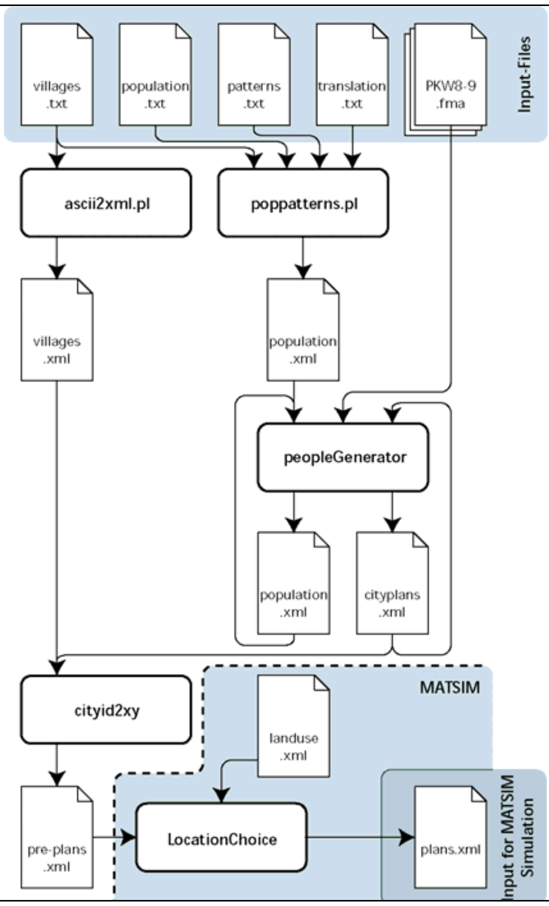


Generating Day Plans Based on Origin-Destination Matrices

A comparison between VISUM and MATSIM based on Kanton Zurich data

Michael Balmer, IVT, ETH Zurich
Marcel Rieser, D-INFK, ETH Zurich
Arnd Vogel, VSP, TU-Berlin
Kay W. Axhausen, IVT, ETH Zurich
Kai Nagel, VSP, TU-Berlin

Conference paper STRC 2005

STRC | 5th Swiss Transport Research Conference
Monte Verità / Ascona, March 9-11, 2005

Generating Day Plans Based on Origin-Destination Matrices

Michael Balmer
IVT
Zurich

Phone: +41-1-633 27 80
Fax: +41-1-633 10 57
email:
balmer@ivt.baug.ethz.ch

Marcel Rieser
D-INFK
ETH Zurich

email:
mrieser@student.ethz.ch

Arnd Vogel
VSP
Berlin

Phone: +49-30-314-29522
Fax: +49-30-314-26269
email:
vogel@vsp.tu-berlin.de

Kay W. Axhausen
IVT
ETH Zurich

Phone: +41-1-633 39 43
Fax: +41-1-633 10 57
email:
axhausen@ivt.baug.ethz.ch

Kai Nagel
VSP
Berlin

Phone : +49-30-314-23308
Fax: +49-30-314-26269
email:
nagel@vsp.tu-berlin.de

March 2005

Abstract

Microsimulation tools are becoming increasingly important in traffic demand modeling. The major advantage over traditional assignment models lies in the fact that each traveler is simulated individually. This means, for example, that decision making processes can be included for each individual. The traffic demand is a result of different decisions made by each individual. Those decisions lead to plans which the individuals try to optimize. Because microsimulation includes decision making processes, some information about the individual's plans has to be given.

On the other hand, traditional assignment tools like VISUM or EMME/2 use origin-destination matrices (OD-matrices) as inputs. Those matrices do not have any information about the chains of activities which define plans for individuals. This raises the question of reconstructing plans from OD-matrices, which is also discussed in this paper. The plans generated by this approach are simulated with MATSIM—a microsimulation model—and the outcome are compared to the results of a VISUM—a traditional assignment model—of the same OD-matrices. The comparison is carried out by a case study of the greater Zurich area of Switzerland.

Keywords

microsimulation – demand modelling – origin-destination matrix – VISUM – MATSIM – plans generation –Swiss Transport Research Conference – STRC 2005 – Monte Verità

1. Introduction

Microsimulation is becoming more and more important in traffic simulation, traffic analysis and traffic forecasting (see Vovsha, 2002). Some advantages over conventional assignment models are:

- Computational savings in the calculation and storage of large multidimensional probability arrays.
- Larger range of output options, from overall statistics to precise information about each specific synthetic traveller in the simulation.
- Explicit modelling of the decision making processes of the individuals.

The last point is important since it is not a vehicle which produces traffic; it is the person who drives it. Persons do not just produce traffic; instead each of them tries to manage his day (week, life) in a profitable way. They go to work to gain money, they go hiking for their health and pleasure, they visit their relatives for pleasure or because they feel obliged to do so, they shop to cook a nice dinner at home, and so on. Since not all of this can be done at the same location they have to travel, which produces traffic. To plan a day efficiently, many decisions are made by each person:

- Which route should I take to get to work? - *Route choice decision*
- Which mode should I use to go to the lake? - *Mode choice decision*
- Should I drink another beer before going home? - *Activity duration choice decision*
- Should I go shopping near my home or at the mall? - *Location choice decision*
- When should I do sports today? - *Activity starting time choice decision*
- Should I go to visit my friend? - *Activity choice decision*
- Whom should I take along? - *Group composition decision*
- Should I go swimming before or after work? - *Activity chain choice decision*

There are many more decisions to make; some of them are made hours (days, months) in advance while others are made as spontaneous reactions to specific circumstances. Many decisions induce other decisions. For example, if I am late for work, I am supposed to work longer, so there's no time left to go shopping today, so I need some time tomorrow to do the shopping. This example shows the importance of describing plans for each individual in a simulation model, because it is the plan and the decisions made by the person who adhere to this plan that produces the traffic.

To simulate a typical day in an urban area, microsimulation tools need precise information about the plans of each individual and also some knowledge about people's decision making process. But to produce daily plans (activity chains) for persons is not a trivial task. On the other hand conventional assignment models like VISUM (http://www.ptv.de) or EMME/2 (http://www.inro.ca) typically use demand matrices which hold the information about the trips from an origin to a destination (usually zones or regions), but these do not provide links back to individuals. This leads to the main question of this paper:

Is it possible to impute (reconstruct) people's activity chains from a set of origin-destination matrices (OD-matrices)?

This paper demonstrates first approaches to answer this question. It is organized as follows: first, the case study to which the approaches are applied is discussed including the description of the OD-matrices and how they have been generated. This is followed by a detailed description of the activity chain generation process based on the given OD-matrices. An alternative activity chain generation process without using OD-matrices is described afterwards which then will be used—together with other, simpler approaches—to compare the quality of the results. VISUM (http://www.ptv.de) and MATSIM (http://www.matsim.org) are used to compute traffic volumes which also will be compared with real count data. Finally, concluding remarks and recommendations for future research are outlined.

2. Case Study of the Zurich Area

The Zurich area has about 1.3 million inhabitants and is divided into 170 municipalities and 12 additional districts inside the city of Zurich. The following subsections describe briefly how those OD-matrices were generated using VISEM (http://www.ptv.de).

