Application of the VISEVA demand generation software to Berlin using publicly available behavioral data

Submission date: 15-Nov-06

Words: 5973
Figures and tables: 6 (= 1500 Words)
Total: approx. 7500

Andreas Justen

DLR - German Aerospace Center
Institute of Transport Research
Rutherfordstraße 2
12489 Berlin

Phone: +49 30 670 55 23 4
Fax: +49 30 670 55 20 2
andreas.justen@dlr.de

Ulrike Beuck

TU Berlin
Institute for Land and Sea Transport Systems
Transport Systems Planning and Transport Telematics (VSP)
Salzufer 17-19, SG12
Germany

Phone: +49 30 314 29 52 1
Fax: +49 30 314 26 26 9
Email: beuck@vsp.tu-berlin.de

Kai Nagel

TU Berlin

Institute for Land and Sea Transport Systems
Transport Systems Planning and Transport Telematics (VSP)
Germany

Phone: +49 30 314 23 30 8
Fax: +49 30 314 26 26 9
Email: nagel@vsp.tu-berlin.de

ABSTRACT

In this paper the EVA algorithm developed by Lohse is applied in order to generate Berlin's average workday traffic based on a minimum of data input. Behavioral parameters are derived from the German travel survey "Mobilität in Deutschland (MiD)". The EVA approach allows generating trip purpose and time dependent OD matrices from general input data used in transport modeling. This model output can be used for standard OD-matrix-based static or dynamic assignment, but provides us with primary activity location choice and scheduling information necessary to generate initial conditions for agent-based transport simulation packages like MATSIM.

The paper describes the basic concept of the EVA model and specifications of the Berlin scenario. Since the range of possible input data for demand generation is limited, our aim was to use the established demand generation model VISEVA with a minimum of input data, which has to be commonly available and easy to purchase (making transfer of transport models to other study areas easier).

The model output is displayed and compared with output resulting from Berlin's official demand generation model. Besides that, the simulation results are compared to real-world data from traffic counts. It can be shown that even though we reduce data requirements to a minimum, the results have a structure adequate for Berlin and could serve as input for initial condition generation for MATSIM.

INTRODUCTION

Most transport models used in practice apply the four-step process. The first three steps—trip generation, destination choice, mode choice—concern modeling the demand, finally described in terms of origin-destination (OD) matrices. In these three steps, various characteristics of the traveler, the land-use, and the network are brought together. In the fourth step, the demand is assigned to the network.

There is widespread agreement that the four-step process, in its conventional form, is unable to capture important aspects of transport planning. This concerns, in particular, all temporal aspects, such as peak spreading, congestion toll modeling, or important environmental aspects (e.g. tailpipe emissions, depending on engine temperature).

The first step to improve this situation is to run separate demand generation and network assignments for the morning and the afternoon peak. This is, however, increasingly problematic with the increase of non-home-based trips. A reaction to this situation is activity-based demand generation (ABDG; see, e.g., (*1, 2*)), where travel is seen as demand that is *derived* from the demand to perform different activities at different locations. However, despite much progress, ABDG is at this point not very much standardized: there are many different models and implementations around (*3, 4, 5, 6, 7, 8*). This is due to a wide variety of different approaches, for example concerning the methods (e.g. Random Utility Modeling vs. rule-based systems) or the level of detail/resolution (e.g. based on half tours, full tours, or complete day plans). Experience and diligent investigation will hopefully demonstrate the respective advantages and disadvantages of each method.

In the meantime, it makes sense to consider alternative methods, which remove some of the disadvantages of the four-step process, while not going the full distance towards ABDG. One such model is the EVA modeling approach of Lohse et al. (*9*). It extends the methodology of the traditional four-step process, which is essentially a method to generate OD matrices for home-based trips, to a methodology to generate OD matrices that connect arbitrary trip purposes. For example, there will be the typical OD matrices for home–work and work–home, but there will also be matrices for, say, work–shop or work–leisure. OD matrices may, in addition, be segmented by demographic groups. More details are provided later in the paper. EVA has been developed over many decades, including a sound mathematical foundation based on probability theory. A further advantage is that it is now publicly available as VISEVA as part of the PTV transportation planning package (*10*), thus providing a standardized access to the package allowing scientifically sound comparisons of results.

In this situation, it would be convenient if it ran from standardized and easily available data. In Germany, such a data set is the “Mobilität in Deutschland (MiD)” data set (*11, 12*). It is essentially a micro-data sample of the German population, with special emphasis on transport-related questions. Unfortunately, most geo-coding was removed from the data set before it was made available to us. Nevertheless, it is a good starting point, in particular since it is available in standardized form for all of Germany. The main question to be answered in the present paper is, in consequence, in how far this data set, possibly augmented by other publicly available sources, is able to provide useful input for the VISEVA demand generation package.

An additional use of such a VISEVA run would be to use it as input to our multi-agent traffic simulation package, MATSIM (*13, 14, 15*)*.* This is particularly appropriate since Germany does not provide data-driven “commuting matrices” that are resolved beyond the city level, and in consequence the coupling between residences and work locations has to be model generated. In this situation, having a standard package such as VISEVA based on standard input data such as MiD appears like a good first step to make progress.

