D6.2 Zurich Case Study

Project acronym : EUNOIA
Project title : Evolutive User-centric Networks for Intraurban Accessibility
Grant Agreement number : 318367
Funding Scheme : Collaborative project
Project start date / Duration: 01 Oct 2012 / 26 months
Call Topic : FP7-ICT-2011-8
Project web-site : http://www.eunoia-project.eu/
Deliverable : D6.2 Zurich Case Study
Edition : 1
Date : 08/01/2014
Status : Final
Dissemination Level : Public
Authoring and Approval

Prepared by
Thibaut Dubernet (ETH – Researcher)
Kay Axhausen (ETH – Principal Researcher)

Reviewed by
José Ramasco (Deputy Scientific Coordinator)
Ricardo Herranz (Deputy Management Coordinator)

Approved by
Maxi San Miguel (Project Coordinator)

Record of Revisions

<table>
<thead>
<tr>
<th>Edition</th>
<th>Date</th>
<th>Description/Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edition 1.Draft 1</td>
<td>04/12/2014</td>
<td>-</td>
</tr>
<tr>
<td>Edition 1</td>
<td>08/01/2015</td>
<td>Updated Bike Sharing results + minor reformulations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Approved for submission to the EC</td>
</tr>
</tbody>
</table>
# Table of Contents

1. **INTRODUCTION** ......................................................................................................................... 3

2. **BUILDING THE CASE STUDY** ..................................................................................................... 5
   1.1 TRAVEL DEMAND FOR ZURICH ................................................................................................. 5
   1.2 THE 2005 DEMAND .................................................................................................................... 5
   1.3 THE 2010 DEMAND ..................................................................................................................... 5
   1.4 SYNTHETIC SOCIAL NETWORK ................................................................................................. 6
   1.5 FROM SWITZERLAND TO ZURICH ............................................................................................. 8
   1.6 SUPPLY NETWORK: REPRESENTING THE TRANSPORT INFRASTRUCTURE FOR ZURICH ............. 8
   1.7 CALIBRATING MATSIM FOR ZURICH ......................................................................................... 10
   1.8 VALIDATING MATSIM FOR ZURICH ........................................................................................... 12

3. **SCENARIOS** ................................................................................................................................. 18
   1.9 HOUSEHOLDS ............................................................................................................................ 18
      1.9.1 Objective of the Scenario ..................................................................................................... 18
      1.9.2 Outputs of the Scenario ....................................................................................................... 19
   1.10 EXPLORING THE INFLUENCE OF MODELING SOCIAL CONTACTS ........................................... 21
      1.10.1 Objective of the Scenario .................................................................................................... 21
      1.10.2 Outputs of the Scenario ...................................................................................................... 22
   1.11 EVALUATION OF A BIKE SHARING SYSTEM ............................................................................ 28
      1.11.1 Objective of the Scenario .................................................................................................... 28
      1.11.2 Outputs of the Scenario ...................................................................................................... 30

4. **OUTLOOK** ................................................................................................................................. 38

ANNEX I. A SHORT PRESENTATION OF MATSIM ............................................................................... 39
   1.12 PHILOSOPHY ............................................................................................................................ 39
   1.13 MODELING PRINCIPLES: GAME THEORY AND CO-EVOLUTIONARY COMPUTATION .................... 39
   1.14 THE PHYSICAL WORLD IN MATSIM ....................................................................................... 41
   1.15 TO SUM UP ............................................................................................................................. 41

5. **REFERENCES** ............................................................................................................................ 42
1. Introduction

Zurich is one of the three case study cities (together with London and Barcelona) included in EUNOIA WP6 “Case studies”. The purpose of WP6 was to bring the data collected and analysed in WP3 “Data collection” and WP4 “Data analysis”, respectively, and the models and tools developed in WP5 “Modelling and simulation platform” to the point where stakeholders can employ them to assess different mobility policies. More specifically, the goal was to demonstrate the capabilities of the agent-based transport simulation model MATSim — implementing the paradigms of activity-based modelling and dynamic traffic assignment — enhanced with the results obtained by EUNOIA to test specific mobility policies at the three cities participating in the project either as partners (Barcelona) or through the EUNOIA Advisory Board (London and Zurich). After discussion with different representatives of the Mobility Department of the Barcelona City Council, Transport for London and the Division of Transport at the City of Zurich, we decided to focus the simulation effort on the analysis of different configurations of the public bike sharing schemes of the three cities. The reason for this choice is twofold:

- From a scientific/technical perspective, the selected case study has allowed to test some of the main results of the previous work packages:
  - In terms of the data sources used to build, calibrate and validate the model, one of our objectives was to make use of some of different combination of conventional data (e.g., travel surveys) and the non-conventional, ICT-based data collected and analysed in WP3 and WP4. In this sense, the London case study provides an application for the active use of the Oyster card data sample made available by TfL, as well as to explore the usage of data collected on the London bike sharing scheme; the Barcelona case study includes two main novelties: the use of phone call records for generating the agents’ activity plans, and the use of the data from the Barcelona public bike sharing smart card for the calibration of the travel demand model; and finally, the Zurich case study is built using more conventional, survey data — but uses the results of an unconventional snowball sample to generate a synthetic social network, to be used in the simulation. Such a survey was performed in Barcelona within the EUNOIA Project.
  - In terms of modelling and simulation, we have used a novel MATSim bike-sharing module developed in the context of WP5, which to our knowledge constitutes one of the first attempts to include public bike sharing systems in travel demand models.
  - Finally, the case studies have also made use of the new visualisation tools developed in WP5 with the aim to provide clear and meaningful interactive information about public bike sharing usage flows.

