load differences between **configuration 2** and **configuration 3.2**. For illustration all stores of the larger region of Zurich, that exhibit a load difference (between configuration 2 and configuration 3.2) of more than 80 % are plotted in the map. It can be seen that the load increase has higher clustering and is concentrated to urban

regions, whereas load decrease can be observed more frequently in rural regions. The natural assumption that especially non-grocery shopping and weekly shopping are accomplished more frequently in urban regions or regions with a high stores density leads to the assumption that the results of configuration 3.2 (taking into account stores density) are more realistic than those of configuration 2. However, this assumption has to be validated in the future by means of customer frequencies data.

Regarding count data the relative error  $\epsilon_{rel}$  is slightly reduced (configuration 2: -36.36 %; configuration 3.1: -35.90 %; configuration 3.2: -35.91 %), where the small improvements of 0.46 % and 0.45 % respectively correspond with the maximal expected values  $\epsilon_{rel,max}(config2, config3.1) = 0.62\%$  and  $\epsilon_{rel,max}(config2, config3.2) = 0.39\%$  respectively.

### 3.2.5 All Configurations

In general the trip distances show a satisfying match with the Swiss microcensus, whereas the trip durations, although improved, are still strongly underestimated. It is expected that this is primarily caused by the missing intersection dynamics (e. g., traffic lights), car-following dynamics (including acceleration processes) and the missing border-crossing and freight traffic in our model. Additionally the access time to the transportation system (e. g., leaving the parking lots) is in our model systematically underestimated especially for coarse networks. On the other hand, the reported travel times in the microcensus have to be handled cautiously. Reported trip duration data are expected to be much less reliable than the trip distances, that are computed after geocoding the reported destinations. Better estimations of travel speeds and travel times are expected to be available after integrating the GPS-survey described in Hackney *et al.* (2007).

## 3.3 Conclusions and Outlook

As expected and demonstrated in the last section, modeling location choice for shop and leisure activities, that is based exclusively on travel times generates a substantial underestimation of the traffic volume.

Applying hollow space time prisms for leisure location choice is productive in the sense that it enables a more precise and thus more relevant evaluation of the shopping utility function extension and shopping activity differentiation.

Extending the utility function by the two described determinants is theoretically-based, natural and actually shows positive—but small—effects. Despite the small size of these effects, in

summary this means that the hypothesis that store size and store density in a given neighborhood are (weak) determinants of shopping location choice is *not falsified* by this test using a large real world scenario. Thereby, the potential power of MATSim for hypotheses testing is revealed.

However, to further improve the relevance, i. e., the significance, of the test the following issues have to be addressed in the future: First, the general underestimation of traffic due to the missing border-crossing traffic of foreigners and freight traffic in MATSim has to be corrected. Second, a reasonable statement of a null hypothesis supporting the testing of the determinants of shopping location choice has to be found. Third, count data is an aggregated measure which is of limited value for the validation of multi-agent simulations. Thus, in addition to the Swiss microcensus, further disaggregated data have to be integrated. As mentioned earlier, in MATSim only the motorized individual traffic of the Swiss population is modeled, whereas the traffic count stations record all of the motorized traffic. Including other means of transport in MATSim is thus expected to produce another improvement of the results with respect to the traffic count data.

### **3.3.1 Leisure Location Choice and Hollow space time prisms**

Constraining leisure location choice by hollow space time prisms, derived from empirical data is in essence an aggregated approach. As the goal of disaggregated methods, such as multi-agent simulations, is to provide arbitrary aggregation levels, the implications of this approach for disaggregated methods have to be researched. Another fundamental topic for further research is the question about the extent to which leisure location choice can be modeled on a disaggregated level and by means of utility maximization. This also includes the question about the implications of reducing leisure behavior that is subject to a high day-to-day variability to a cross-sectional sample.