VISEVA—BASIC CONCEPT

The applied EVA algorithm developed by Lohse (*9*) handles trip generation, trip distribution and mode choice simultaneously. This algorithm—implemented in the commercial software package VISEVA (*16*) distributed by PTV AG—is a disaggregate description of the demand. The demand is disaggregated into activity-purpose pairs at origin and destination zones. Trip generation, distribution and mode choice are based on the activity-purpose pair classification. Each activity-purpose pair associates with a certain trip purpose, e.g. the home-work pair contains trips from home to work, and can be associated with all or a subgroup (behaviorally homogenous groups) of travelers. Only employed persons leave home to work, this means that the activity-purpose pair home-work is associated with employees. Other pairs like home-education contain trips of, for instance, high school and university students. The concept of activity-purpose pairs allows obtaining matrices by trip purpose. Summing up these matrices gives total demand in a defined period.

Basically, one can define any classification of activity-purpose pairs that serves the specific problem best. Some standard classifications have already proven their usefulness. We chose one with 6 activity-purposes resulting in 13 activity-purpose pairs (table 1). The pairs can be grouped into types according to the location of the home activity at origin or destination (type 1 and 2). Work can also be the home activity, when the pair lacks the original home activity. Activity purpose pairs containing neither home nor work at origin or destination are of type 3.

Trip production is calculated with trip rates per activity-purpose pair at origin according to its type. At destinations, trip attractions are calculated as a proportional to the capacity of activity opportunities. These capacities can be used as hard or soft constraints. Generally, for primary activities capacities are modeled as hard constraints. Soft constraints allow exceeding the given capacity to a certain degree. Thus, only an upper limit can be set at first; the final number of attractions (number of trips of a certain activity-purpose pair attracted by a zone) cannot be defined without joint trip distribution and mode choice. That is, spatial competition can be modeled (e.g. different shopping locations). To distinguish hard and soft constraints is an advantage of the EVA approach compared to simple destination choice models, which only enforce the constraint at the origin.

As just mentioned, the calculation starts with calculating trip production for each activity-purpose pair at the home location (according to the activity-purpose pair's type) in each zone.

$$H_e = \sum_p TP_p \cdot BP_{ep} \cdot u_p \qquad V = \sum_e H_e \qquad (1)$$

with:
TP_p production rate of person group p
BP_{ep} number of persons of group p in zone e
u_p share of intrazonal trips for group p in zone e

with:
H_e trips at home location (according to the activity-purpose pair's type)
V total number of trips of the activity purpose pair in the study area

Trip attractions can be derived (normalized according to the sum of trip productions over all zones) when hard constraints are given.

$$Z_j = \frac{\sum_r ER_r \cdot SZ_{rj}}{\sum_{j'} \sum_r ER_r \cdot SZ_{rj'}} \cdot V \qquad (2)$$

with:
Z_j traffic attracted to zone j
ER_r attraction rate of attractor r
SZ_{rj} volume of attractor r in zone j

For soft constraints only an upper limit value is calculated as already mentioned.

$$Z\max_j = \frac{\sum_r F_{rj} \cdot ER_r \cdot SZ_{rj}}{\sum_{j'} \sum_r ER_r \cdot SZ_{rj'}} \cdot V \qquad (3)$$

with:
F_{rj} additional load factor of zone j considering attractor r

The EVA model applies its activity-purpose pair approach per subgroup of travelers to the joint destination and mode choice as well. The marginals of the generated matrices are known (in case of soft constraints as maximum number of trips) and the share of trips with mode k between zones i and j are calculated as a function of the generalized costs of travel using different model forms. This conditional probability is:

$$BW_{ijk} = P\left(W \left|\left(A_i \cap E_j \cap M_k\right)\right.\right) \tag{4}$$

with randomly chosen probabilities that

A_i zone i is origin
E_i zone j is destination
M_k mode k is used
W trip from i to j using k is accepted with regard to the generalized costs

Although an arbitrary function can be used transform transport costs w into probabilities, we used the EVA function (*16*), which obtains with its three parameters E, F, G a flexible shape of the elasticity ε over the range of the generalized costs.

$$BW = f(w) = \left[1+\left(\frac{w}{F}\right)^G\right]^{-\frac{E}{G}} \qquad \varepsilon(w) = \frac{df/f}{dw/w} = -E\cdot\frac{w^G}{F^G+w^G} \tag{5}$$

The parameters of the EVA function can further be differentiated according to the subgroups of travelers. Traffic flows (v_{ijk}) are calculated considering simultaneously the generalized costs, the probabilities of the events $P(A_i)$, $P(E_j)$ and $P(M_k)$ and the constraints with respect to (maximum) traffic volumes at origin and destination zones. The formulation is structurally a Bayesian model:

$$v_{ijk} = \frac{P\left(\left(A_i \cap E_j \cap M_k\right)\middle|W\right)}{\sum_{i'}\sum_{j'}\sum_{k'} P\left(\left(A_{i'} \cap E_{j'} \cap M_{k'}\right)\middle|W\right)}\cdot V = \frac{P(A_i)\cdot P(E_j)\cdot P(M_k)\cdot P\left(W\middle|\left(A_i \cap E_j \cap M_k\right)\right)}{\sum_{i'}\sum_{j'}\sum_{k'} P(A_{i'})\cdot P(E_{j'})\cdot P(M_{k'})\cdot P\left(W\middle|\left(A_{i'} \cap E_{j'} \cap M_{k'}\right)\right)}\cdot V \tag{6}$$

Further explanations and solution algorithms are described more in detail in (*9*), (*17*) or (*18*).