2.1 Data Resources

The microcensus (ARE, 2001) is a periodic survey of the travel behaviour of the Swiss population. It has been run every five years since 1974. The microcensus is carried out by the Federal Office for Spatial Development ARE (see http://www.are.admin.ch) in cooperation with the Swiss Federal Statistical Office BFS (see http://www.statistik.admin.ch/index.htm) and contains detailed information about the mobility behaviour of 30'000 persons from all over Switzerland. Information about population distribution in the municipalities comes from the federal population census (http://www.volkszaehlung.ch). BFS also provides land-use data. The regional transport models (road, rail and other public transport) (see Vrtic *et al.*, 2002) describe the effective travel times between each municipality. In summary the following is available:

Municipalities

- Location
- Structural properties
- Travel distance matrices
- Travel time matrices for private and public transport
- Accessibility classes
- Local access and egress times to the modelled networks

Population

- Home locations
- Population groups (children, worker, non-worker, senior)
- Mobility (car, season ticket ownership, bike, walking)

- Activity chain distributions by subgroup

About 1670 different activity chains can be found in Micocensus2000. Most of them appear very rarely, therefore only the 100 most frequently occurring activity chains are used later, which still cover more than 90% of all days.

2.2 VISEM output

VISEM (http://www.ptv.de) is conceptually an activity-based traffic demand generation model. It differentiates population sub-classes (e.g. juniors, working adults, non-working adults, and seniors) by the respective distribution of activity patterns (e.g. from the microcensus2000). VISEM then generates a synthetic population from the census data and assigns an activity chain to each subgroup of the population. Locations for activities are assigned using gravity models, while a logit model is used for mode choice. The results are aggregated into hourly zonal OD-matrices (including the purpose of each trip) on the basis of time-specific transition probabilities between activity types.

It is to mention that the resulting OD-matrices of the case study of the Zurich area contain more trips than depicted by other research studies of that area (i.e. Pendlermatrix, see Vrtic and Axhausen, 2003). Three conspicuous traffic peaks (morning, noon and evening) occur from the OD-matrices. For more detailed information see Rieser (2004).

3. Activity Chain Generation Process Based on OD-Matrices

The main input are OD-matrices (hourly and by mode), as discussed above. The matrices do not show which person is generating a specific trip and what the following trips will be. Even more, we do not know how long an activity will be performed before the next trip starts. The question is if it is possible to reconstruct plausible complete daily activity chains from that material and some additional information and assumptions.

Given activity chain distribution information e.g. from the microsensus2000, (see ARE, 2001), population distribution and land uses (say, from the BFS) it should be possible to impute activity chains from the given matrices. Unfortunately the additional information available does not contain information about the duration of the individual activities. But to combine different trips to form a complete chain, one needs at least some information about the average time an individual spends on an activity. The activity chain generation process requires such information. So, assumptions have to be made as long no other data is available.

3.1 Target (MATSIM Input Plans)

MATSIM accepts different kinds of plans: from simple one-trip-plans (corresponds to OD-trips) to full daily activity chains. But to use the advantage of enhanced re-planning rules for individuals daily chains are favoured. The first activity should include an end-time which defines when the first trip of the plan starts. In the case, of the daily chain the last activity does not need any timing information because it defines the end of the day. All other activities must have a defined duration. The location (given as x and y coordinates) and the activity type has to be included for each activity. Table 1 shows an example of a day for a person. Please notice that MATSIM handles only motorized individual transport mode at the moment. Therefore the mode for each leg is defined as “car”.

Table 1 Example of a MATSIM day plan (XML format)

```
<person id="14">
  <plan selected="no">
    <act type="h" x100="683021" y100="261826" end_time="06:40:00" />
    <leg mode="car"/>
    <act type="w" x100="687408" y100="256250" dur="04:00" />
    <leg mode="car"/>
    <act type="h" x100="683021" y100="261826" dur="01:00" />
    <leg mode="car"/>
    <act type="s" x100="683750" y100="246450" dur="01:30" />
    <leg mode="car"/>
    <act type="w" x100="685350" y100="254950" dur="03:00" />
    <leg mode="car"/>
    <act type="h" x100="683021" y100="261826" />
  </plan>
</person>
```

3.2 The Process

The process creates the MATSIM input plans for each user-defined mode separately. Figure 1 shows the process to create the activity chains for a specified mode. Each step uses specific input files and also produces well defined output files. Each step can be started independently, which makes it easier to extend the process with additional steps, methods or information. This section describes the different files first, followed by descriptions of each step:

- *villages.txt*: Holds the 182 municipalities / districts, their names and the coordinates of their centroids.

- *population.txt*: Holds the number of persons for each population group and each municipality.

- *patterns.txt*: Holds the share of each given pattern by population group.

- *translation.txt*: correspondence list between German and English activity codes.

- *villages.xml*: The same information as in villages.txt but in XML format.

- *fma files*: The hourly OD-matrices for a specified mode generated from VISEM.

- *population.xml*: Contains the persons by municipality who do not have an activity chain assigned yet.

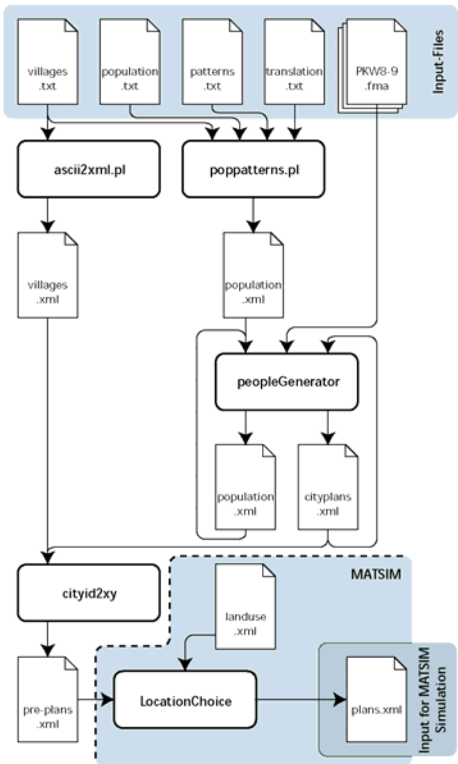
- *cityplans.xml*: Chains for persons similar to Table 1, but the activities do not have a location assigned yet. However, each person retains the information about their home location and the specified mode.

- *pre-plans.xml*: Chains for all persons similar to Table 1 but only the primary activities have a location assigned. (The definition of the primary activity will be explained in the following section.)

- *landuse.xml*: This file contains land use data. It essentially describes the attractiveness of each area. Since for Switzerland land use information is available for hectares (100 meters times 100 meters), the land use information also refers to that resolution. Attractiveness is described as the number of opportunities by activity type.

Further more, the modal split is given to reduce the population by the specified mode.