Figure 2: Avg. trip travel distances and durations of shopping and leisure trips

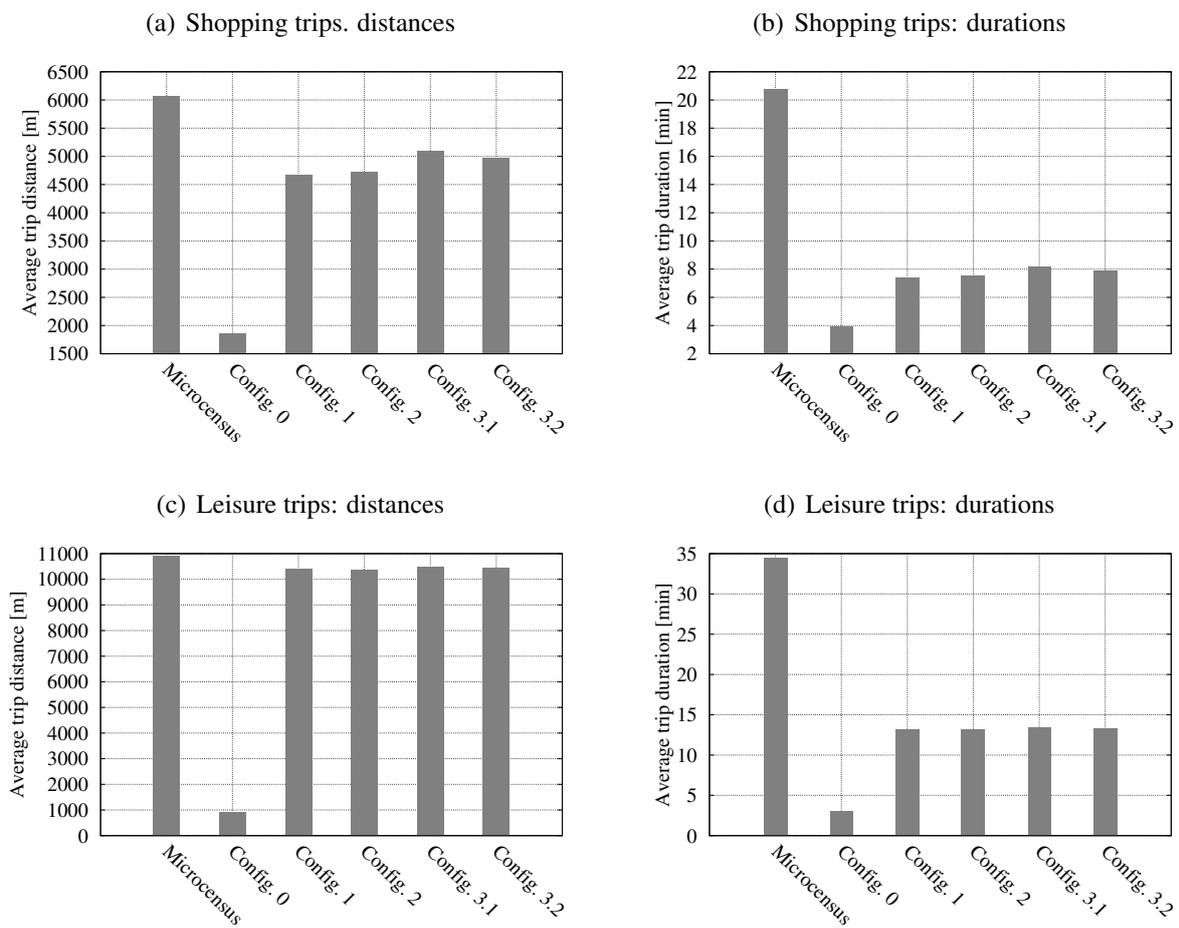


Figure 3: Shopping trip