AUTOMATIC CALIBRATION OF MICROSCOPIC, ACTIVITY-BASED DEMAND FOR A PUBLIC TRANSIT LINE

Manuel Moyo Oliveros (corresponding author)
Berlin Institute of Technology
Transport Systems Planning and Transport Telematics
Salzufer 17-19
10587 Berlin
Germany
Telephone: (+49)3031423308
FAX: (+49)3031426269
moyo@vsp.tu-berlin.de
http://www.vsp.tu-berlin.de

Kai Nagel
Berlin Institute of Technology
Transport Systems Planning and Transport Telematics
Salzufer 17-19
10587 Berlin
Germany
Telephone: (+49)3031423308
FAX: (+49)3031426269
nagel@vsp.tu-berlin.de
http://www.vsp.tu-berlin.de

Submission date: November 15, 2011.

Abstract

This work describes the methodology to get a more realistic transit route choice estimation, evaluated with occupancy counts comparison at bus stops in a real world scenario. First, a search of travel priorities for passengers was carried out with the selection of travel behavioral parameters whose values correction would help for the manual calibration. Moreover, the dynamic transport calibrator Cadyts was integrated on the transit simulation to help to estimate the trip generation and route choice. A bus line of the city of Berlin with real data of daily usage was taken as test scenario. The calibration experiments reduced the occupancy counts comparison error by 35% from about 50% to about 15%.

INTRODUCTION

The importance of transit systems is marked by transport science and other disciplines. It is a relevant topic on energy consumption reduction discussions, an alternative for urban environmental issues, and a point in question for urban planners. Besides, public vehicles availability is an indispensable option for low-income households(1).

The implementation of a microscopic approach for the simulation of passengers' travel behavior represents a valuable tool for route choice analysis. In this way, transit assignments models recognize more constituent elements than route choice approaches for private cars. When people make use of the public transport infrastructure to fulfill their daily activities in different locations, the individual route choice considers the adaptation to actual timetables with the minimization of some travel properties like time, distance, number of vehicle changes. Thus, a realistic microsimulation must recognize passengers' travel preferences. In order to model them, they are to be parameterized, measured and validated with real transit usage data.

MATSim (2) is an agent based transport simulation framework that is able to handle scenarios with millions of agents. An innovative methodology is the use of a calibration tool to estimate the travel demand. Cadyts (Calibration of Dynamic Traffic Simulations)(3) is an open-source calibrator originally developed for the estimation of vehicular travel demand. This work reports the integration of both tools to use passenger counts at stop facilities for a microscopic public transport demand estimation. The study started from a population sample and actual data of transit usage of bus line M44 in Neukölln district in Berlin, Germany.

RELATED WORKS

Traffic demand calibration is a prevailing topic in transport research. Chu et al.(4) proposed a traffic network-level calibration procedure for PARAMICS. Route choice diversification was achieved by costs modifications on link decreasing speed limit values, link cost factors and link tolls. Vaze(5) used a mesosopic simulation to prove the calibration improvement with automatic vehicle identification techniques using simultaneous perturbation stochastic approximation, genetic and particle filter algorithms. Zhang et al. (6) described an implementation of genetic algorithm-based calibration tools for local, global and departure-route choice parameters.

However, few works are found that deal directly with the estimation of passenger travel demand in transit simulation. A Fuzzy-Neuro approach is proposed by Yaldi et al.(7) to improve accuracy in travel demand modeling. Tamin and Sulistyorini (8) used Non-Linear-Least Squares to calibrate parameters to estimate O-D matrix. Li et al. (9) estimated also OD matrix based route choice through passenger counts. Parveen et al. (10) presented the calibration of the aggregate transit-assignment model used in EMME/2, which is based on the minimization of travel time with five parameters: boarding time, wait-time factor, wait-time weight, auxiliary time weight and boarding-time weight. In order to match on-board counts, it uses a genetic algorithm where each chromosome represents a set of parameter values generated randomly. A more recent work by Wahba and Shalaby (11) presented the calibration of the transit scenario of Toronto with MILATRAS. The learning model is based on mental models for every passenger where travel experiences are updated and evaluated in order to adjust waiting and in-vehicle time. The calibration defines nine parameters related to trip purpose and transit vehicle type. It is done with the integration of the genetic algorithm GenoTrans engine.

BACKGROUND

This section describes the two main elements that were combined for this article: (1) the public transit router (and simulation) for MATSim, and (2) the demand calibration tool Cadyts.