- From a policy perspective, this is a subject of interest for the three cities: the public bike sharing schemes of Barcelona (Bicing) and London (popularly known as Boris Bikes, after Boris Johnson, who was the Mayor of London when the scheme was launched) have been in operation since 2007 and 2010 respectively, and the need to optimise the operation of the system (e.g., the bike relocation strategies) is a common concern for both cities; in the case of Zurich, there is not a public bike sharing system yet, but there is a plan for its implementation. Therefore the subject allows the analysis of problems that are common for London and Barcelona, as well as of the differences between both systems, and can in turn be useful to inform the deployment of the Zurich system.
This report presents the building, calibration and results of the Zurich EUNOIA case study. The aim of this case study is twofold:

- benchmark the developments of WP5 for real-world scenarios
- study policy questions, in particular bike sharing.

The building of the scenario is part of a continuous effort to build and maintain a Zurich scenario, for use across different projects and studies. A novelty of this scenario is the inclusion of a synthetic social network, enabling the explicit representation of social decision dynamics within the simulations, and their influence on mobility behavior.
2. Building the Case Study

2.1 Travel Demand for Zurich

The Zurich case study uses two different travel demands, due to different needs:

- An old travel demand, generated for a former project, using 2005 data. The advantage of this demand is that it is based on the last full Swiss census, containing household membership information. It is thus used to test the joint decision developments of WP5 for the household case.

- A newly generated travel demand, using the most recent data. This new demand does not include households, but does contain a synthetic social network, using the approach from Arentze, Kowald & Axhausen (2013).

2.2 The 2005 Demand

The initial demand comes from the full-Switzerland scenario described by Meister et al. (2010). It was generated by allocating activity chains from the national travel survey from the years 2000 and 2005 to records from the national census 2000, which is a 100% survey of the Swiss population, containing in particular information on home location at the hectare level and work location at the municipality level, as well as household membership information. The agents are grouped according to the household information from the census, and only the households having at least one member performing a trip passing less than 30km from the Bellevue Place, in the center of Zurich, are kept. A sample of 10% of those households is used for the simulation. This results in a scenario containing 206943 agents, grouped in 88439 households, and performing a total of 788931 trips.

2.3 The 2010 Demand

Between years 2005 and 2010, the Federal Office of Statistics stopped collecting detailed data for every Swiss resident for the census. The available data now consists of a full register survey, complemented by detailed questionnaires for 1% of the Swiss population. The register survey contains information from the civil registry, in particular home location at the hectare level and basic socio-demographic attributes.

The National Travel Survey is still performed every five years, and results from 2010 are used here.

The register survey and the National Travel Survey were combined using statistical matching, using age, gender and a nine-level commune classification, to obtain an allocation of activity chains to each entry of the register survey. The process is described in Müller & Axhausen (2014).

Activity chains consist of a sequence of activity types (home, work, education, leisure and shopping), initial modes for the trips between those activities, and desired duration for each activity type. Modes only serve as a starting solution for the MATSim optimisation process; durations (corresponding to the durations reported in the National Travel Survey) are used as parameters for the scoring function, ensuring that agents do perform each activity type for a duration roughly equivalent to the reported durations.
Once each agent has been allocated an activity chain, those activities need to be located into space. Home activities are simply located according to the Register Survey. Work location are allocated the following way:

1. agents are allocated a random work commune, such that municipality-to-municipality commuting flows reproduce the commuter matrix
2. within a commune, the exact location is sampled randomly, considering work place capacity constraints from the enterprise census.

In the absence of data, education location was chosen as the closest known facility providing education.

Initial location for shopping and leisure were sampled randomly using a space-time prism approach, using the surrounding mandatory (home, work, education) activity locations. In addition, leisure location choice is performed during the MATSim simulation, using individual preferences, following the approach of Horni (2013).

Individuals living in Switzerland are not the only ones to use the Swiss transport system. To take this fact into account, agents reproducing the data from OD matrices for freight and cross-border traffic were generated for a previous project (Shah, 2010). This background traffic is used in the current scenario, as it is.

### 2.4 Synthetic Social Network

An innovation of the scenario used in this study is the integration of a synthetic social network, allowing to test for the first time the influence of representing joint decision processes on the realism of the results, in a real-world scenario.

This synthetic social network was generated using the approach from Arentze et al. (2013). This model was estimated using data from a snowball sample collected in Switzerland (Kowald & Axhausen, 2012). The results of this survey gave a valuable description of the leisure contact network in Switzerland, using an iterative approach:

- a random sample of individuals were contacted, and asked to name persons with whom they often performed leisure activities
- those persons were then contacted, and asked the same questions.

In total, 5 iterations were performed.

This survey approach was reproduced in Barcelona during the EUNOIA project. Unfortunately, the new results obtained for Barcelona could not yet be compared and used here, for two reasons:

- The survey was completed close to the end of the project, letting only little time to use it
- Response rate in Barcelona was much lower than in Switzerland, making the accuracy of the social network topology questionable.
The model of Arentze et al. (2013) uses this data to estimate a model able to generate a synthetic leisure contact network from population data. This model was however until now only applied for populations way smaller than the usual number of agents in a simulation scenario.

The model from Arentze et al. (2013) takes into account homophily and transitivity in the following way:

- each potential tie is associated a (random) utility, which depends on age difference, gender difference and geographic distance between home locations
- the probability that a given tie exists is the probability that its utility is greater than a calibrated threshold.
- a second round is performed, considering only individuals with common contacts between which no tie exists: a tie exists then if the utility of the tie is higher than a reduced threshold.