travel distance distribution

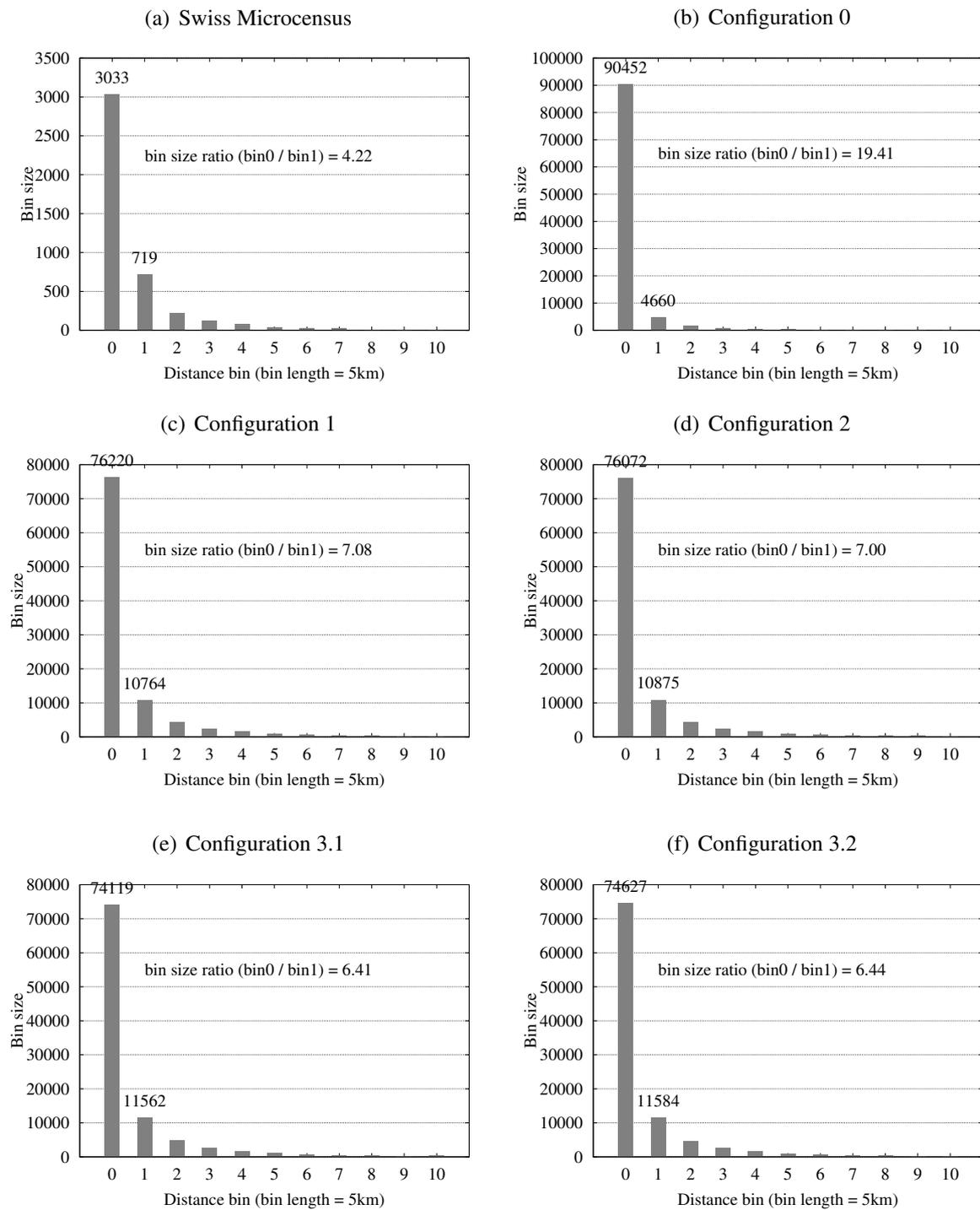


Figure 4: Counted volumes versus simulated volumes: evening peak (18:00 - 19:00)