It has to be mentioned that there is a need to iterate between the travel demand calculations, following the EVA approach, and the assignment, generating the travel cost values, to obtain a mutually consistent solution. The software tool VISEVA (*16*) provides tools to implement an iteration scheme in conjunction with the assignment software VISUM (*10*).

ESTIMATION OF DEMAND FOR AVERAGE WORKDAY TRAFFIC—BERLIN SCENARIO

Following the above described steps in VISEVA, average workday traffic is derived from general input data on land-use and population. Two different models were built. In the first model (model I), demand was derived from land-use data and the population without any further differentiation. The second model (model II) distinguishes subgroups of the population (homogeneous behavioral groups). The second model is more flexible, but we were also interested how well demand derived from an undifferentiated population will perform.

Both models have the same definition of activity-purpose pairs. Based on 6 activities, 13 activity-purpose pairs can be distinguished.

<table>
<tr><th></th><th>Home
(H)</th><th>Work
(W)</th><th>Kindergarten
(K)</th><th>Education
(E)</th><th>Shopping
(S)</th><th>Others
(O)</th></tr>
<tr><th>Home
(H)</th><td>-</td><td>HW</td><td>HK</td><td>HE</td><td>HS</td><td>HO</td></tr>
<tr><th>Work
(W)</th><td>WH</td><td>-</td><td colspan="4">WO</td></tr>
<tr><th>Kindergarten
(K)</th><td>KH</td><td rowspan="4">OW</td><td colspan="4" rowspan="4">OO</td></tr>
<tr><th>Education
(E)</th><td>EH</td></tr>
<tr><th>Shopping
(S)</th><td>SH</td></tr>
<tr><th>Other
(O)</th><td>OH</td></tr>
</table>

TABLE 1: Activity-purpose pairs definition

In total for all four modeled modes (motorized private travel, transit, walking, biking) 52 matrices of 889 x 889 zones are calculated for the first model with no population differentiation. The second model calculates for each subgroup 52 matrices of 889 x 889 zones. The following subgroups were used:

- Employees with car
- Employees without car
- Non-employed people with car
- Non-employed people without car
- University student
- Pupils/ high school students
- Children < 6 years old
- Apprentice

Even though it would have been possible to use different parameters of the EVA function and of the modal split values according to the specific subgroup of travelers and activity-purpose pair, this functionality was not used for the present study. In both models, EVA parameters are chosen based on experiences provided by Lohse (*16, 17*), modal split values were used according to the activity-purpose pair only.

Assigning departure time intervals to demand is possible by using estimations of hourly demand according to the trip purposes. This kind of information can be found in surveys as well. Lohse (*17*) gives an example from previous projects. In order to compare the demand generated in the project described here to the data from the model used by Berlin's planning department, 24-h matrices containing all purposes are calculated (planning department uses only 24-h OD matrices).

Travel costs in terms of travel times are mainly derived from a VISUM model of the road infrastructure of 1998. Further information on available input data and necessary data processing is described in the next section.

INPUT DATA

As already mentioned, a minimum range of input data should be used. Input data is gathered for a period of time around the year 2000. Data for demand generation has to contain information on land-use, composition of population and travel behavior, and the network to obtain travel costs between zones. Data requirements depend among other things on the defined activity-purpose pair classification. Since only a road network was available, travel costs of the other modes (transit, walking, biking) were estimated based on linear distance measured and average mode speed estimated and schedules of the transit, available on the internet (*19*).

Input data were available on different spatially aggregated levels. The largest entities are districts. Up to the year 2000, the study area of Berlin consisted of 23 districts, 195 statistical areas, 881 traffic analysis zones (TAZ), and 15.101 blocks. The 881 TAZ were already included in our road network of 1998, which the planning department provided us with. Outside Berlin 8 zones were defined in order to model commuting travel.

The population is available on block level, but as the only attribute age is given. The land-use of Berlin is based on traffic analysis zone level and consists of capacities for "home", "work", and "education". "Shopping" capacities had to be defined manually (see below). Unfortunately, there is no commuting matrix available for Berlin. Information on commuting is gathered by the authorities between municipalities only and Berlin as a whole is only one municipality. For this scenario there is no commuter matrix available, which would be desirable for a calibration procedure of the VISEVA models.

According to the population subgroups used in production and attraction calculations, behavioral data has to be derived from surveys. The city's planning department based its model on an extensive data record, but we did not have access to it. There is however the German travel survey "Mobilität in Deutschland (MiD)" (*11, 12*), which is easy to access for scientific purposes. Former surveys of 1976, 1982 and 1989 were conducted in western Germany only; MiD 2002 is the first travel survey embracing unified Germany.

The following sections describe how behavioral parameters are derived based on MiD and aspects of land-use data acquisition and processing. Besides giving an overview of relevant land-use data and their sources, the description focuses on processing information about shop floor areas in Berlin in order to model activity-purpose pairs other than work or education related.

Behavioral data

The main source for the generation of behavioral parameters was the nation-wide survey MiD conducted in 2002 by the Federal Ministry of Transport, Building and Urban Affairs (BMVBS). Objective of this survey is the generation of reliable and representative information about socio-demographic aspects of persons and households in combination with their daily travel behavior. This data record distinguishes between the German Federal States. As Berlin is not only a municipality but also a Federal State, data evaluation could be done for Berlin directly, which should assure that specific behavioral aspects of Berlin are recognized. Additionally, MiD provides aggregated data based on certain spatial categories. Spatial categories are classifications by habitants' density, delivered by the Federal Office for Building and Regional Planning.