Transit simulation

The key processes of MATSim for the transit simulation(12, 13) are briefly described in the following.

Required **input data** are: transportation demand data, description of street network, and timetable information of the considered scenario, including description of vehicles and stops. MATSim considers the normal daily itinerary of every person, represented by a plan data structure. In it, the sequence of daily activities like being home, at work, education, shopping or leisure is described with their start and end times and geographic locations. The trips that the persons accomplish between the planned activities are represented by legs described with travel time, route and transport mode.

The **transit schedule** is a data structure containing the public transport system information. A *transit line* is understood here as an organized public transport supply normally labeled with an alphanumeric or color identifier that covers a defined area with a set of transit routes. A *transit route* denotes a distinctive fixed trip between an initial stop and a final stop. As a rule, two transit routes of the same transit line travel the same path but in opposite directions, but also more transit routes with slightly different paths may be included in a transit line. A *stop facility* or just *"stop"* is a defined location where transit vehicles make a time-planned pause to pick up or drop off passengers.

A **transit network** for the router is created, with nodes representing the stops, and directed *transit links* between them, according to each transit route information. *Transfer links* are added to allow transfers between stops that are next to each other. The transit network represents a logical layer used for routing passengers. It is merged with another directed graph of the street physical layer to create a multimodal network that is used for the complete transit and traffic flow simulation.

The **transit user route calculation** is described in Sec. 4.3 of (12) and (very similarly) in Sec. 7.4 of (13). The transit router uses an adaptation of Dijkstra's algorithm (14) which allows multiple starting and ending nodes. The compound least cost path considers walk time, in-vehicle travel time, travel distance and vehicle transfers. Unfortunately, the default values of the parameters are not given in those texts, but they can be extracted from the MATSim software repository (matsim.svn.sf.net). They are:

- Marginal Utility of Travel Time Transit (MUTTT): $-6/3600s$.
- Marginal Utility of Travel Time Walk (MUTTW): $-6/3600s$.
- Marginal Utility of Travel Distance Transit (MUTDT): $0/m$.
  It is the product of Marginal Utility of Money default value 1.0 and Monetary Distance Cost Rate default value 0.0.
- Utility of Line Switch (ULS): $60s*$ MUTTT. Note that this typically is negative, since MUTTT typically is negative.

Utilities are taken as dimensionless quantities; $-6/3600s$, say, means "minus six utils per hour". In the approach, time spent waiting at stations is included into the travel time transit and thus weighted with the same factor as travelling; this is a property of the underlying routing algorithm that may need revision in the future. For the present study, it is to be expected that the marginal cost of waiting is absorbed into the work costs for access and transfer.

Other route search configurable options are:

- Initial search distance: radius length in meters for stop facilities search, having starting or destination points as center. Its default value is 1,000 (1,093.61 yd).
- Extended search radius: an extra distance in meters to be added in case that an insufficient number of stops are found only with the initial distance. Its default value is 200 (218.72 yd).
- transfer connection distance: radius distance in meters to search potential transferring stops in a circle around a change point. Its default value is 100 (109.36 yd).

The transit router finds a transit connection at a given time between two locations including the necessary walks to, between and from stops, and description of transit legs between starting and final stops. Once the routing process is done, the details of the found connection are described there: new transit activities and legs are added to the original plan to depict actions like walking to stop facilities, boarding, transferring and alighting.

The **traffic flow simulation** executes all plans simultaneously by moving agents in the physical network. Streets are represented in the queue model by links with free speed travel time, flow capacity and storage capacity as constraints. Vehicles are differentiated as private or public, so that bus driver agents are incorporated in the simulation to execute their own plans that consist in driving public vehicles according to route schedules. Public vehicles stop at the fixed stop facilities located at the end of the links where passengers wait for them in a waiting queue. Passengers can get on the arriving vehicle, if it is the one selected by their route choice, and if it has not reached its maximum capacity. The microsimulation approach handles every agent that can be tracked and thus, the flow of private and transit agents can be measured.

Plans are executed through many **iterations**, so that re-planning strategies can be subsequently applied. A strategy defines the mechanism whereby some properties of plans are modified in an iteration so that they may be later evaluated to assess the plan performance at a day. This may include route choice or time departure, for example. In this process, agents may learn having the experience of the results of previous iterations and applying modification to the new ones. The strategy also includes the plan choice approach.