Figure 1 Activity chain generation process based on OD-matrices



3.3 Steps of the Process

This section describes each step of the process shown in Figure 1 in details:

ascii2xml.pl

This is a pre-process to convert the villages.txt file to XML format.

poppatterns.pl

This is also a pre-process which calculates the number of persons by municipality with a particular activity chain and by the specified mode. The given distribution of the activity chains is used to assign a chain to the persons. Notice that minor rounding errors of the number of people can occur because of distributing the population to the municipality, reducing the population according the given modal split and assigning the activity chains.

peopleGenerator

This is the main step of the process. It sequentially goes through each hour of the OD-matrices for the specified mode and generates partially completed chains.

The peopleGenerator process reads (sequentially, hour by hour) one line after the other from the OD-matrix. Every cell of the OD matrix represents a number of persons which travel from one location to another. The algorithm selects those persons from the population.xml file of the origin of the OD-trip such that their primary activity matches the purpose of the OD-trip. The match also defines the start-time (given by the current hour of the OD-matrix). Since each matrix represents a full hour of the day, the chosen start-time is calculated as a uniform distribution of that hour. Also the number of people in the population list is decreased by the number of assigned persons. Recall that every person already has an activity pattern assigned from the pre-processing stage.

This module is used only to assign locations to the primary activities. In our simplified approach, there is only one primary activity per chain. If there is more than one out-of-home activity (peopleGenerator recognizes only 'work', 'education', 'leisure' and 'shop' out-of-home activities at the moment), then one of them is declared the primary activity. The idea here is to select the location of the primary activity based on the VISEM output, and then use some other module to select secondary activity locations. This leaves the question of defining the primary activity when there are multiple out-of-home activities in a row. The following simplified approach was adopted:

The first occurrence of 'work' is set as primary. If 'work' is not part of the chain, the first occurrence of 'education' is set as primary. If 'education' is not part of the chain, the first *occurrence of 'leisure' is set as primary. If 'leisure' is not part of the chain, the first occurrence of 'shopping' is set as primary.*

As the count of available people is decreased while reading one OD-matrix entry after the other, it can happen that there are no more people available in a location to fulfil the travel predicted by VISEM. In this case, no new plans are generated.

Every chosen person together with their plan is written to the population.xml file. When writing, each activity gets a 'standard' duration assigned. Recall that no data about durations of activities are available assumption about the average duration of an activity type spent per plan has to be made (user-defined input of this process). If an activity type occurs more than once in a plan, the duration for each single activity is divided by the number of occurrences of the activity type. This way, the sum of durations of all the activities of one type in the plans matches the assumed default durations.

Even though those assumptions are sensible, sometimes the durations can be quite wrong. For example, the duration of the shopping activity in the activity chain 'home-shop-home' could be just some minutes (i.e. buying a bread at the bakery next door) or it could also be about 10 hours (i.e. a shopping day at the mall). The same question applies for the leisure activities. On the other hand, the average duration of 'work' and 'education' usually has less variance. Nevertheless, for lack of more detailed assumptions the generation process will be employing these assumptions for the time being.

cityid2xy

Until now, all plans have the attribute “cityid”, which MATSIM does not recognize. Instead, it expects the x- and y-coordinates for a location. Cityid2xy achieves this by choosing a random point within the given municipality (here: municipalities and districts), thus disaggregating the locations.

To choose random points the algorithm simplifies the shape of cities into circles. For every municipality the algorithm searches for the nearest neighbour and takes the half of that distance as radius for a circle around the respective centre. Using this procedure, the areas of cities will not overlap each other.

LocationChoice

MATSIM contains a module (see Marchal and Nagel, 2004), which chooses locations for secondary activities in such a way that each agent improves the daily chain. The home-location and the location of the primary activity remain unchanged.

The application uses variant of a genetic algorithms to determine the locations for each agent. Each agent (based on the previously created chain) has limited knowledge about places in the area of interest and searches within these places for a path, such that his chain improves its utility. Additionally, each agent “knows” some other agents and exchanges information about places after every iteration with them. This way, a social network is built that helps the agents improve the utility of their chains.

The original application as described in Marchal and Nagel (2004) assumed that all plans contain the activity 'work', which was interpreted as primary activity. Additionally, the application expected every chain to have one or two secondary activities. As the activity chains (see Rieser, 2004) do not fulfil these restrictive requirements, the application was modified. In a first step, it was ensured that short plans with no secondary activities (i.e. home-work-home) were processed correctly. Such chains will now be ignored for the iterative calculations, and left unchanged.

In a second step, the code was modified to respect the primary tag for activities instead of checking whether an activity is of the type 'work'. The primary tag is set by the above described process.

After these modifications the LocationChoice module is able to generate the missing locations for the secondary activities for the plans given by the cityid2xy process. The resulting plans file has the required information MATSIM needs (see Table 1).

We need to point out one important fact about this location choice process: Even it does use the information of the land use data; it does not respect the constraints given by that. In other words it can happen that more people work at a given raster element than working places are available.

It must be noted that it this whole plan generation process often does not reach a perfect match with the given OD-matrices. As an example, the output matrix for a whole day generated by VISEM (http://www.ptv.de) as explained above prevents a perfect match: for many municipalities, the number of trip starts differs from the number of trip ends making it impossible for persons to be at home at the end of a day. On the other hand, the activity chain generation process generates only roundtrips, so the number of trip starts and trip ends will be the same in each village.

4. Activity Chain Generation Process without using OD-Matrices

The activity chain generation process described in Section 3 is using origin-destination matrices produced by VISEM. On the other hand, the available data resources allow us to create those activity chains directly. This section describes one possible way to do that. Notice that the generation process uses three different level of detail:

- the municipality / district level (182 in the Zurich area) denoted with letter "M",
- the land use raster level (100 x 100 square meters) denoted with letter "R" and
- the street network level (locations are mapped to a link of the network) denoted with "L_ID".

Since different data resources refer to different level of details, it is necessary to "switch" between those levels.

The activity chain generation process is split up into eight sequential steps (see Figure 2 for a graphical overview). Each process step (except the first one) uses one specific data resource to extend a plan similar to the example shown in Table 1. The following describes the steps in details:

1. *Conversion*: This process step just converts the input population flat files (i.e. population.txt file as shown in Section 3.2) into the XML person description file. None of the person tags holds a plan yet. But additional attributes like age, sex, etc are included. It also holds the information in which land use raster element this person lives.

2. *Mapping*: Each raster element of the land use data belongs exactly to one municipality. By using this mapping, each agent can be assigned to the municipality of his home location.

3. *Distribution*: Given the distribution of the activity chains described in Section 2.1, we assign one of the chains to each person according to the given distribution. We also respect the fact that children do not go to work, therefore persons of young age are not allowed to hold an activity chain including a "work" activity.

4. *Distribution*: The Pendlermatrix 2000 (Vrtic and Axhausen, 2003) holds the information about work and education commuters on municipality level of detail. With the assumption about primary activities given in Section 3.3 we are able to add

the locations of the primary activities “work” and “education”. Unfortunately we do not have similar data for the primary activities “shop” and “leisure”. As long no better data is available we just assume that those activities are done in the same municipality / district where the person lives.