The two thresholds include both the willingness to form a friendship, as well as the probability for two individuals to actually meet. Thus, not all potential friendships need to be evaluated: one can only evaluate a sample of those, and calibrate the thresholds to obtain target average degree and clustering coefficient.

Arentze et al. (2013) demonstrated the quality of the social network their models generates, by comparing the results of their model with the output of a null model. They however performed their tests on a sample population containing 3301 individuals. They tested different sampling rates, from 7.5% to 15%, and observed that the distance distribution of ties tends to be better reproduced with low sampling rates — when considering too many friendships, the model tends to over-represent ties with agents living close to each other.

The Swiss 10% synthetic population however contains 736894 synthetic individuals, which is 223 times more, and results in more than $2 \times 10^{11}$ potential (bidirectional) ties. This forces to use much lower sampling rates than Arentze et al. (2013), and the quality of the results in this case have to be assessed. The highest sampling rate allowing to get the results in reasonable time was achieved with a sampling rate of 0.05% (1 tie over 2000). Comparison of the Switzerland-wide synthetic network with data from the snowball sample is presented in table. The restriction of the social network to the Zurich area is presented in the next section. Overall, the statistics for the two sampling rates analyzed (0.025% and 0.05%) are pretty similar. Clustering and average number of social contacts are well reproduced, as the two thresholds of the models were calibrated to this end. Homophily is however lower than in the data, in terms of age and gender: the sampling rate does not allow to test enough “good” ties, and the calibration of the threshold thus leads to the selection of less desirable ties. This interpretation applies also to the highest home-to-home distance between alters. A complete re-writing of the code, allowing faster calibration and allowing to leverage the power of parallelization, is being implemented. It could unfortunately not be prepared early enough for use in the current study, and this yet-to-perfect synthetic network is thus used here.
Table 1. Synthetic Social Networks Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Snowball</th>
<th>Synth (0.025%)</th>
<th>Synth (0.05%)</th>
<th>Synth (ZH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering</td>
<td>0.206</td>
<td>0.195</td>
<td>0.190</td>
<td>0.187</td>
</tr>
<tr>
<td>Avg. Degree</td>
<td>22.0</td>
<td>22.0</td>
<td>22.0</td>
<td>20.6</td>
</tr>
<tr>
<td>Same Age Ratio (%)</td>
<td>46.3</td>
<td>30.2</td>
<td>30.7</td>
<td>29.4</td>
</tr>
<tr>
<td>Same Gender Ratio (%)</td>
<td>61.7</td>
<td>56.1</td>
<td>56.5</td>
<td>55.7</td>
</tr>
<tr>
<td>Avg. Distance (km)</td>
<td>26.6</td>
<td>54.3</td>
<td>49.1</td>
<td>17.8</td>
</tr>
<tr>
<td>Avg. Social Distance</td>
<td></td>
<td>4.85</td>
<td>4.83</td>
<td>4.4</td>
</tr>
</tbody>
</table>

2.5 From Switzerland to Zurich

The case studies presented here focus on Zurich, and thus do not need to simulate mobility in all of Switzerland. To this end, an area of interest, defined as a 30 km circle centered on Bellevue Place in central Zurich, is defined. Only agents going through this area at least once during the day are retained in the final scenario. This includes agents having trips originating and/or ending in the area, as well as agents having to cross this area at least once during the day. This results in a reduced social network, where only social contacts being part of this reduced population are retained. Keeping all social contacts of agents part of the reduced population does not make much sense, as it results in most agents being retained, resulting in no big advantage over a full-Switzerland scenario in terms of computation time. The statistics of this reduced network are presented in the last column of table.

Interestingly enough, this reduction of the network does not dramatically changes topological properties of the network: clustering remains in the same range, agents only lose 2 contacts on average, and the average social distance remains pretty similar. Homophily statistics also remain quite stable. Only average home-to-home distance is strongly decreased, which is not surprising given that the average home-to-home distances produced by the model is close to the diameter of the area of interest (60 km).

2.6 Supply Network: Representing the Transport Infrastructure for Zurich

Both variants of the demand are simulated with the same supply, build with 2010 data.

Two versions of the road network are used. The first version is a planning network, from the National Transport Model (Vrtic, Fröhlich & Axhausen, 2003). It models the Swiss network at a coarse resolution, as well as the major arterials in the neighboring countries. The second, more detailed one, is a navigation network, representing all roads. It is used for the bike sharing studies, for which network
information is used for walk and bike, requiring accuracy for short distances. The two networks are represented on Figure 1.

Figure 1. Planning and Navigation Networks

The public transport schedule from the Cantonal Transport Model is used to get realistic travel time estimates (Amt für Verkehr, 2011). Vehicles with a capacity are actually simulated, and agents have to wait if a vehicle is full. However, vehicles do not suffer from congestion, as they are simulated on a separate network.

In addition, “facilities” containing the information about opening times for different activity types, and roughly correspond to buildings, are used. Data comes from the federal enterprise census. Finally, a bike sharing system, that is planned for the city of Zurich, is represented. It is planned to establish the system in 3 steps:

1. 100 stations (1500 bikes) in the city center
2. 100 additional stations (1500 bikes) to densify the North part of the city
3. 100 additional stations (1500 bikes) extending to further residential areas
As in most current bike sharing systems, the fare is planned to be time-dependent, with the first 30 minutes being free, before increasing every 30 minutes. Analysis of the usage of current systems showed that this results in a large majority of trips being below this time limit — 92% for instance in Paris (Nair, Miller-Hooks, Hampshire & Bušić, 2013). Hence, though it is difficult to calibrate a scenario to represent prices accurately, an approximate representation should be enough to reproduce this effect. The planned fares are included assuming that one Swiss Franc corresponds to one utility unit.