The overall national sample captured 25,000 households, 62,000 persons and 190,000 reported trips. The Berlin sample of the survey holds 7,616 trips reported by 2,163 Persons (taken only the trip data set as reference—the persons' data set is based on 2,849 interviewed persons). The following parameters were extracted from MiD to be used in our two VISEVA models:

- specific trip production rate of the population of Berlin differentiated by trip purposes (model I)
- specific trip production rates of the population of Berlin differentiated by homogeneous behavioral groups and trip purposes (model II)
- modal split values differentiated according to the needs of model I and II

As modal split values could be extracted rapidly by using standard data processing software, further comments do exclusively concentrate on the generation of the specific trip production rates. Extensive coding of the original MiD trip data set was necessary in order to calculate these essential parameters calculating demand according to the VISEVA approach.

Trip production rate

The specific trip production rate is defined as the average number of trips made by person per day and purpose and is primarily determined by socio-demographic characteristics and availability of a private car. In order to calculate trip production rates to be used in model II, size of subgroups had to be extracted from MiD as well. Average specific trip production rate used for first evaluations of the Berlin trip data set is calculated as follows:

$$TP_p = \frac{Number\ of\ trips\ per\ person\ (of\ subgroup\ p)\ over\ all\ activity\ purpose\ pairs}{Number\ of\ persons\ in\ group\ p} \quad (7)$$

Specific trip production rates by activity-purpose pair to be used in our VISEVA models are calculated as:

$$TP_{px} = \frac{Number\ of\ trips\ per\ person\ (of\ group\ p)\ in\ the\ specific\ activity\ purpose\ pair\ x}{Number\ of\ persons\ in\ group\ p} \quad (8)$$

Person groups

In model II, trip production rates had to be calculated for each subgroup defined. Thus, subgroup size had to be determined (in each activity-purpose pair) as well. In MiD, subgroups are already defined, and two different classifications are available. MiD distinguishes 9 or 12 behaviorally homogeneous groups. For this study, the original classification consisting of 12 subgroups was reduced to 9 subgroups, which is not identical with the original MiD classification into 9 groups. On top of that, the survey itself provides the necessary variables to generate more sophisticated groups. But in context of this study, a classification into 9 subgroups seems sufficient. Table 2 summarizes the size of subgroups and their proportion in total of the trip data set. The group of "others" was not evaluated any further.

Behavioral homogeneous groups	Number of people	Share in %
Employed person with car	674	35.66
Employed person without car	161	8.52
Non-employed person with car	419	22.17
Non-employed person without car	230	12.17
Students with & without car	72	3.81
Apprentice with & without car	38	2.01
Children under the age of 6	99	5.24
Pupils	197	10.42
N (excluding „Others“)	1,890	
Others	273	
Total	**2,163**	**100**

TABLE 2: Group sizes

Number of trips per person and activity-purpose pair

Joining the information of group sizes and trips realized per day (by person and purpose) the specific production rate can be calculated. Substantial coding modifications of the original data set of reported trips were necessary, because the activity-purpose pair definition was not part of MiD. Activity pairs had to be computed based on the reported trip chains of household members. Additionally, we had to aggregate certain trip purposes to match our activity-purpose pair classification. Table 3 summarizes trip production rates by activity-purpose pairs for two different data sets.

	MiD 2002: trips in agglomerations	Share in %	trip production rate		trip production rate	Share in %	MiD 2002: trips in Berlin	
HW	2,873	7.52	**0.260**		**0.231**	6.78	486	**HW**
WH	2,376	6.22	**0.215**		**0.177**	5.21	373	**WH**
HE	523	1.37	**0.047**		**0.059**	1.74	125	**HE**
EH	453	1.19	**0.041**		**0.050**	1.47	105	**EH**
HK	906	2.37	**0.082**		**0.070**	2.07	148	**HK**
KH	786	2.06	**0.071**		**0.059**	1.74	125	**KH**
HS	3,539	9.26	**0.320**		**0.326**	9.56	685	**HS**
SH	4,003	10.47	**0.362**		**0.373**	10.96	785	**SH**
HO	8,922	23.34	**0.807**		**0.776**	22.77	1,631	**HO**
OH	7,769	20.32	**0.702**		**0.681**	19.97	1,431	**OH**
WO	815	2.13	**0.074**		**0.088**	2.57	184	**WO**
OW	443	1.16	**0.040**		**0.041**	1.21	87	**OW**
OO	4,818	12.60	**0.436**		**0.475**	13.94	999	**OO**
Sum	38,226	100	**3.46**		**3.41**	100	7,164	**Sum**
N	11,060						2,102	**N**

TABLE 3: Comparison of trip production rates derived from two different MiD samples (before correction regarding immobile persons)

On the right, production rates derived from the Berlin sample can be seen. Both the average trip production rate with 3.41 daily trips per person and the specific trip rate calculated according to the definition of activity-purpose pairs have plausible values. But the small sample size (N denotes number of persons in the data) is problematic. Therefore, we also derived the trip production rates from the larger sample of all regions classified as spatial category “agglomerations with an outstanding centre“ (left half of the table). Berlin is classified as one such outstanding center.