5. *Disaggregation*: The process step 6 uses the location choice (see Section 3.3) which is working on the raster level of detail. Therefore, we need to disaggregate the locations of the secondary activities to that level. It is done by uniformly picking one of the raster elements of the given municipality / district.

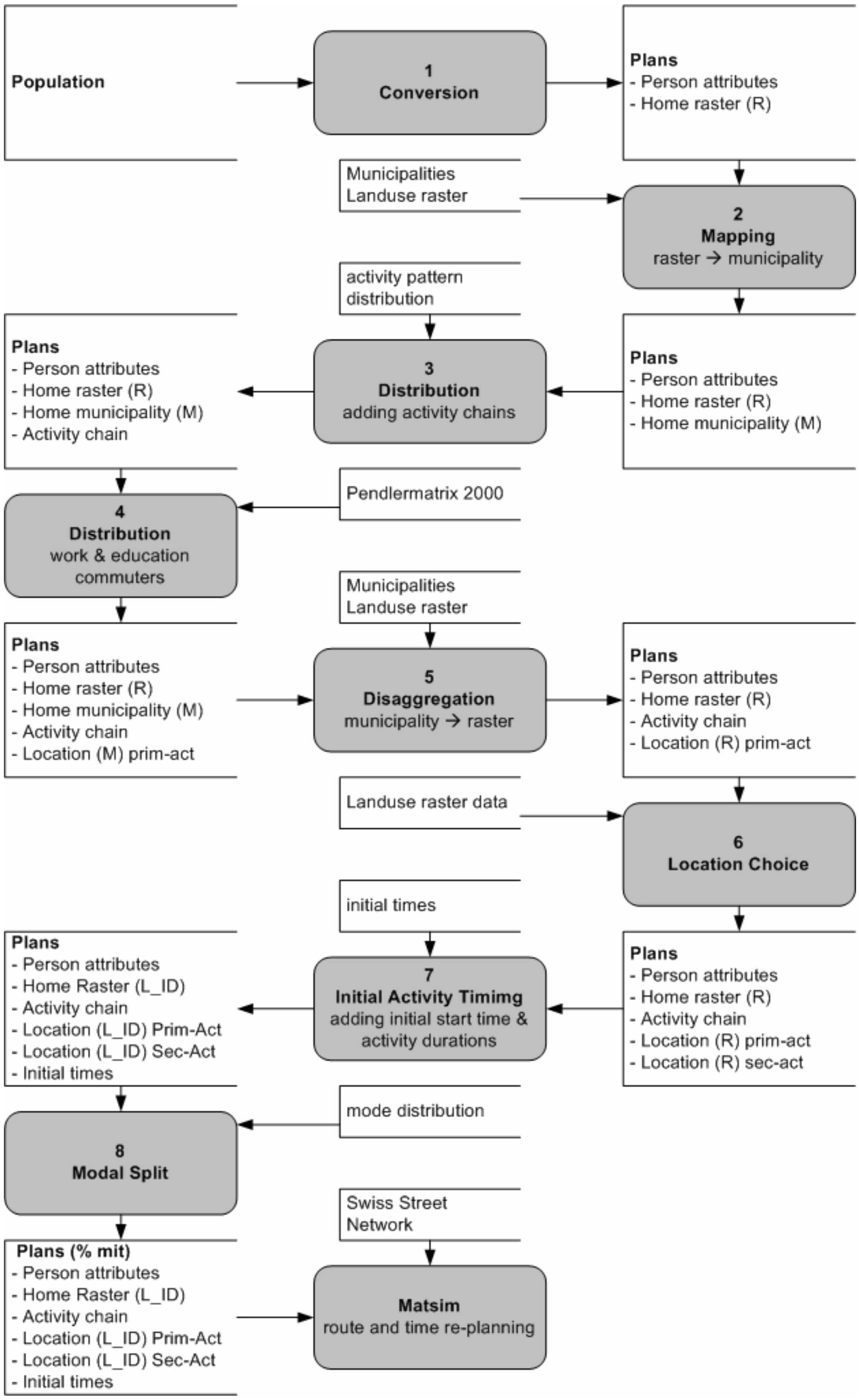
6. *Location choice*: The XML person description file holds now the same amount of information as shown in Table 1 except that the end time and the durations are missing. Since the location choice process (Marchal and Nagel, 2004) does not need timing information we can use that process to add the missing location of the secondary activities like described in Section 3.3.

7. *Initial activity timing*: As mentioned above, the end time and the durations of the activities are missing. Like described in Section 3 we define standard durations. In difference to Section 3 we also define the end time of the first activity. Since MATSIM already showed that it can generate reasonable end times by using a time allocation re-planning module (Balmer *et al.*, 2004) those defined end times are just initial values which will be changed during the MATSIM iteration process.

8. *Modal split*: MATSIM currently can only handle motorized individual transport mode. Until this process step we were handling the whole population. We now need to reduce the number of persons to a certain percentage given by modal split data. It is done by randomly picking persons according to the given percentage of the motorized individual transport mode.

In the resulting XML file all required data is available to use as the initial input for MATSIM.

Figure 2 Gane-Sarson process plan for generating daily activity chains

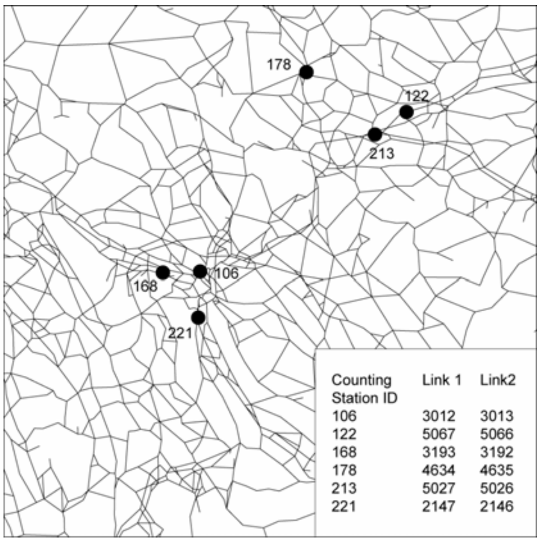


5. Comparisons

In this section, the results of an assignment procedure done with VISUM (http://www.ptv.ch) using the OD-matrices are compared to a simulation run with MATSIM (Raney and Nagel, 2004) using generated plans by the two processes described in Section 3 and 4. In addition, a direct conversion of the OD-matrices into one-trip-plans was performed and simulated with MATSIM. The results of this simulation run will also be compared to the VISUM results. The comparison between VISUM and MATSIM will be performed only for the motorized private transport since MATSIM is only able to handle this mode at the moment.

Both simulations are using the same underlying street network from ARE (2001). Some changes made to this network are described in (Balmer *et al.*, 2004). The comparison of the outcome of the simulations will be done for twelve links (see Figure 3) of the given network where also count data is available (ASTRA, 2001).

Figure 3 Six count stations (12 links) in the Zurich area



5.1 Data used

The case study used is the same as in Rieser (2004) containing 182 municipalities / districts with a total population of 1'247'566 persons. Using a model share of 45.44% for motorized individual transport, about 566000 persons are simulated.