### 2.7 Calibrating MATSim for Zurich

The scenario was calibrated according to the following statistics:

- Mode shares
- Trip distance distribution per mode
- Traveled distance for leisure purpose

Figure 2 shows the cumulative mode share per distance class. The overall fit of the base model (last distance category) is pretty good. However, the scenario exhibits too few short trips, which has an influence on the accuracy of mode shares for shorter distances.
Figure 3 shows the trip distance distribution per mode. The over representation of long trips in the simulation is reflected by longer trips for all modes. The most significant difference occurs for public transport.

Figure 4 shows the trip distance distribution per purpose. For aggregate purposes (from “all” to each type and from each type to “all”), fit is rather satisfying. Significant differences can however be found for specific type-to-type categories. In particular, distances between leisure and home/education are dramatically overestimated. In this run with leisure destination choice, preference for a leisure location comes from a trade-off between travel disutility and a random utility associated to each agent-location pair. Though this approach allows to fit overall travel distribution pretty well, it ignores the possible specificities of different activities, for instance the possible preference for sport facilities or pubs close to a university, that allow students to perform those activities with their friends. This is the idea behind the inclusion of social networks in the simulation.
2.8 Validating MATSim for Zurich

In order to validate the results, the output of the base scenario are compared here with counting stations. This scenario is run with the planning network: it is to be expected that count values are higher than in reality, as one link in the simplified network “represents” several actual links.

It is to be noted that this data was not used for calibration, and thus can help assess the quality of the parameters, without a fear of overfitting.

Figure 5 shows the relative error of simulated counts compared to measured counts, with time of day. As expected, most stations overestimate counts, by a factor roughly 1.5. Traffic early in the morning and late in the afternoon tends to be overestimated.
Figure 5. Error on counts with time of day

Figures 6 and 7 show the example of counts for “good” stations, with time of day. In those counting stations, not only are the overall values well reproduces, but the dynamics of traffic also fit pretty well.
Figure 6. Example of counts for a “good” station: link 100414
Figure 7. Example of counts for a “good” station: link 100757

Figures 8 and 9 show the example of stations with dramatically overestimated traffic. The first station still reproduces the temporal pattern of traffic. However, there are also stations such as the second one, where both volumes and temporal dynamics are off.
Figure 8. Example of counts for a “bad” station: link 101207
Figure 9. Example of counts for a “bad” station: link 101203
3. Scenarios

This section presents and analyses simulation results, with two purposes:

- Benchmark the developments of WP5
- Explore policy questions, namely the establishment of a Bike Sharing scheme.

3.1 Households

3.1.1 Objective of the Scenario

The aim of this scenario is more to benchmark the WP5 development than to assess the viability of a policy. It aims at testing the joint decision simulation framework, developed as a part of WP5, in the case of households.

The desire to simulate joint decisions within households comes from a large body of literature, that makes clear the importance of household processes in mobility behavior. Examples include Junyi Zhang, Timmermans & Borgers (2005); Jun Zhang, Timmermans & Borgers (2007); Kato & Matsumoto (2009); Bradley & Vovsha (2005); Gliebe & Koppelman (2005); Gliebe & Koppelman (2002); Ho & Mulley (2013); Vovsha & Gupta (2013); Golob & McNally (1997); or Golob (2000). In particular, Ho & Mulley (2013), for Australian data, show high joint household activity participation on weekends, and a high dependence of joint travel on trip purpose and household mobility resources.

For the Swiss case, Stauffacher, Schlich, Axhausen & Scholz (2005) analyse a twelve-week travel diary and find that the share of joint travel with household members is particularly high for some activity purposes.

This scenario uses the 2005 demand with households. A limitation of this demand is that though the households come from data, and are thus realistic, the activity chains allocated to each agent are sampled independently — that is, the activity chain of an agent does not depend on the activity chains of other agents in the household. This is mainly due to the unavailability of a Swiss household travel diary with a sample large enough, but might be seen as a limitation of this scenario, when attempting to identify joint trips or activities.

During the simulation, in addition to the usual behavior, agents can re-order their activities, choose a location for a leisure activity together with household members, and plan joint trips with household members.

Analysis of those results is presented in more details in Dubernet & Axhausen (forthcoming), currently under review, as well as Dubernet & Axhausen (2014).
3.1.2 Outputs of the Scenario

Figure 10 shows the travelled distances per mode, with and without allowing agents to change leisure location, as well as in the National Travel Survey. Travelled distances are pretty similar in both cases. Very short trips are under-represented for drivers with a passenger ("driver" mode), but travelled distances remain pretty close to the observed ones. This is not the case for passengers, that get driven much shorter distances than in the data. This implies that though drivers do not drive much more, the part of their trips actually performed with a passenger is much lower than in the data. This might come from the under-representation of fully joint trips, that is, trips performed between the same origin and destination, probably due to the independence between agents when allocating daily plans, as highlighted above.

![Distance Distribution per Mode](image)

**Figure 10. Distance Distribution per Mode**

Figures 11 and 12 show the proportion of passenger trips that are performed for each purpose, in the National Travel Survey and in the simulation. Data from the National Travel Survey does not differentiate between traveling with household members or other persons — the graph thus contains all kind of social relationships. An interesting feature of the simulation output is the emergence of escorting.
to education, with roughly 30% of passenger trips being performed between education and home. Those trips are not enforced in any way, but rather emerge from the constraints specific to households: children do not own a car, and are part of the “contacts” that the parents are willing to drive.