For **scoring**, a utility based approach is followed to evaluate plans' performance after their execution with a quantitative score. In it, the utility of a plan *V(i)* is calculated as the sum of positive utilities (in a logarithmic form) achieved by carrying out activities, plus the sum of negative utilities of traveling between activities locations.

$$V(i) = \sum_{act \in i} \beta_{perf} \cdot t^*_{act} \cdot \ln t_{perf,act} + \sum_{leg \in i} V_{tr,leg} \tag{1}$$

where:
β_{perf} is the activity marginal utility at its typical duration.
t^*_{act} is the activity typical duration.
$t_{perf,act}$ the activity duration in the simulation.

$V_{tr,leg}$ is the utility (typically negative) of a leg (see below).
Sometimes, there are also penalties for schedule delay, such as arriving late or departing (too) early.

Agents can have more than one plan. **Plan choice** is done as follows:

- If the agent has at least one non-scored (i.e. never executed) plan, a random choice between the non-scored plans is performed.
- If an agent has all plans scored, then a score based selection in a multinomial logit (15) form is performed.

Behavioral calibration of transport simulations

Cadyts (16) is a transport demand estimation tool that can be integrated to any stochastic and iterative assignment microsimulator. It calibrates the behavior in a Bayesian setting from real counts data. A previous interaction between MATSim and Cadyts to estimate private car traffic in the Zurich scenario is described in (17).

Cadyts is not a stand-alone framework but a calibration tool for dynamic traffic assignment simulators originally developed for the estimation of vehicular travel demand. Detailed theoretical description can be found in (18). Only a summary description of the calibration steps is introduced here in order to help to illustrate its integration with MATSim transit simulation:

- **Initialization:** At the beginning of the run, the calibrator method *addMeasurement* collects all available traffic counts.

  addMeasurement(L l, int start_s, int end_s, double value, double stddev, Measurement.TYPE type)

  The method is called for each location *l* with available traffic counts during the time bin specified from *start_s* to *end_s*, with counts classified in accordance to the data structure *Measurement.TYPE* whose instance *type* may denote either the average flow rate or the total traffic count *value* during the time interval.

  Thereby, *L* is a template variable, defined by the object instantiation

  MATSimUtilityModificationCalibrator<L> calibrator = new MATSimUtilityModificationCalibrator<L>(...); .

  There are no restrictions to the type of *L*, which means that measurements can be attached to arbitrary objects. They just need to be the same as the objects that are traversed by the plans (see next).

- **Plan Choice:** A MATSim-Cadyts-adapter would create instances of an interface called *Plan<L>* for Cadyts' own internal representations of travel demand. The calibration of a simulation with utility-based demand works by computing utility adjustments for every agent at each iteration. After calculating the utility modification with the method

  calcLinearPlanEffect(Plan<L> plan),

the agent selects a plan based on the modified utility of plans. How the choice model uses this information is left to the choice model, but in many cases, for every plan the utility modification is added to the uncorrected utility, and the resulting modified utilities are used for the choice model (also see below).

The selected plan is presented to the calibration with the method

registerChoice(Plan<L> plan) .

Cadyts runs a regression model for every featured location l and time bin, where the number of agents that intend to cross that location is the explanatory variable and the actual flow across the same location is the dependent variable. The slope of the resulting regression line provides sensitivity information to the calibration. The *registerChoice(Plan<L> plan)* method is necessary to identify the explanatory variable.

- **Update:** At the end of each iteration, the calibrator reads the output network loading situation through a container *SimResults* which takes in a set of hourly resulting traffic volumes of a location <*L*>.

afterNetworkLoading(SimResults<L> simResults)

The Cadyts posterior choice model is outlined in Section 3.1 of (17) and specially at 3.2 where it is presented the formulation of its distribution embedded with the MATSim demand simulation multinomial logit model, assuming moderate congestion with independently normal distributed traffic counts. The core equation for the purposes here is

$$P(i|y) \sim exp\left(V(i) + \sum_{ak\in i} \frac{y_a(k) - q_a(k)}{\sigma_a^2(k)}\right) \qquad (2)$$

where:
y is the actual traffic count on link a.
$P(i|y)$ is the posterior plan choice distribution given y.
$V_n(i)$ is the score of a plan i as formulated in Eq.(1).
$y_a(k)$ is the actual traffic count at link a during time k.
$q_a(k)$ is the simulated traffic count at link a during simulated time k.
$\sigma_a^2(k)$ is the variance of the traffic count.