Instead of using all 100 activity chains from Rieser (2004), only a subset was chosen. In a first step, the number of activities was reduced by merging several rarely occuring activities into the "work"-activity. These were the activities "Begleitung" (escort), "Service" (service), "Geschaeftsreise" (business trip) and "Dienstreise" (travelling on company business). The remaining activity types are now "h" (home), "e" (education), "l" (leisure) and "s" (shopping). In a second step, all activity chains matching at least one of the following requirements were chosen:

- Activity chains consist of four or less activities.
- The number of occurrences of the activity chain is at least 1% or more of all occurrences.

This way, 21 activity patterns remained, representing nearly 93% of all occurrences of the 100 chains. It is to mention that the 100 activity chains from Rieser (2004) do not include "home" activities in between (which are also called journeys), since VISEM (http://www.ptv.de) does not allow that.

5.2 Comparison between MATSIM, VISUM and traffic counts

The VISUM output and the traffic counts are the references to which the volumes produced by the different MATSIM runs will be compared. MATSIM is known to produce results similar to VISUM. Nevertheless, a similarity check is also done by a one-to-one conversion of each trip of the OD-matrices into "one-trip-plans", meaning each trip corresponds to a separate person. Of course, those are not "real" plans but the produced traffic should be similar to the one VISUM generated. We ran the chain generation process that is using the OD-matrices three times with the different average duration assumptions. The three generated plan files are used as initial plan files for MATSIM iterating 50 times by using only the route re-planning module (Raney and Nagel, 2002). Therefore the following results will be compared to the VISUM reference output:

- *OD2Trips*: Direct conversion from trips to one-trip-plans.
- *OD2PlansW8*: Activity chain generation with average durations: 8 hours work, 4 hours education, 4 hours leisure, 1 hour shopping.
- *OD2PlansW9*: Activity chain generation with average durations: 9 hours work, 6 hours education, 5 hours leisure, 2 hours shopping.
- *OD2PlansW10*: Activity chain generation with average durations: 10 hours work, 8 hours education, 6 hours leisure, 3 hours shopping.

For the second activity chain generation process described in Section 4 one additional plan file is generated:

- *MIV*: The end time of the first activity of each plan is uniformly distributed between 6am and 8am. The durations are set as follows: 8 hours work, 6 hours education, 2 hours for each leisure and shop activity.

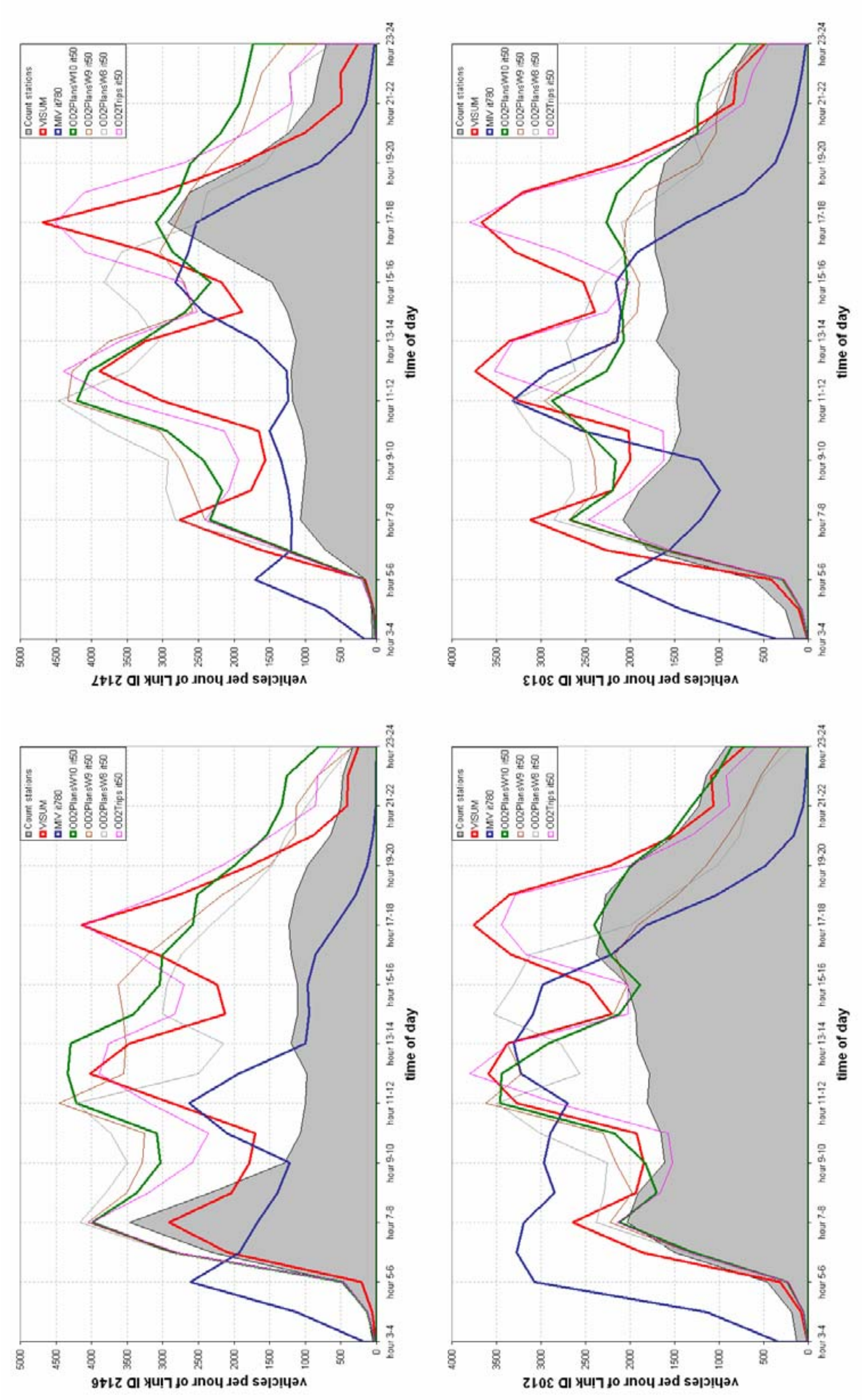
As already mentioned in Section 4 the end time of the first activity is only an initial value. For this run we also use the time re-planning module of MATSIM (Balmer *et al.*, 2004) to allow adaptation of the timing. To calculate adequate scores (utilities) for each executed plan (Raney and Nagel, 2004), MATSIM requires constraints of each type of activity. In this case we define them as follow:

- The work and education opening time is set to 7am
- The work and education closing time is set to 6pm.
- The latest work arrival and education arrival time for the first work activity, education activity resp. is set to 9am.
- The shop opening time is set to 8am.
- The shop closing time is set to 8pm.
- The leisure opening time is set to 1am.
- The leisure closing time is set to 11pm.

5.3 Hourly Volumes

Figure 4 shows the hourly volumes of four of the twelve links for which counts are available. VISUM (bold red curves of Figure 4) shows the three peaks already mentioned in Section 2.2. The volumes differ substantially from the hourly volumes recorded at the count stations (grey area of Figure 4), so a comparison to real world data does not make sense until the calibration of VISEM has been improved.