Figure 11. Repartition of passenger trips per purpose: NTS
Though accuracy is lacking, those results are encouraging for the approach: in particular, the emergence of escorting to school in the absence of specific parameter to this end is a strong point of micro-simulation, as it allows to represent the constraints that lead to this result.

### 3.2 Exploring the Influence of Modeling Social Contacts

#### 3.2.1 Objective of the Scenario

As the precedent, this scenario aims at benchmarking the developments of WP5, focusing on the influence of social contacts on the results of a simulation, using the developments from WP5.

The study of other kinds of social contacts is more limited, most probably due to the difficulty to collect reliable data. Several studies however focused on the influence of social contacts, and joint activities, on
travel behavior (Arentze, Kowald & Axhausen, 2012; J. A. Carrasco & Habib, 2009; e.g. J. A. Carrasco, Miller & Wellman, 2008; Deutsch & Goulias, 2013; Habib & Carrasco, 2011; Moore, Carrasco & Tudela, 2013). Such studies reveal significant influences of spatial distribution of social contacts, or of the number of participants in a joint activity, on the traveling behavior of the participants. This is particularly important for leisure purpose trips, which represent an increasing share of the trips performed in western countries (Axhausen, 2005; Schlich, Schönfelder, Hanson & Axhausen, 2004).

In addition to the usual behavior, agents can re-order their activities, choose a location for a leisure activity together with social contacts, and plan joint trips with social contacts.

In addition, the willingness of individuals to meet their social contacts is represented by an additional term in the scoring function. Representing this willingness represents challenges, and requires formulating hypotheses on why individuals meet contacts. Two complementary interpretations are possible:

- individuals value higher time passed with friends than alone
- the purpose of some activities is to meet friends, and time passed without friends in such activities is perceived as waiting time.

This second formulation was found to lead the better results. In the implementation, for each leisure activity, each agent has a desired number of social contacts to meet. Time passed with less contacts then desired results in a penalty, driving agents to minimize this time — which in turn pushes agents to travel together to their leisure activities, to avoid this undesirable waiting time.

### 3.2.2 Outputs of the Scenario

Figure 13 shows the trip distance distribution per purpose, in two cases: with all agents desiring at least one co-participant in leisure activities (scenario $g$), and without any preference for group size (scenario $ng$). The most important differences with the base case can be observed for leisure purpose. Desire for co-participants pushes agents to go slightly further. An interesting fact is that the strong over-estimation of traveled distances between leisure and home or education does not appear anymore in this scenario. This is probably due to the fact that here, willingness to travel further does not come from noise put on the utility of an activity, which results de facto in an admissible additional travel time an agent is willing to endure to access this location, whatever the minimal traveled distance might be. Here, it comes from a tradeoff between travel and willingness to meet social contacts: an agent will not have the incentive to increase too much the traveled distance for a short trip, if he has the possibility to meet a social contact in an activity on the way.
Figure 13. Distance Distribution per Purpose, all modes

Figure 14 shows the same kind of information, only for trips performed as a car passenger. Desire for co-participants has here a strong impact on the accuracy of traveled distance: trips performed as a passenger are much longer in this case, and much closer to the observed data. Without this desire, agents just pick-up and drop-off contacts on their way, mostly for very short trips. This results is again encouraging for the stream of work looking at the usage of microsimulation to predict joint travel: car passenger traveled distance, and, thus, car occupancy, can result from desires about the activities, represented explicitly.
Figure 15 shows the traveled distances per mode in the two scenarios. Contrary to the household case, differences with the data are important for the “driver” mode. This might come from the too high home-to-home distances between social contacts, as highlighted above. A good point is the influence of the willingness to meet contacts at leisure activities, that results in a very accurate traveled distance distribution for the *passenger* mode. This comes, as already highlighted, from the willingness of social contacts to arrive *together* at the joint activity, making it more attractive for an agent to be driven, even if it implies performing other trips by public transport or walk, rather than by car.
Figures 16 and 17 show the proportion of passenger trips that are performed for each purpose, in the National Travel Survey and in the simulation with desire for co-participants. The desire for co-participants allows to reproduce the observed pattern that most joint travel occurs for leisure purpose.
Figure 16. Repartition of passenger trips per purpose: NTS
Figure 17. Repartition of passenger trips per purpose: Simulation

The scenario however has flaws, in particular a low share of joint travel. This might in part come from the flaws in the social network, addressed in the first part. This problem of the social network generation — namely the computational efficiency, that forbids testing enough friendships — is being tackled, and will hopefully help improve this problem.

Those results are encouraging, this approach helping further refine the predicted traveled distances, as well as allowing a first glance at joint traveling and car occupancy in micro-simulation models.
3.3 Evaluation of a Bike Sharing System

3.3.1 Objective of the Scenario

This scenario is the most applied. It aims at testing a planned bike sharing system in the city of Zurich.

In the recent years, more and more cities implement this kind of scheme, where a set of bikes is available for short-term rental between an origin and a destination station. Shaheen, Guzman & Zhang (2010) estimated the number of cities implementing such a system to be 125 worldwide, for a total of around 140000 bicycles. Though it is sometimes marketed as “ecologically-friendly”, based on the simple assumption that trips performed by bike-sharing would have been performed by car, the impact of bike sharing schemes on the mode shares are limited: not only is the capacity of the system rather small, but it was also observed that users of bike sharing systems use it mainly to perform trips that they would have performed with “sustainable” modes (Fishman, Washington & Haworth, 2013). What was observed however is a strong impact of those systems on the image of biking in the city, pushing more individuals to use their private bikes regularly: after introduction of the two big French bike sharing systems, a significant increase of bike use was observed: 44% in Lyon, and 70% in Paris (Shaheen et al., 2010). Such an effect is usually seen as positive, for environmental, public health and “livability” reasons. Another strong point in favor of such systems is that they might help to solve the “first/last mile problem”, that is, the loss of attractiveness of public transport due to the difficulty to get to the initial station or go from the end stop to the destination of the trip. Indeed, a study in Dublin showed that 55% of the bike sharing trips were actually performed as part of a more complex multimodal chain (Fishman et al., 2013).