---

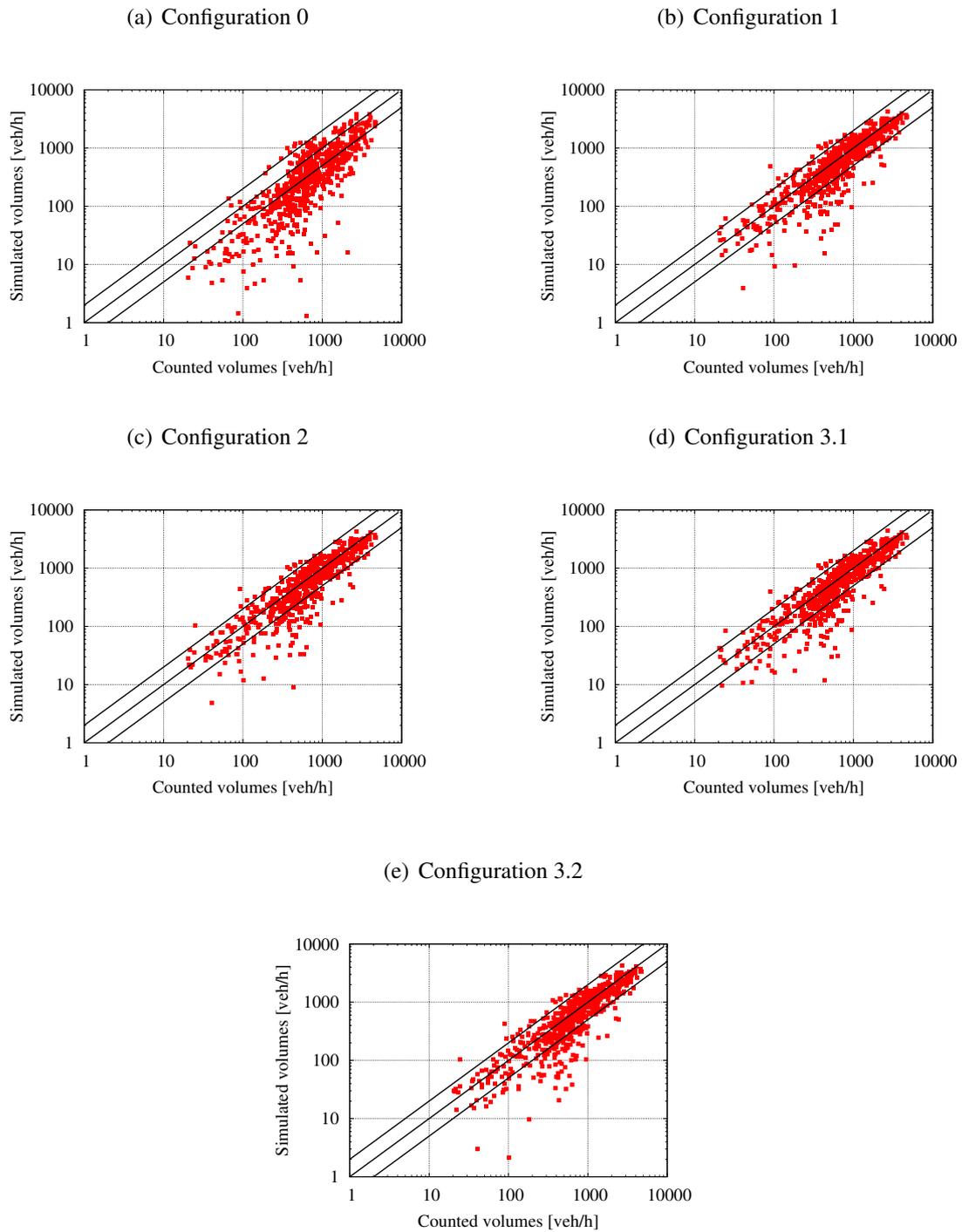


Figure 5: Count data stations: Hourly signed relative error and signed absolute error