Figure 4 Hourly volumes of the VISUM and the five MATSIM runs compared to field volumes



VISUM vs. MATSIM with input plans created by chain generation process using OD-matrices

The comparison between VISUM and OD2Trips (light pink curves) shows similarities comparable to those reported in (Raney and Nagel, 2002). That again confirms that VISUM and MATSIM produce comparable results.

At a first sight, the hourly volumes of OD2PlansW8 (light grey curves of Figure 4) differ strongly from the VISUM volumes. This is not really surprising since the activity durations are just assumptions. Anyways, some similarities are still observable:

- The three peaks are present in both curves.
- The morning peak matches pretty well.
- The other peaks appear too early in the day (noon peak around 11am, evening peak around 4pm).

This leads us to the other two duration definitions (OD2PlansW9 and OD2PlansW10). Especially OD2PlansW10 (bold green curves of Figure 4) matches the first two peaks quite well. Also, the third peak appears at the right time, but with too low volumes while later in the night, the volumes are too large.

The reason for this difference is again in the duration definition. The duration of a chain can differ between 3 hours (h-s-h chain) to 16 hours (h-w-l-w-h). If more activity chains were used, the duration of a chain could go up to 27 hours (h-w-l-e-s-h)! In one sentence, short chains usually produce too short out-of-home durations while the reverse is true for longer ones. The results are still surprisingly good for the crudeness of the assumptions made.

VISUM vs. MATSIM with input plans created by chain generation process without using OD-matrices

It is not surprising that the volumes produced by MATSIM using the demand from the MIV generation process (bold blue curve of Figure 4) are quite different to VISUM since they use different input datasets. Nevertheless, it is quite questionable why the MIV-process also produces a peak around noon. Since the network links on which the volumes are measured are located in an urban area (the city of Zurich), the noon peak could be produced by agents who are performing shop and leisure activities between two work or education activities. Since the simulation does not allow mode change, all agents are using their vehicle to reach such an intermediate activity. On the other hand, it is more typical that individuals who work in the city are having i.e. lunch break next to their offices.

It is also to mention that in three of the four volume measurements MIV produces much less all day volumes than the VISUM run. The reason for that lies into the fact that the agents are not that strongly bound to a fixed time for leaving an activity. They are able to rearrange the times when they travel and also which route they choose.

Field Data vs. MATSIM with input plans created by chain generation process without using OD-matrices

Apart from the above mentioned noon peak the volumes produced by the MIV model (bold blue curve of Figure 4) are quite similar to the field data (grey area of Figure 4), but are showing one difference: The MIV curve has a shift of about two hours compared to the field data. The cause of this shift could be the fixed time windows described in Section 5.2. Instead of defining the work activity opening time between 7am and 9am, we could shift it more towards noon. Note that the MATSIM simulation process had not yet reached a stable state when it was producing the outcome of iteration 780. With the defined initial end time of the first activity between 6am and 8am the simulation process of MATSIM is still attempting to reach the time 7am to 9am window. Therefore a more appropriate setup will reduce the occurred bias.

Comparison of volumes on other links of the Scenario

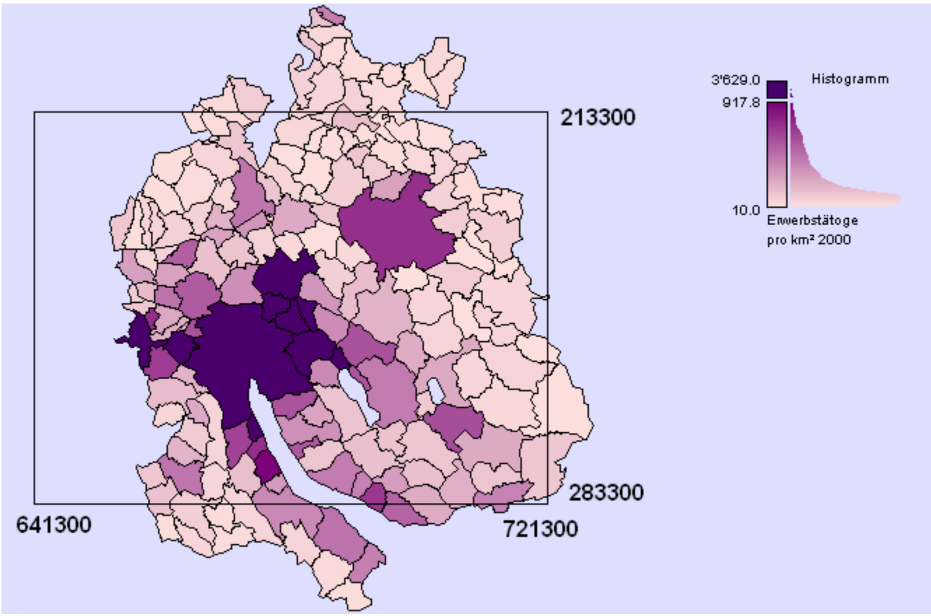
Comparing the volumes produces by VISUM, the different MATSIM runs and available count data on other links in the city of Zurich (i.e. link ID 3192 and 3193 of Figure 3) similar plots can be generated as described above. But the more we get to the border of the Zurich area (count station ID 178, 213 and 221 of Figure 3) the less traffic is produced by MATSIM and VISUM compared to the traffic counts. This is expected since we modelled only intra-kantonal demand. Therefore, all transit traffic passing through the Kanton of Zurich is missing. This hints on the importance of the right definition of the study area for the outcome of a scenario analysis.

5.4 Comparison to Land use data

Both activity chain generation processes uses the location choice process described in Section 3.3. As mentioned it does not respect the capacity constraints of work places, shopping areas, etc. In this section we now want to analyse what the differences are between the given land use data and the occupancy generated by the two processes.

Figure 5 shows the labour force density for each municipality of the Zurich area. We can clearly see that Zurich itself and also the region of Winterthur contain a higher density. The drawn square indicates the area in which we will compare occupancy against land use data.