Those systems come however with operational challenges. One particularly difficult problem is bike redistribution. Indeed, asymmetry in the travel demand and aversion for uphill riding often leads to shortage of bikes in stations with lots of departures and shortage of available slots in stations with lots of arrivals. To address those problems, several approaches can be used. The most common is the usage of relocating trucks, that redistribute bikes during the day. This kind of strategy however comes with a high monetary cost, and might threaten the environmental credibility of bike sharing (Fishman et al., 2013). Another approach is the usage of incentives to push the system to regulate itself better. Those incentives can typically take the form of station-dependent fares, coupled with an information system. In the case of Paris, for instance, while trips are normally free up to 30 minutes, some stations located uphill or in the outskirts of the city, labeled “Velib’+”, give right to 45 minutes of free travel time (Nair et al., 2013).

The design of redistribution strategies to allow the provision of a reasonable Level of Service while keeping the costs as low as possible has of course received attention from the research community. Raviv, Tzur & Forma (2013) for instance consider the static redistribution problem, when bikes are only redistributed once during the night. Other Operations Research approaches have included some model of demand, based on historical data, to design dynamic redistribution strategies, where trucks relocate bicycles throughout the whole day (Contardo, Morency & Rousseau, 2012; Fricker & Gast, 2014; Schuijbroek, Hampshire & van Hoeve, 2013) to satisfy most of the demand. Similar work is also done in the car-sharing field (Barth & Todd, 1999; Kek, Cheu & Chor, 2006). Fricker & Gast (2014) also consider
the possibility of “incentives”, which take the form of directives from the operator, such as proposition of an alternative drop-off station, that the user is assumed to follow. Pfromer, Warrington, Schildbach & Morari (2014) considers dynamic price incentives, and estimate their effect by considering customers with a linear value of time. However, none of those approaches contains a model of the reaction of the demand to a change in the quality of the redistribution process. The approach used here thus takes the problem from a different angle: rather than designing optimal relocation strategies given a demand, it proposes a method to estimate the reactions of demand to a change in the relocation strategy.

Among behavioral models, multi-agent activity-based models, such as MATSim, are particularly well suited to represent reactions of demand to a change in supply, by seeing individuals as competing for limited joint resources. Though the usual resource considered is road capacity, this framework naturally extends to the bike sharing case: individuals consider bike sharing not only based on its intrinsic value, but also on the availability of bikes, which depends on the behavior of others. The simulation system of WP5 allows to represent such effects (Dubernet & Axhausen, 2014).

Though the simulation framework from WP5 allows to represent any kind of redistribution strategy, no complex strategy was yet implemented. There is however another strength of multi-agent activity-based models yet to be leveraged, that is, their fully disaggregate representation of supply and demand. It allows to identify potentially problematic stations, taking into account asymmetric mobility flows, as well as factors specific to cycling, such as aversion for cycling uphill, allowing to guide the design of the bike sharing system before its actual implementation.

The scenario used here thus uses an optimal bike redistribution strategy, where if there is a bike parked in the system and a user needs to pick up a bike, the user will find a bike at the desired station.

For each link in the (navigation) network, the average slope is known. The importance of slope is considered in the simulation via two distinct additions:

- Bike and walk travel times are calculated, using slope information, along the least-cost-path in the network, using person-dependent speeds. Details of this approach are detailed in Dobler & Lämmel (2012)

- by considering slope in the disutility of travel. Several route choice models in the literature consider slope one way or another. Menghini, Carrasco, Schüssler & Axhausen (2010) for instance estimate a route choice model using GPS traces for Zurich, and include the maximum gradient in the utility of a route; Stinson & Bhat (2005) include three levels of hilliness (mountainous, hilly and flat) in a stated preference survey in the USA; Hood, Erhardt, Frazier & Schenk (2014) use the cumulated gain in a model based on GPS traces for the Monterey Bay Area. Though the model of Menghini et al. (2010) was estimated for Zurich, its formulation using a global route characteristic (maximum gradient along the path), rather than a link-level property, makes it difficult to use in standard shortest path algorithms. On the contrary, the formulation from Hood et al. (2014) is relatively straightforward to include in the scoring function, as it allows to convert additional uphill gain into additional distance traveled, and is thus the one that was chosen for this scenario. It results in a parameter saying that riding one additional meter uphill is equivalent to riding 13 additional meters on flat land.
3.3.2 Outputs of the Scenario

Zurich being a hilly city, topography is likely to have a strong influence on bike sharing demand. Identifying the effects of topographical factors, isolating them from other causes of spatial variability of demand, can thus help design guidelines for the establishment of bike sharing services. Two behavioral parameter sets are thus considered, depending on whether the influence of slope on the disutility of traveling is considered. 10 simulation runs, each one with a different random seed, are ran for the “realistic” case, with disutility of slope. An additional run, without this disutility, is performed as a reference — computing resources limitations made it impossible to realise 10 runs for this case.