When comparing Berlin’s trip production rates to the larger sample size on the left, one sees that even though the Berlin sample is small, plausible trip production rates were calculated for our models. Finally, the resulting trip production rates had to be modified according to the information on how many interviewees performed no out of home activity. This modification was necessary, because the production rates were derived from the trip data of mobile persons only. The resulting rates were used in both VISEVA models.

Further model relevant parameters

The MiD data base offers more possibilities for the generation of model relevant parameters. In particular, relevant information for model calibration can be obtained. For example, distributions of travel distances and times can be calculated from the MiD data set. This kind of data can be used both for VISUM runs and for multi-agent simulations. When using MiD for calibration/validation purposes, one has to take care whether very short distance trips can be found in the model.

Synthetic population and land-use data

As basic spatial unit in our models, the traffic analysis zones have to be described by their land-use and population data. Therefore, information from both areas has to be aggregated or disaggregated to this level depending on their original spatial resolution. It was important to us to use only publicly accessible data as input, in order to demonstrate possibilities and restrictions in data acquisition and to guarantee a self-determined process with data in modified applications of the transport model. The extensive processing of shop floor area as an important land-use data used by the model is exemplified in the following section. This data will be also important in multi-agent applications, since shopping opportunities can differ considerably but detailed and disaggregated data on it is hard to get.

Land-use data processing: shop floor areas

Original information about distribution of shop floor areas was available on the spatial level of the 12 new districts (aggregation of the former 23 districts) (*20*). Additional figures existed for selected shopping areas with a high proportion of shop floor area in m² (*21*). Combining both data sets made it possible to distribute data spatially on block level by a sequential approach, as follows. Figure 1 illustrates the procedure for the inner-city district Mitte.

In a first step, the shop floor area (SFA) of specific, individually known stores ("high concentration" in figure 1) was manually assigned to the specific blocks. In a second step, the remaining SFA for each selected shopping area was manually assigned to surrounding blocks ("medium concentration" in figure 1). In a third step, remaining amounts of SFA on the *district* level were distributed proportional to population density per block. Finally, the SFA was re-aggregated from the block level to the TAZ level.

These steps result in a hierarchically organized assignment of information about shop floor areas. Thus, effects of commercial concentration can be integrated into a transportation model. The hierarchical assignment of shop floor area provides a more realistic mapping of transportation attractions throughout the city. In the future, also different shopping activities with different frequencies can be modeled based on the data set created.