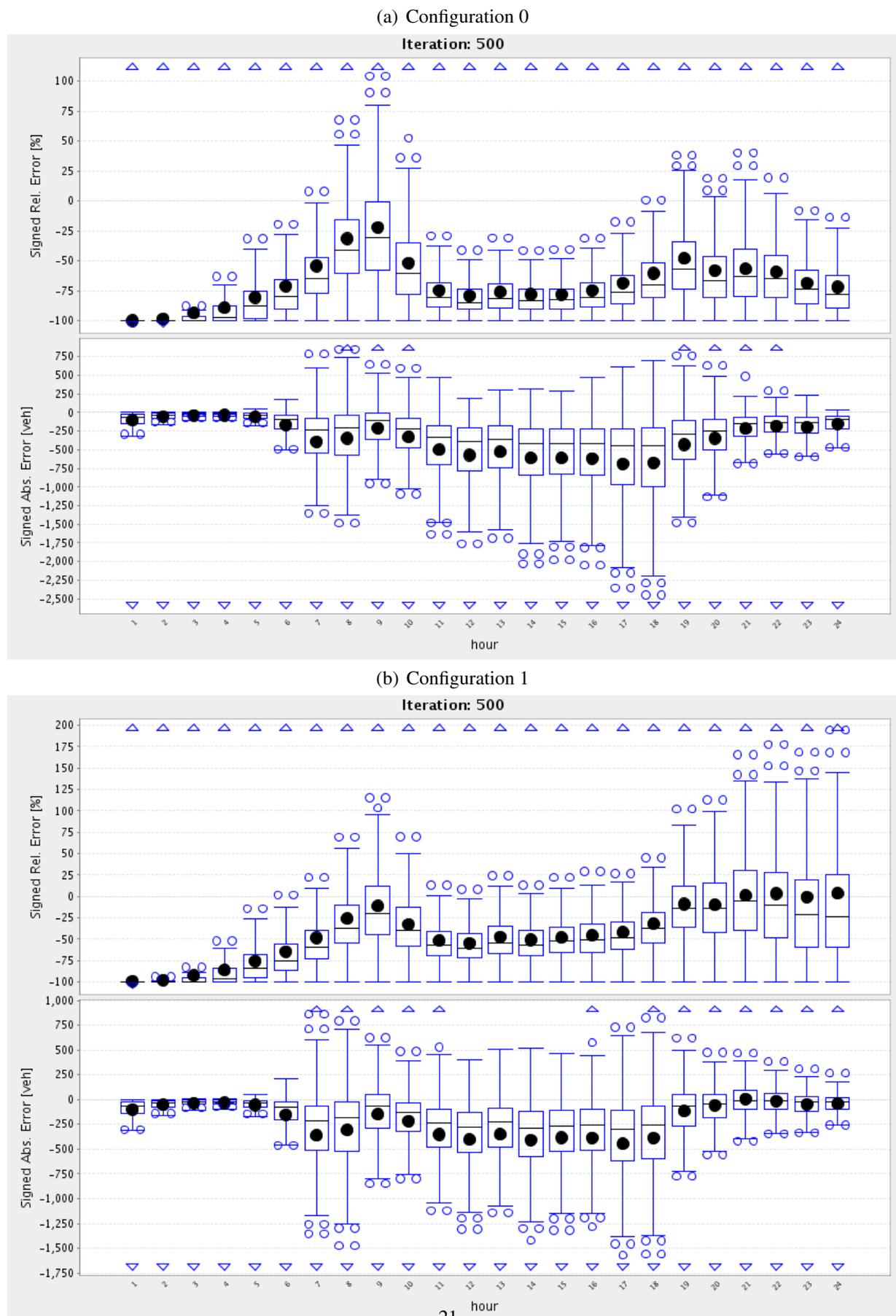
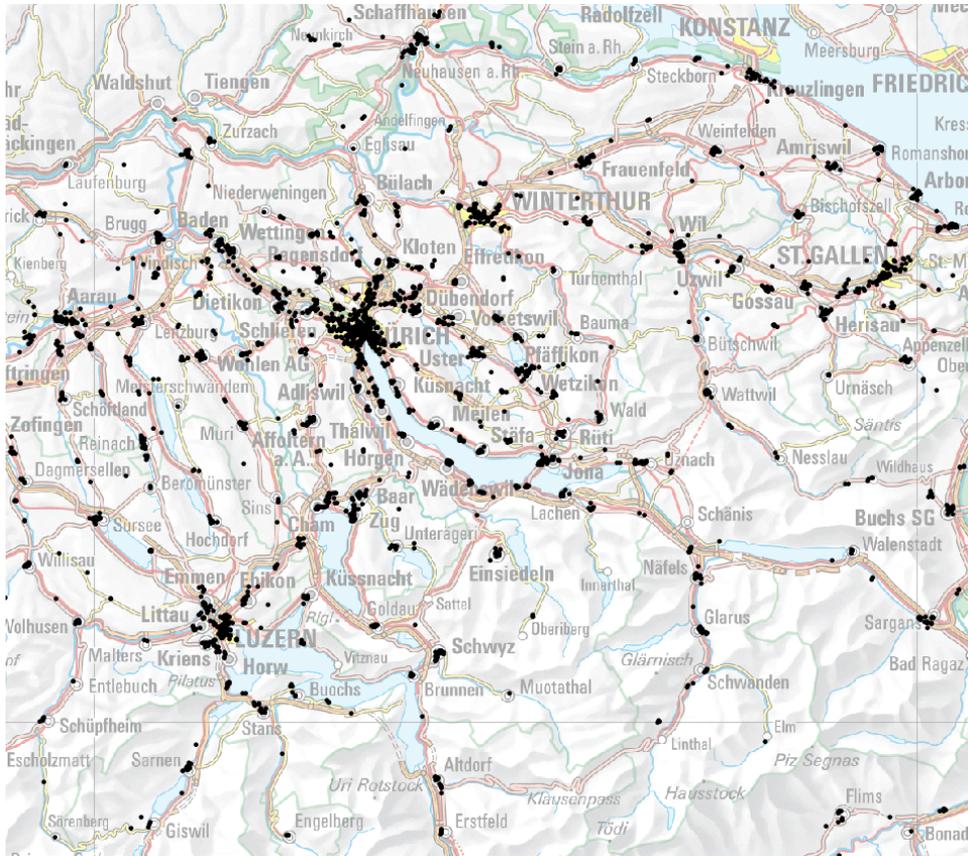