Figure 18 shows the average, min and max bike sharing usage across the simulated day, for the runs with disutility of riding uphill. The results of the 10 runs are pretty similar. The number of en-route bike remains significantly lower than the capacity of the system (4500 bikes). Whether this is a capacity effect or the demand correspond to the maximal demand, could not be investigated due to computation capacity limitations. The demand is relatively constant throughout the day, without the usual morning and evening peaks, pointing more towards a capacity effect.

![Figure 18](image.png)

Figure 18. Number of en route bike sharing bikes throughout the day

An important question, when looking at the influence of topography on the performance of the system, is the following: are stations uphill really generating more trips than downhill stations, and how much?
Figure 19 answers this question, by plotting the average of the difference between inflow and outflow for stations across the 10 runs, as a function of altitude. Overall, stations downhill do indeed attract more trips than they generate, and stations uphill generate more than they attract.

![Figure 19. Difference between inflow and outflow with altitude](image)

Figures 20 and 21 show the in and outflow per station, respectively. The shades of blue represent topography (the darker the color, the higher), and the area of dots is proportional to the number of trips originating or terminating in a given station, respectively. Figure 22 shows the difference between inflow and outflow. Red stations generate more trips than they attract, green stations attract more trips than they generate, and the area of dots is proportional to the difference between in and out flow. Figures 23 to 25 are similar, but for the run with aversion for uphill riding.

The difference between in and outflow for uphill stations can here be observed again. Imbalance follows a pretty clear pattern, where the city center attracts trips, and the periphery generates. This does not happen for the runs without uphill aversion, pointing towards an effect of topography. Another pattern can be observed in the runs with uphill aversion, but not in the run without: the demand in the northern part of the city is pretty low. Indeed, the city is cut in two by hills, with the city center at the south, and more residential areas in the north, resulting in a large part of the trips generated or attracted by the northern part of the city passing from one side of the hills to the other — making bike sharing much less attractive in this area, as most of those trips result in a uphill ride.
Figure 20. Outflow per station
Figure 21. Inflow per station
Figure 22. Out excess per station
Figure 23. Outflow per station (no uphill aversion)
Figure 24. Inflow per station (no uphill aversion)
Those results show which influence the hilliness of Zurich might have on the operation of a bike sharing system. Indeed, in Zurich, the city center is located at border of the lake, and the periphery climbs the surrounding hills. This has a strong impact on the bike sharing demand, as the two causes of demand imbalance (asymmetric mobility flows and aversion for uphill rides) reinforce each other. This might make relocation of bicycle particularly challenging. Another identified effect is a low demand in the northern part of the city, separated from the city center by hills. This area of the city might thus not need as high a density of stations as the city center, and might even be left out in the first step on the 3-steps process planned to build the full system, decreasing operational difficulties without reducing a lot the part of the demand that cannot be satisfied.
4. Outlook

The Zurich base scenario is build using standard data sources, mainly Census and Travel Diary Survey for the demand, and infrastructure data for the supply. In addition, it includes a synthetic leisure contact network, generated with a model estimated on a snowball sample, a kind a dataset collected in Barcelona for the EUNOIA Project. The usage of such a social network in an operational transport model is here done for the first time, and was shown to have significant influence on the quality of predicted traveled distance distributions, in particular for the car passenger mode, absent of most current model.

The joint decisions simulation framework from WP5 was tested in two cases: households and leisure contacts. Interesting results could be obtained:

- escorting trips to education emerge naturally from the constraints specific to households, without any specific parameter having to be involved.
- the willingness to meet social contacts during leisure could produce realistic traveled distances for car passengers, due to the willingness to arrive jointly at destination.

The most directly applicable development of WP5, the bike sharing module, was used for the Zurich area, taking into account the aversion for uphill riding. It could be shown that the topographical specificities of Zurich, with peripheral urban areas located on hills surrounding the city center, has a strong impact on bike sharing demand imbalance, with topography reinforcing the natural asymmetry of mobility flows.

Those scenarios however still lack several important features:

- In the case of households, activity chains are still generated independently for each agent — mainly due to the lack of appropriate data for the Swiss population. Such household level generation is however possible in the case of the London input data; time unfortunately did not allow to test the approach in this area.
- For computation time reasons, the generation of the social network for a synthetic population as large as the Swiss 10% sample had to be performed evaluating only a very low fraction of the possible ties (0.05%), compared to the rate that was shown to have good results on a small sample (7.5%) by Arentze et al. (2013). A more efficient implementation of the social network generation software is in preparation, but could not be ready early enough.
- The bike sharing scenario includes slope-dependent travel times and disutilities and capacity constraints, but not yet a realistic redistribution strategy, nor the possibility to represent induced demand. The challenging task of representing induced demand in MATSim has been experimented by Feil (2010), but was never taken to the operational level, and never actually validated.
- The aversion for uphill riding was transferred from a route choice model from the literature, but the effect of this parameter was not validated: only the global mode share for the bike mode was recalibrated. The intensity of the various observed effects might thus lack accuracy.
- The analysis would have benefited from additional runs, for instance varying the bike sharing supply according to the recommendations in the conclusion of the previous section, to test them. Limited computing resources unfortunately made this impossible for this report.
- Though the simulations consider the dynamic nature of mobility, the analyses remained limited to aggregations at the day level. It is planned to go further in the analysis, by looking at the dynamics.
Annex I. A short presentation of MATSim

Philosophy

MATSim is a software framework, aiming at representing interactions between transport supply and mobility behavior with a high level of details. It is based on the convergence of three established streams of research, namely:
- activity-based mobility behavior analysis, that explicitly looks at travel as a consequence of the desire of individuals to perform activities, to represent demand,
- microscopic traffic simulation methods, allowing to represent complex but important effects in traffic flows, such as queue spillback, as well as other effects of individuals' interactions, such as crowding of public transport vehicles,
- game theory and co-evolutionary computation methods, that provide a theoretical ground for modeling the interactions of supply and demand, as well as practical and efficient ways to find approximate solutions for the resulting theoretical model.