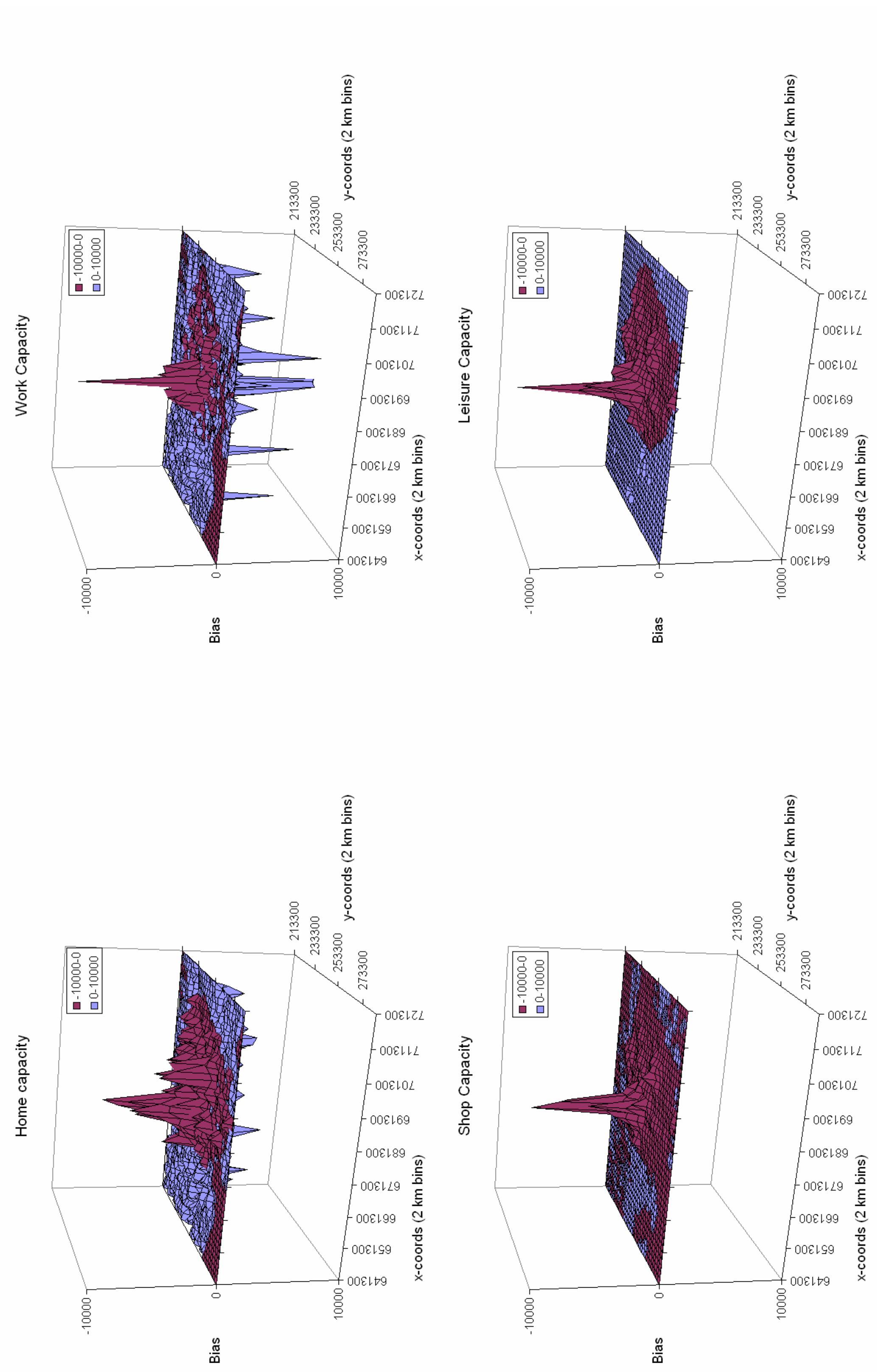
Figure 5 Labour force density for the municipalities of Kanton Zurich



Source: Statistisches Amt Kanton Zurich (http://www.statistik.zh.ch/)

Since both demand generation process uses only 45.44% of the whole population we need to aggregate the out coming occupancies to 100% again otherwise a direct comparison to land use is not expressive. Figure 6 and Figure 7 shows the difference of occupancy generated by the processes and land use spilt up into four of the five activity types, namely “home”, “work”, “shop” and “leisure”. The resolution is a square of size 2x2 kilometres. The light blue areas indicate that the capacity given by the land use is not reached by the process while the dark red area shows that the process produced occupancy is higher than the given capacity.