If the variance $\sigma_a^2(k)$ is not specified separately, it is assumed to be proportional to the measured value in order to be consistent with the assumption of Poisson distributed measurements. Measured values that are registered without an explicit variance are multiplied by a configurable positive factor called `varianceScale` and the product is assumed as variance. The `varianceScale` default value 1.0 was used for this paper. In order to avoid numerical problems, Cadyts bounds the effective values of $\sigma_a(k)$ from below. The configurable `minStddev` value defines the smallest allowed standard deviation for measurements. After some experimentation, it was set to 8 for this paper. This effectively means that relative errors at count values below $8^2 = 64$ are under-weighted accordingly.

When the whole process is converged, *calcLinearPlanEffect* effectively returns the result of the $\sum_{ak \in i} \ldots$ computation of Eq. (2). That is, at the choice step Cadyts affects the agent plan choice in this way: the plan choice is under normal circumstances a function of the plan performance reflected on its score, but in the case of the calibration it is also a function of the real data counts reproduction. That is, the utility of a plan gets a higher value with the utility correction if it helps to improve the reproduce the real counts. And on the contrary, a plan receives a lower score value if it deteriorates the counts reproduction during simulation. In order to make sure that utility corrections at more representative count stations have a more significative effect than at unimportant stations, the error contributions of every individual counting station are scaled with $1/\sigma_a^2(k)$.

The combined effect of the σ_a also is that it balances between the prior utility $V(i)$ and the Cadyts utility correction: larger σ_a mean less trust in the measurements, and thus a larger weight to the prior.

SCENARIO

Supply

The transport system of the city of Berlin was chosen as scenario for calibration tests. 906 million per year, 2.4 million per day is the number of passengers that the local public transport firm BVG (Berliner Verkehrsbetriebe) reports in 2011(19). The demand is satisfied by 10 metro (U-Bahn) lines, 149 bus lines, 22 tramway lines in the east districts, and 6 ferry lines. Moreover, 15 suburban metro lines (S-Bahn, not operated by BVG) cover also the transit demand along the most important stop facilities in the city and its surroundings.

The *multimodal* network created for the transit simulation consists altogether of 37,591 links, where 25,704 of them represent the main avenues and 11,887 all transit links created out from the BVG timetable information. The number of nodes in the transit network representing any kind of stop facility is 4,791.

The passengers with activities in the area around the bus line M44 have also other alternatives. Fig. 1 shows the line M44 path with nearby lines: the bus line 181 that overlaps with M44 in two stations, and the bus line 744(also named 736) that overlaps in 4 stations. The passengers located east from line M44 may be also attracted to use subway line U7. This subway runs in a parallel path to M44 with transversal distances around 900 meters. Moreover, passengers travelling from the area around bus stop Britzer Damm/Tempelhofer Weg in direction northwest (for example to S-Bahn stations Südkreuz) or Schöneberg might use the line M44 and transfer to line S42 at Hermannstraße, or travel directly with line M46.