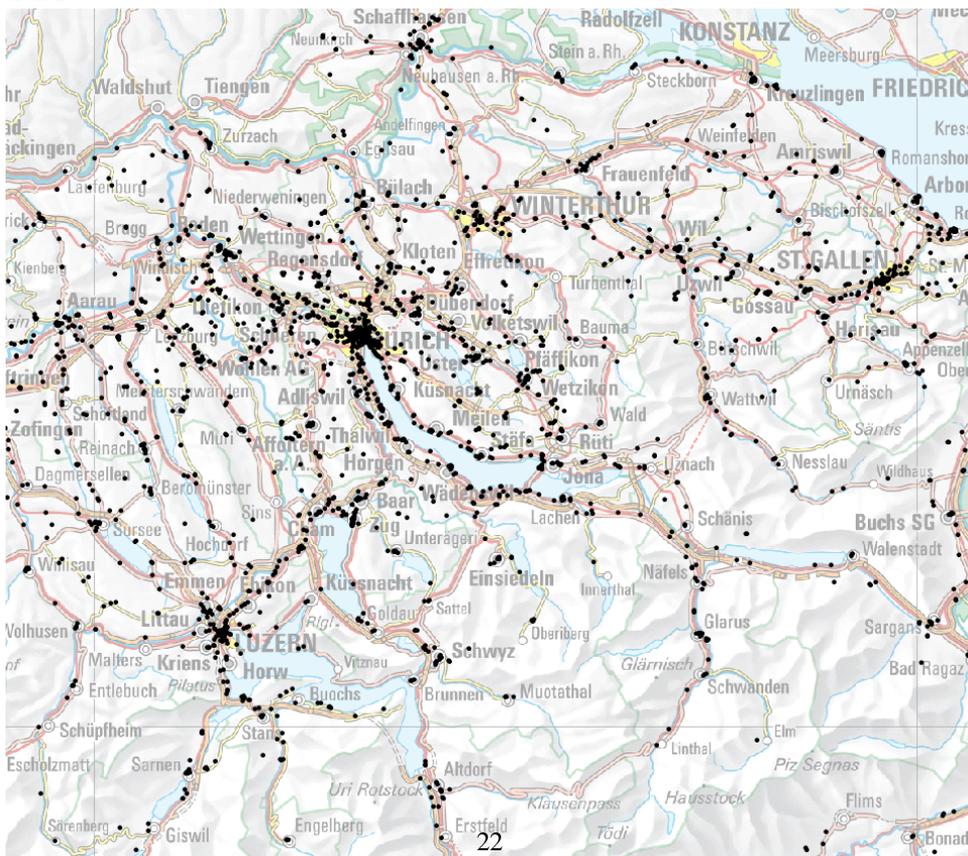