Figure 6 Difference between given land use data and MIV setup



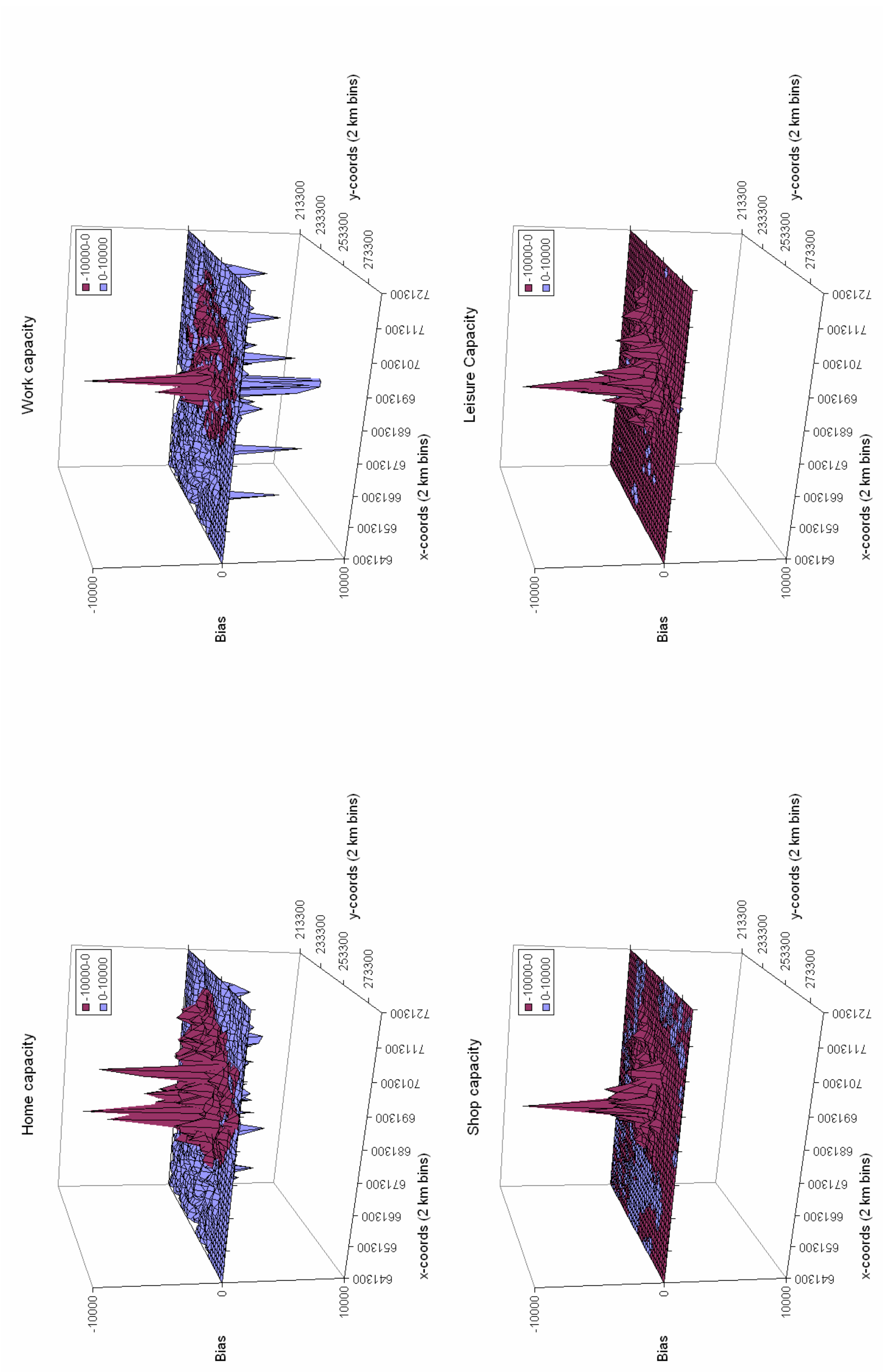
Land use data compared with MIV run

In all four activity types the location choice produces higher occupancy than there is available space for the City of Zurich and partially also for Winterthur (see Figure 6). The other regions are underestimated (especially for “work” activity). Interestingly, in downtown of Zurich there are much less work activities than expected. The reason for that could be of the high peak in the north-eastern part of the city of Zurich. The location choice more often places work activities there because the Zurich industry is taken place at this area. The type “leisure” is almost only chosen at the lakeside of Zurich, meaning that almost everybody travels to the city of Zurich for recreation, which is not true in reality since the land use clearly shows that there are other attractive recreation areas in the Kanton of Zurich.

Land use data compared with OD2PlansW10 run

The conclusions we made for the above comparison are true again. The location choice process does not respect the constraints given by the capacities of the land use. The overestimations occur again at the same areas as above (see Figure 7).

Figure 7 Difference between given land use data and OD2PlansW10 setup



6. Conclusion and Future Work

This paper shows two different ways to produce daily traffic demand of each individual for a given area. The first one is based on given origin-destination matrices, while the second one uses demographical and statistical data of the scenario. Both methods do have some advantages and disadvantages:

Activity chain generation process based on OD-matrices

Given origin-destination matrices on an hourly based resolution and some additional demographic data, this process is able to reconstruct a daily demand for each individual in the examined region. Up to a certain degree it can be shown that traffic produced by the generated daily demand reflect the traffic produced by the OD-matrices with the advantage that the daily demand is based on individual plans for each agent.

Nevertheless, the process uses some assumptions, namely the typical durations of the generated activities. As shown in Section 5 several runs have to be done to find out which assumptions produce the best match compared to the given matrices. And it is difficult to see if the above chosen durations could be further optimized.

The location choice process which is used to calculate the locations for the secondary activities has the disadvantage that it does not respect the constraints given by the land use data. It should be modified or substituted by i.e. a gravity model. Since the general activity chain generation process consists of several independent components it is quite easy to replace parts of the process by more enhanced models.

At last some computational issues are to be mentioned. The activity chain generation process needs about one hour to generate daily demand for about 580000 individuals one a single 1.8 GHz CPU. Additionally about 4 hours is needed to generate the secondary activity locations. MATSIM needs about 18 hours for 20 iterations (reaching a relaxed state using only route re-planning described in Balmer *et al.*, 2004) on a Beowulf cluster using eight CPUs. Therefore the whole computation can be done in about 1 day.

Activity chain generation process without using OD-matrices

If the required input data is available the process described in Section 4 preferred over the other ones since constructing hourly based origin-destination matrices from daily demand for each individual is a trivial task. Also the resulting volumes match the reality quite better than the process described in Section 3.

But again, some assumptions have to be made: The opening and closing times for the different activities have to be chosen carefully. The setup described above made some assumptions which did not produce a good match with reality.

Since this process uses the location choice module as well, the above mentioned problem of not respected constraints from land use also occurs in this case.

An important issue is the computational performance: The entire activity chain generation process needs about 10 minutes with additional 4 hours for the location choice on a 2.4 GHz single CPU, which is quite fast. But the results used for the comparison to other approaches in this study are taken from iteration 780! That means, MATSIM has to run for several weeks to produce the volumes. Even worse, at iteration 780 the process has not yet reached a complete relaxed state. The core of this problem lies into the used time re-planning module described in Raney and Nagel (2004). On the other hand, it can be expected that a more sophisticated re-scheduler decreases the amount of iteration substantially.

This paper shows that—in principle—the activity chain generation approach produces realistic demand, but is highly dependent on reliable input data and sophisticated process steps. It is also of interests to see that some parts of the processes are crucial while others (i.e. distribution or disaggregation shown in Figure 2) are not.

7. References

ARE - Bundesamt für Raumentwicklung und Bundesamt für Statistik (2001). Mobilität in der Schweiz, Ergebnisse des Mikrozensus 2000 zum Verkehrsverhalten, Bern und Neuchâtel.

ASTRA - Swiss Federal Road Authority (2000). Automatic traffic counts 1999. Bern, Switzerland. http://www.astra.admin.ch (Accessed Feb. 2005).

Balmer, M., B. Raney, and K. Nagel. (2004) Agent-Based Activities Planning for an Iterative Traffic Simulation of Switzerland – Activity Time Allocation. *Swiss Transport Research Conference (STRC), Monte Verita, Switzerland*. http://www.strc.ch (Accessed Feb. 2005).

BFS - Swiss Federal Statistical Office. http://www.statistik.admin.ch/eindex.htm (Accessed Feb. 2005).

Charypar, D. and K. Nagel.(2003) Generating complete all-day activity plans with genetic algorithms, *10th International Conferece on Travel Behaviour Research, Lucerne, Switzerland*. http://www.ivt.baum.ethz.ch/allgemein/iatbr2003.html. (Accessed Feb. 2005).

Marchal, M. and K. Nagel (2004). Modeling location choice in activity-based models with cooperative agents. *Swiss Transport Research Conference (STRC), Monte Verita, Switzerland*. http://www.strc.ch (Accessed Feb. 2005).

MATSIM - multi agent traffic simulation. http://www.matsim.org (Accessed Feb. 2005).

Raney, B. and K. Nagel (2002). Iterative route planning for modular transportation simulation. *Swiss Transport Research Conference (STRC), Monte Verita, Switzerland*. http://www.strc.ch (Accessed Feb. 2005).

Raney, B. and K. Nagel. (2004) An improved framework for large-scale multi-agent simulations of travel behaviour. *Swiss Transport Research Conference (STRC), Monte Verita, Switzerland 2004*. http://www.strc.ch (Accessed Feb. 2005).

Rieser, M. (2004) Berechnung von Nachfragematrizen mit VISEM. Semester thesis, Institute for Transportation Planning and Systems (IVT), ETH Zurich, Switzerland.

Vovsha, P.; E. Petersen and R. Donnelly (2002). Microsimulation in Travel Demand Modeling: Lessons Learned from the New York Best Practice Model. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1805*, pp. 68-77, TRB, National Research Council, Washington, D.C.

Vrtic, M., Ph. Fröhlich and K.W. Axhausen (2002). Schweizerische Netzmodelle für Strassen- und Schienenverkehr, in *T. Bieger, C. Laesser und R. Maggi (Hrsg.) Jahrbuch 2002/2003 Schweizerische Verkehrswirtschaft,* 119-140, SVWG, St. Gallen, Schweiz.

Vrtic, M. and K.W. Axhausen (2003) Experiment mit einem dynamischen Umlegungsverfahren, *Strassenverkehrstechnik*, **47**(3), pp 121-126.