The software was implemented in the Java programming language, that is multi-platform and relatively easy to learn for modelers with little programming background. It is free and open-source, meaning any interested researcher or modeler can analyse, use or extend it for his or her use (see www.matsim.org). The software is designed to allow usage in large-scale simulations, in particular by adopting a minimal yet powerful traffic flow model.

Modeling Principles: Game Theory and Co-evolutionary Computation

Given the philosophy above, the aim of MATSim can be summarized as modeling the influence of individuals interactions in the transport network on their activity and trip planning decisions. Modeling the interaction of individuals with possibly conflicting objectives has been the subject of game theory for decades, making this theoretical framework particularly well suited for the problem at hand.

A game theoretic view of transportation systems has indeed been popular for more than half a century, since the seminal work of Wardrop (1952) The essential idea behind it is to see the transportation system as a set of shared resources (road space, public transport vehicle seats...), for which individuals compete, individuals in the population trying to maximize their own satisfaction, given the resources left available by others. Game theory studies solution concepts for such strategic interactions. A game theoretic solution concept is a definition of which states are equilibria, that is, stable under assumption of rationality — a state being considered stable if no agent/player has an incentive to change its behavior. The static, trip-based approach of Wardrop (1952) has been refined and extended with time. In particular, the equilibrium idea can be pretty naturally transfered to the activity based framework: individuals do not just try to optimize their trips, but their whole day. The implementation of this idea in a practical, usable software framework, was the aim of the MATSim software framework from the beginning (Axhausen, 2006; Nagel & Flötteröd, 2009). MATSim simulates agents having daily activity plans, consisting of activities located in space, and the trips between those activities.

Co-evolutionary algorithms are particularly well suited for searching for states approximately satisfying game theoretic solution concepts, when the search space is too big for exhaustive search (Ficici, 2004; Popovici, Wiegand & De Jong, 2012). In this kinds of algorithm, an evolutionary process is performed on
sub-solutions, that are evaluated by interaction. In the case of MATSim, those sub-solutions are individual daily plans, that are evaluated by interaction in a mobility simulation, that represents congestion. This process can be seen as an emulation of a learning process (Nagel & Marchal, 2006). The steps of this process, represented on Figure 26 are the following:

1. **Initial demand** All agents have an initial daily plan, which will serve as a starting point for the iterative improvement process. Some characteristics of the plans are left untouched during the simulation, and should therefore come from data or external model. This is typically the case of long term decisions, such as home and work locations, or decisions involving a larger time frame than a single day (e.g. do the weekly shopping or not).

2. **Mobility simulation** Plans of all agents are executed concurrently, to allow estimating the influence of the plans of the agents on each other.

3. **Scoring** The information from the simulation is used to estimate the score of each individual plan. This information typically takes the form of travel times and time spent performing activities; experiments also included information such as facility crowding (Horni, Scott, Balmer & Axhausen, 2009). The functional form is the one used by Charypar & Nagel (2005). It uses a linear disutility of travel time, and a logarithmic utility of time passed performing activities. Different parameters can be defined for each mode/activity type.

4. **Replanning** This step actually groups two of the important components of co-evolutionary algorithms: selection of the interactions for evaluation, and application of the evolutionary operators (selection and mutation).

   To do so, part of the agents select a past plan based on the experienced score, following a Logit selection probability. This will have two consequences: the state of the transport system, used for evaluation, will only evolve slowly from iteration to iteration, giving the time to the agents to adapt, and those plans will be re-evaluated, given the new plans of the other agents. The other agents copy and mutate one of their past plans. If the number of plans in an agent’s memory exceeds a predefined threshold (usually 4 or 5), the worst plan is deleted, pushing the evolution towards plans with higher scores. Steps 2 to 4 are then iterated until the system reaches a stable state.

   What kind of mutation is performed determines which alternative plans will be tried out by the agent. Typical replanning strategies include least cost rerouting using travel time estimates from the previous iteration, departure time mutation, and mode mutation at the subtour level, considering mode chaining constraints. A tour is a sequence of consecutive trips starting and ending at the same location, named anchor point. A subtour is a tour, possibly without other tours it contains. Vehicular modes can only be performed for whole subtours, which must be anchored at home or in subtours of the same mode. Experiments included secondary activity location choice (Horni et al., 2009) and activity sequence (Feil, 2010).

   Those steps are iterated until a stationary state is reached, and the state of the system in this stationary state is taken as a result.
The Physical World in MATSim

With the scoring function, an important element is the *mobility simulation*, that is used to evaluate the influence of agents on each other’s satisfaction.

The standard mobility simulation uses a step-base queue model to simulate road infrastructure, that agents might use with a car, or by boarding a public transport vehicle. Other modes are assumed not to suffer from congestion, and are simply “teleported” with a pre-computed travel time. This simulation framework was found to allow fast large-scale simulation, while still representing the essential dynamics of congestion.

To Sum Up

MATSim is a freely available open-source simulation software, that has already been used in Switzerland (Meister et al., 2010), Berlin, or Singapore (Erath et al., 2012), to cite only a few examples.

It searches for an *equilibrium* for daily plans in a synthetic population, where individuals cannot improve the satisfaction they get from their day, given the state of the transport system, that depends on how individuals use it.
References