Figure 6: Shopping locations daily load changes (geodata ©swisstopo)

(a) Shopping locations with increased load in configuration 3.2 compared to configuration 2



(b) Shopping locations with decreased load in configuration 3.2 compared to configuration 2



## References

- Arnott, R., A. de Palma and R. Lindsey (1993) A structural model of peak-period congestion: A traffic bottleneck with elastic demand, *The American Economic Review*, **83** (1) 161–179.
- ASTRA (2006) Swiss Federal Roads Authority, webpage, <http://www.astra.admin.ch/>.
- Balmer, M. (2007) Travel demand modeling for multi-agent traffic simulations: Algorithms and systems, Ph.D. Thesis, ETH Zurich, Zurich, May 2007.
- Balmer, M., A. Horni, K. Meister, F. Ciari, D. Charypar and K. W. Axhausen (2009) Wirkungen der Westumfahrung Zürich: Eine Analyse mit einer Agenten-basierten Mikrosimulation, *Final Report*, Baudirektion Kanton Zurich, IVT, ETH Zurich, Zurich, February 2009, <http://www.ivt.ethz.ch/vpl/publications/reports/ab550.pdf>.
- Balmer, M., B. Raney and K. Nagel (2005) Adjustments of activity timing and duration in an agent-based traffic flow simulation, in H. J. P. Timmermans (ed.) *Progress in Activity-Based Analysis*, chap. 5, 91–114, Elsevier, Oxford.
- Brunner, J. A. and J. L. Mason (1968) The influence of driving time upon shopping center preference, *Journal of Marketing*, **32** (2) 57–61.
- Cetin, N. (2005) Large-scale parallel graph-based simulations, Ph.D. Thesis, ETH Zurich, Zurich.
- Charypar, D. and K. Nagel (2005) Generating complete all-day activity plans with genetic algorithms, *Transportation*, **32** (4) 369–397.
- Hackney, J. K., M. Bernard, and K. W. Axhausen (2007) Predicting road system speeds using spatial structure variables and network characteristics, *Journal of Geographical Systems*, **9** (4) 397–417.
- Hägerstrand, T. (1989) Reflections on “What about people in regional science?”, *Papers of the Regional Science Association*, **66**, 1–6.
- Horni, A., D. M. Scott, M. Balmer and K. W. Axhausen (2008) Location choice modeling for shopping and leisure activities with MATSim: Combining micro-simulation and time geography, *Working Paper*, **527**, IVT, ETH Zurich, Zurich, <http://www.ivt.ethz.ch/vpl/publications/reports/ab527.pdf>.
- Iragael, J. (2007) The role of travel time budgets - representation of a demand derived from activity participation., paper presented at *the 11th World Conference on Transportation Research*, Berkeley, June 2007.

- Kitamura, R., C. Chen and R. Narayanan (1998) Traveler destination choice behavior: Effects of time of day, activity duration, and home location, *Transportation Research Record*, **1645**, 76–81.
- Lademann, R. (2007) Zum Einfluss von Verkaufsfläche und Standort auf die Einkaufswahrscheinlichkeit, in M. Schuckel and W. Toporowski (eds.) *Theoretische Fundierung und praktische Relevanz der Handelsforschung*, 143–162, Springer, Heidelberg.
- Lefebvre, N. and M. Balmer (2007) Fast shortest path computation in time-dependent traffic networks, paper presented at *the 7th Swiss Transport Research Conference*, Ascona, September 2007, [http://www.strc.ch/conferences/2007/2007\\_lefebvre.pdf](http://www.strc.ch/conferences/2007/2007_lefebvre.pdf).
- MATSim-T (2008) Multi Agent Transportation Simulation Toolkit, webpage, <http://www.matsim.org>.
- Meister, K., M. Rieser, F. Ciari, A. Horni, M. Balmer and K. W. Axhausen (2009) Anwendung eines agentenbasierten Modells der Verkehrsnachfrage auf die Schweiz, *Straßenverkehrstechnik*, **53** (5) 269–280.
- Nash, J. (1951) Non-cooperative games, *The Annals of Mathematics*, **54** (2) 286–295.
- Recker, W. W. and L. P. Kostyniuk (1978) Factors influencing destination choice for the urban grocery shopping trip, *Transportation*, **7**, 19–33.
- Rieser, M., K. Nagel, U. Beuck, M. Balmer and J. Rügenapp (2007) Agent-oriented coupling of activity-based demand generation with multiagent traffic simulation, *Transportation Research Record*, **2021**, 10–17.
- Robinson, R. and R. Vickerman (1976) The demand for shopping travel: a theoretical and empirical study, *Applied Economics*, **8** (4) 267–281.
- Scott, D. M. (2006) Constrained destination choice set generation: A comparison of GIS-based approaches, paper presented at *the 85th Annual Meeting of the Transportation Research Board*, Washington, D.C., Jan, Oliver 2006.
- Swiss Federal Statistical Office (2000) Eidgenössische Volkszählung 2000, [http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen\\_quellen/blank/blank/vz/uebersicht.html](http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen_quellen/blank/blank/vz/uebersicht.html).
- Swiss Federal Statistical Office (2001) *Eidgenössische Betriebszählung 2001 - Sektoren 2 und 3*, *GEOSTAT Datenbeschreibung*, Swiss Federal Statistical Office, GEOSTAT, Neuchatel, <http://www.bfs.admin.ch/bfs/portal/de/index/dienstleistungen/geostat/datenbeschreibung/betriebszaehlung05.Document.111424.pdf>.

Swiss Federal Statistical Office (2006) *Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten*, Swiss Federal Statistical Office, Neuchatel.

Vickrey, W. S. (1969) Congestion theory and transport investment, *The American Economic Review*, **59** (2) 251–260.

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