Demand

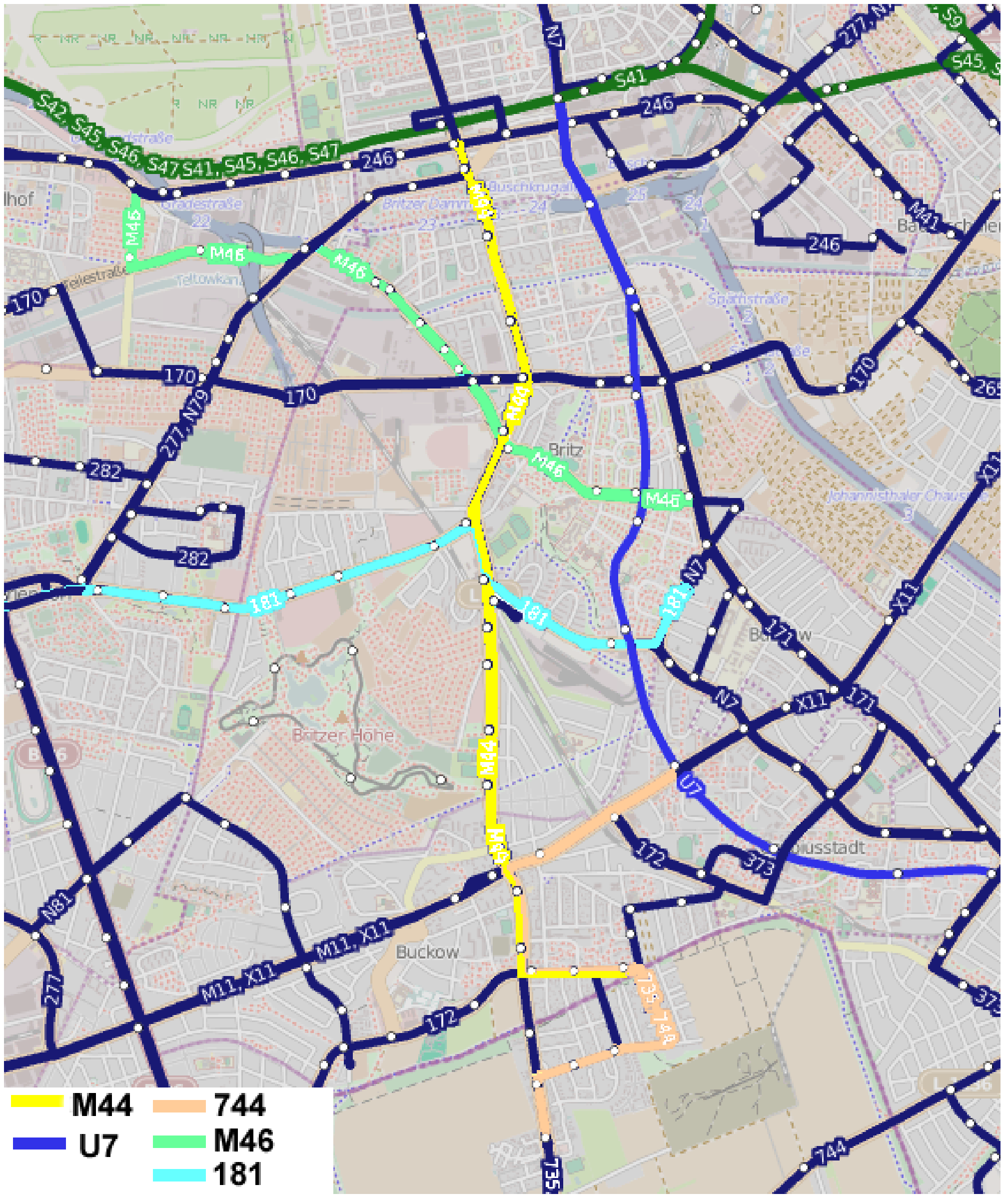
In order to set the demand, a synthetic population sample of agents having activities inside the bus lines M44/344 cover zone was constructed as follows (Neumann,A., unpublished data):

- The starting point is a BVG household survey from 1998 also used in other studies (20, 21, 22). After cleaning, this survey contains the trip diaries from 57,688 persons in the Berlin-Brandenburg area and represents nearly 2% of the population.

- All persons are routed according to their selected mode. For transit mode, the aforementioned parameter default values were used.

- Passengers not having an activity in the area served by the M44/344 bus area are removed. Since in this analysis, entering/leaving a bus counts as activity, all passengers entering or leaving a bus in the M44/344 area are maintained. In contrast, passengers just traveling through the area, either by car or by other means of transport (such as longer bus lines) are removed. For the purposes of the present study, it is assumed that this is acceptable since no mode choice was considered. It is improbable (albeit not impossible) that some long-distance passenger might be available for switching to the M44/344 line; such a passenger would have been removed by the filter.

- In order to get a suitable synthetic population base of large demand, the remaining population is expanded to a “5x” sample, which means that each agent representing a passenger is copied 4 times, and these copies get their activity locations randomly relocated in a 1-kilometer radius circle around the original sites. After the expansion, the new synthetic travelers are routed in the same way according to their mode, and the sample is filtered again to discard new agents outside the area surrounding the line M44/344 path.



(Source: www.openmap.lt in Oct. 2011)

FIGURE 1 Bus line M44 and other nearby lines

- Finally, during execution of routing processes, all agents with car mode are discarded because they do not belong to the scope of this study, which ends up with a final 5x ($= 10\%$) population sample of 36,119 agents.

Counts

Passenger occupancy counts data for 18 stops covered by the bus line M44 come from a survey by BVG in September 2009 and they reflect the usage of the line in a normal weekday.