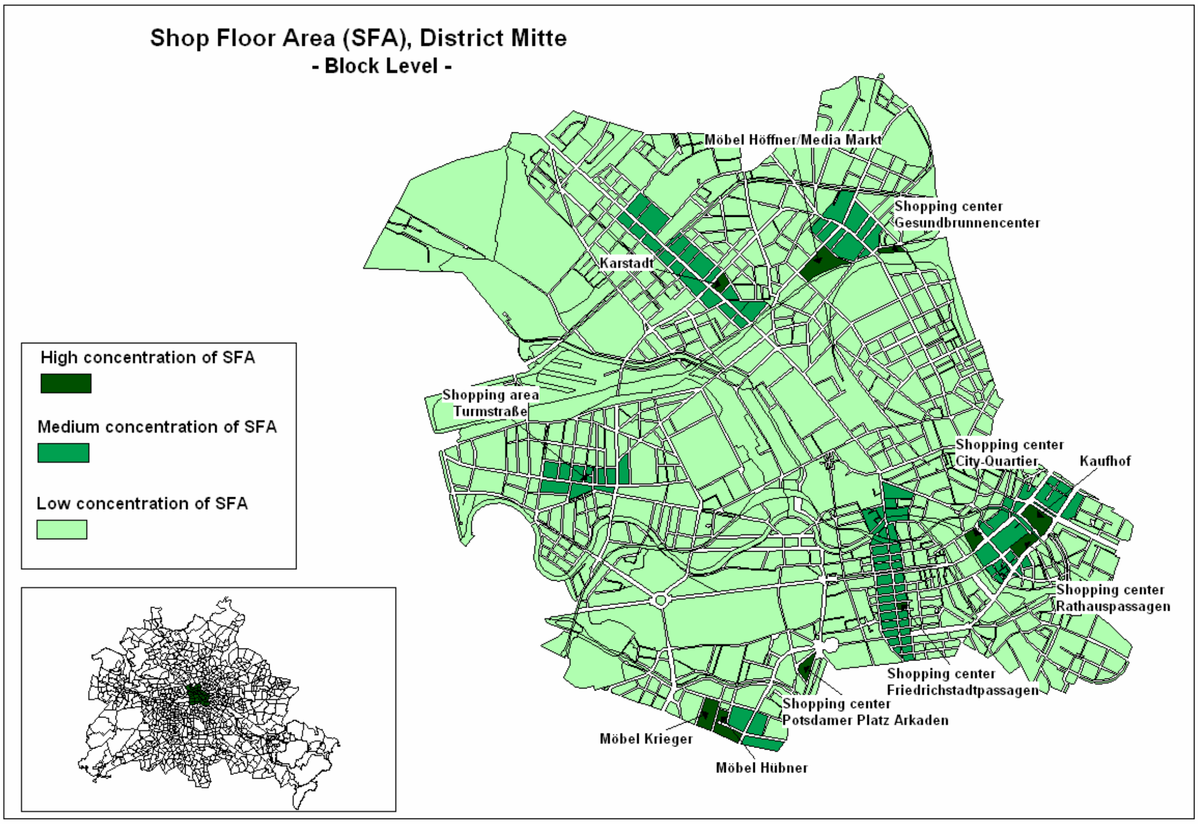

FIGURE 1: Shop floor areas in Berlin's district Mitte

RESULTS

Since the mid 1980s, Berlin's planning department has employed an activity based demand generation model, originally developed by Kutter (*22, 23*), and based on 72 different person groups with common transport behaviors. The detailed information on population composition and behavioral parameters has their source in geo-coded survey data of extensive amount (unfortunately not available to us). This model is used to generate 24-h OD matrices using the same zonal system as we did, at least for Berlin; its surroundings are represented with a higher spatial resolution by the planning department. Since no further detailed output of this official planning model was available, we calculated a 24-h OD matrix of private motorized traffic by summing up the purpose specific matrices in order to compare our model results to the ones of the official model.

First, we compared the values of selected matrix cells of the official planning model and the model I (no subgroup differentiation). Both matrices contain individual motorized traffic of the study area. A comparison of the trip productions is shown in figure 2. As both matrices represent calculation results for Berlin associated traffic, but model the surroundings with the different level of detail, we compare selected 370 TAZ of an inner-city part of Berlin. As it can be seen, the values of the VISEVA model tend to be higher than the comparable trip productions generated with the official planning model. Nonetheless, the structure of Berlin's demand could be reproduced with the simple model I, although the input data was less disaggregated and MiD contained only a small sample size for Berlin.

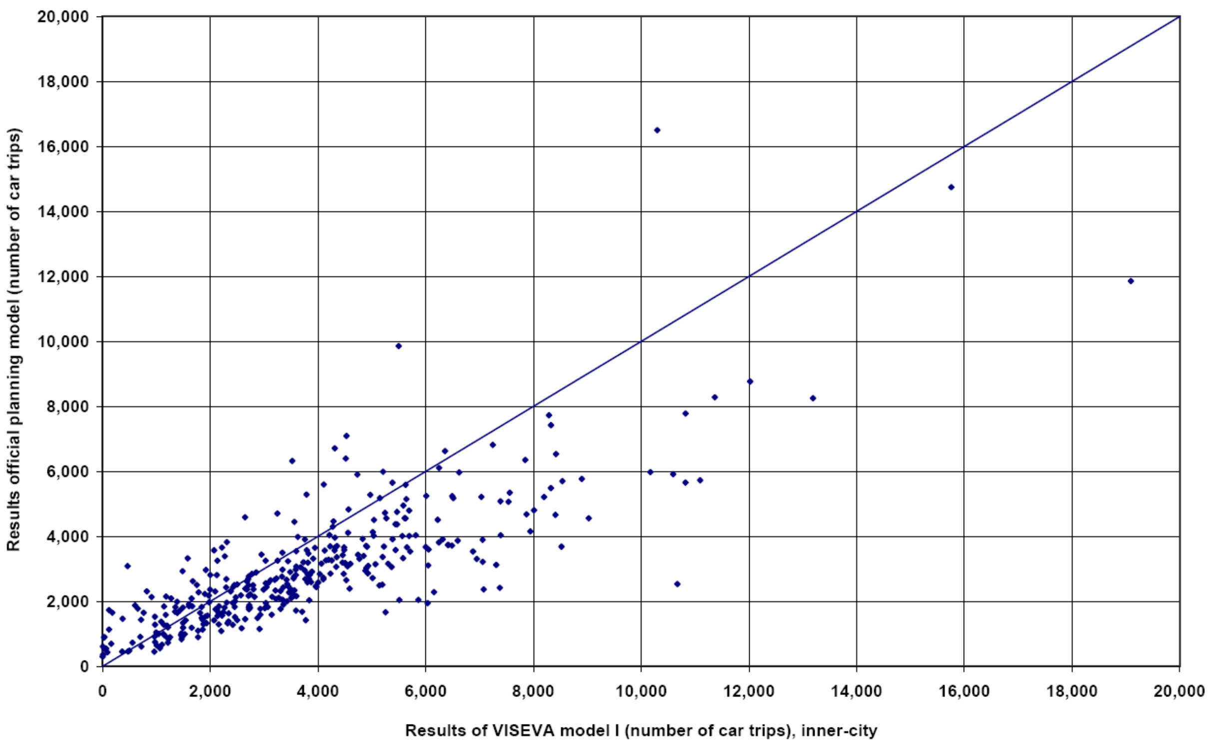

FIGURE 2: Comparison of trip productions generated by VISEVA model I and by the official planning model of Berlin's planning department for 1998 (370 inner-city TAZ included)

Although the reference model was carefully chosen, comparing model output of one model to output of another is not enough. Therefore, for both matrices of individual motorized traffic (generated by the VISEVA model I and the official 24-h matrix of the planning department), we employed the official assignment model and procedure in order to compare the resulting link volumes to real-world counts. The planning department provided us with count data, but only four counting stations had counts for 24 hours. Three of these four 24-h counts could be assigned to the simulation network. Therefore, only these three stations can be used for this comparison. At every station volumes of different vehicle classes of both directions were measured.

Both VISEVA models generate individual motorized traffic only. The official assignment model and procedure include matrices of commercial and long-distance traffic passing through Berlin (through traffic) as well. These additional road user segments affect the route choice of individual motorized traffic. By applying an assignment procedure assigning all three road user segments simultaneously, as in the official assignment model, we make sure to capture this effect.

In order to use the matrices for commercial and long-distance traffic used in the official model, we had to adapt the spatial resolution of the official and the VISEVA model. This concerns the traffic analysis zones of Berlin's surroundings. In the official model, this region is described by 139 TAZ in our VISEVA model we made use of only 8 TAZ in the immediate vicinity of Berlin. Since our study focuses on the city of Berlin, this modeling approach seems suitable. Also the comparison of link volumes is conducted for the Berlin area only. In a pre-test we made sure that the results of the spatially adapted official model are similar to the results of the official model with its original spatial resolution.

Table 4 shows results of this comparison of the official model (spatially adapted) and the VISEVA model I. Presented are simulated 24-h volumes of individual motorized traffic and counts of vehicles classified as cars (which could also be used by commercial and through traffic). At each counting station the measured volumes for both directions are compared to the simulated volumes in the VISEVA model I and the official model. Additionally, relative errors are calculated for each pair of volumes (cars measured and simulated).

Since the available traffic count data does not provide us with the explicit counts for individual motorized traffic, we expected that both, the official planning model and our VISEVA model, will give lower traffic volumes of passenger cars as counted in the real-world. Table 4 proves this assumption right; no simulated volume exceeds the measurements, which can be directly related to the fact that vehicles detected as passenger cars are also used by commercial and through traffic.

	counts	VISEVA model I	official model	district
from node -> to node	1925 -> 2748	2307 -> 3225	2307 -> 3225	Wedding
volume/24h	**25,783**	**19,235**	**15,116**	
rel. error		**0.254**	**0.414**	
from node -> to node	2748 -> 1925	3225 -> 2307	3225 -> 2307	Wedding
volume/24h	**25,418**	**19,252**	**15,343**	
rel. error		**0.243**	**0.396**	
from node -> to node	3372 -> 3301	3910 -> 3908	3910 -> 3908	Neukölln
volume/24h	**17,409**	**14,353**	**11,862**	
rel. error		**0.176**	**0.319**	
from node -> to node	3301 -> 3372	3908 -> 3910	3908 -> 3910	Neukölln
volume/24h	**17,882**	**13,538**	**11,373**	
rel. error		**0.243**	**0.364**	
from node -> to node	1171 -> 1708	1361 -> 2075	1361 -> 2075	Spandau
volume/24h	**23,320**	**13,891**	**16,186**	
rel. error		**0.404**	**0.306**	
from node -> to node	1708 -> 1171	2075 -> 1361	2075 -> 1361	Spandau
volume/24h	**24,943**	**13,219**	**15,523**	
rel. error		**0.470**	**0.378**	

TABLE 4: Comparison with real-world traffic counts

The relative errors are relatively high and it can be questioned whether these errors can be completely explained by the missing road users using passenger cars as well. To answer this question, additional information on location was included in this comparison (last column). All of the measurements were taken from road segments not part of the inner-city area (370 selected TAZ), and two of the three stations are feeder roads. Those are circumstances, which explain the high error values of both models rather well. On average, relative errors of the volumes generated by the VISEVA demand are lower.

CONCLUSIONS AND FUTURE WORK

The results presented in the previous section look promising. With a minimum amount of publicly available input data we derived Berlin's demand, which allows comparison to results derived from the official planning model. The demand was derived by applying the EVA modeling approach. As it was stated in the introduction, this modeling approach can help to overcome some of the shortcomings the traditional four-step modeling approach has. With input data easily to get in almost any study region one can get fairly good results. At the same time it offers possibilities differentiating the demand model. Additionally one can assign departure time intervals to the purpose differentiated matrices. VISEVA produces hourly demand based on hourly shares on daily traffic, specific to activity-purpose pairs. Thus, time and purpose dependent OD matrices produced with VISEVA shall also be sufficient as input for generation of initial agents' plans. Such initial plans could be used as input to MATSIM iterative optimization process. In such a multi-agent simulation framework individualized information is processed and maintained at every level. Such an approach captures the temporal effects of traffic and would allow to model reactions to toll differentiated by demographics. When relaxation is reached in such a simulation, MiD can provide further information for validation. Simulated trip length and time distribution can be compared to the corresponding data reported in MiD.

Of course there are several possibilities to improve the two VISEVA models presented. Obviously, differentiated behavioral parameters the parameters of the EVA function and the modal split, which could not only be differentiated by activity-purpose pairs but also by subgroups model, have to be included in VISEVA model II. Another aspect concerning both VISEVA models is to differentiate traffic related to shopping and other not further specified purposes, but a refining like this could be also introduced while transforming the time and purpose dependent OD matrices into agents' plans. A strong argument supporting the latter is that shop floor areas were made available on block level that is the basis of all spatial entities in Berlin.

Generally, when transforming VISEVA output into agents' plans the demand to date should be represented. Especially the network changed considerably from 1998 up to today. But also land-use data change a lot (also this data should be rather easy to obtain). Both environment changes can be partly explained by Berlin's special situation after the reunification of Germany. New VISEVA runs with updated data could solve this.

Finally we want to point out that Berlin's surroundings—namely the federal state of Brandenburg—were only modeled very roughly (only commuting traffic was modeled). As long as we are interested in the inner-city area this is not problematic. If the focus changes, the surroundings have to be modeled more precisely.

Concluding, it can be stated that it was possible to model the structure of Berlin's demand for individual motorized traffic. When the purpose differentiated matrices are also time dependent, which can be done with VISEVA, we can produce initial plans for our multi-agent simulation of individual motorized traffic in Berlin.

REFERENCES

1. Hensher, D. and J. King, editors. *The Leading Edge of Travel Behavior Research*. Pergamon, Oxford, 2001.
2. Timmermans, H.J.P., editor. *Progress in activity-based analysis*. Elsevier, Oxford, UK, In press.
3. Jonnalagadda, J., N. Freedman, W.A. Davidson, and J.D. Hunt. Development of microsimulation activity-based model for San Francisco: destination and mode choice models. *Transportation Research Record*, 1777:25–35, 2001.
4. Bowman, J. L., M. Bradley, Y. Shiftan, T.K. Lawton, and M. Ben-Akiva. Demonstration of an activity-based model for Portland. In *World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research 1998*, volume 3, pages 171–184. Elsevier, Oxford, 1999.
5. Pendyala, R. M. and R. Kitamura. FAMOS: The Florida Activity Mobility Simulator. In H.J.P. Timmermans, editor, *Progress in activity-based analysis*. Elsevier, Oxford, UK, In press.
6. Miller E. J. and M.J. Roorda. A prototype model of household activity/travel scheduling. *Transportation Research Record*, 1831:114–121, 2003.
7. Arentze, T. A. and H.J.P. Timmermans. *Albatross: A Learning-Based Transportation Oriented Simulation System*. EIRASS (European Institute of Retailing and Services Studies), TU Eindhoven, NL, 2000.
8. Arentze, T. A. and H.J.P. Timmermans. *ALBATROSS – Version 2.0 – A learning based transportation oriented simulation system*. EIRASS (European Institute of Retailing and Services Studies), TU Eindhoven, NL, 2005.
9. Lohse, D. Teichert, B. Dugge and G. Bachner. Ermittlung von Verkehrsströmen mit n-linearen Gleichungssystemen – Verkehrsnachfragemodellierung, *Schriftenreihe des Instituts für Verkehrsplanung und Straßenverkehr der TU Dresden*. Heft 5/1997, Fakultät Verkehrswissenschaften „Friedrich List", TU Dresden 1997.
10. PTV www page, accessed 7/2006. URL www.ptv.de
11. Bundesministerium für Verkehr, Bau- und Wohnungswesen (Ed.). *Ergebnisbericht Mobilität in Deutschland*. Berlin, 2004.
12. Deutsches Institut für Wirtschaftsforschung und Institut für angewandte Sozialwissenschaft (Ed.): *Mobilität in Deutschland 2002. Kontinuierliche Erhebung zum Verkehrsverhalten.* Datensatzbeschreibung. Bonn/Berlin 2004.
13. Balmer M., B. Raney, and K. Nagel. Adjustment of activity timing and duration in an agent-based traffic flow simulation. In H.J.P. Timmermans, editor, *Progress in activity-based analysis*. Elsevier, Oxford, UK, 2005. Also see www.vsp.tu-berlin.de/publications.
14. MATSIM www page, accessed 7/2006. URL www.matsim.org
15. M. Balmer, M. Rieser, A. Vogel, K.W. Axhausen, and K. Nagel. Generating day plans based on origin-destination matrices. A comparison between VISUM and MATSIM based on Kanton Zurich data. In *Proceedings of Swiss Transport Research Conference (STRC)*, Monte Verita, CH, 2005. See www.strc.ch.
16. Lohse, D., H. Teichert and B. Dugge. *Programmsystem VISEVA, Verkehr in Städten und Regionen – Erzeugung, Verteilung, Aufteilung*. developed at TU Dresden, distributed by PTV AG Karlsruhe, 2004.
17. Lohse, D. *Verkehrsplanung*, volume 2 of *Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung*. Verlag für Bauwesen, Berlin, 1997.
18. Vrtic, M., D. Lohse, P. Fröhlich, C. Schiller, N. Schüssler, H. Teichert and K.W. Axhausen. A simultaneous two-dimensionally constraint disaggregate trip generation, distribution and mode choice model: Theory and application for the Swiss national model. Paper presented at the ERSA conference, Amsterdam, July 2005.
19. BVG www page, accessed 7/2006. URL www.bvg.de
20. Senatsverwaltung für Stadtentwicklung and Industrie- und Handelskammer Berlin. *Stadtentwicklungsplan Zentren 2020. Standorte für Einzelhandel und Freizeit*. Berlin, 2005.
21. Gesamtverband des Berliner Einzelhandels. *Zentrenatlas*. Berlin, 2004.
22. Kutter, E. *Integrierte Berechnung städtischen Personenverkehrs – Dokumentation der Entwicklung eines Verkehrsberechnungsmodells für die Verkehrsentwicklungsplanung Berlin (West)*. Berlin, 1984.
23. Kutter, E., H-J. Mikota. *Weiterentwicklung des Personenverkehrsmodells Berlin auf der Basis der Verkehrsentstehungsmatrix 1986 (BVG)*. Untersuchung im Auftrag des Senators für Arbeit, Verkehr und Betriebe. Berlin, 1990.