The bus line M44 contains four transit routes. Two transit routes cover, in opposite directions, the complete set of 18 stops. The other two cover only 13 stops. For simulation and calibration, occupancy counts for all 18 stops in all directions and on all routes were considered. The results presented next aggregate, at each of the 18 stops, the data into hourly bins.

The occupancy is always counted *after* the stop, i.e. when the doors are finally closed for departure. Since the last stop of a bus line implies that all passengers must get off, no occupancy count are produced there and therefore it is not shown in the figures.

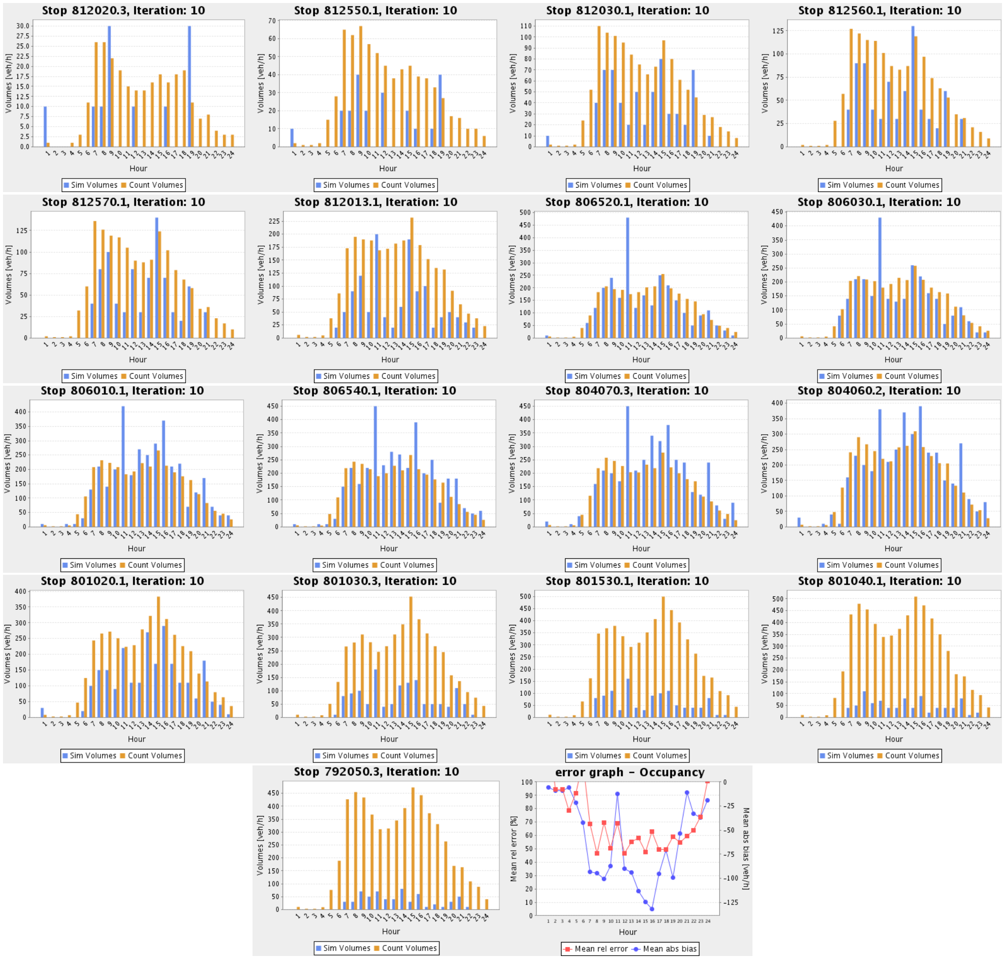


FIGURE 2 Per stop counts data-simulation comparison plots and general error graph before calibration (5x expanded population)

RESULTS

Adaptations to transit router

An adjustment was made to the transfer link creation in that transfer links between nodes of the same line were completely forbidden. A second modification was the progressive search of near stop facilities around the starting and destination points. These changes had overall beneficial effects on the number of found connections and the average travel time, and were kept for that reason. More information is given in the Appendix of (23).

Before calibration

A first test was carried out using the MATSim router travel parameters *default* values (MUTTW $= -6/3600s$, MUTDT $= 0/m$, ULS $= 60s/$MUTTT). The 17 sub-figures of Fig. 2 show the comparison of real occupancy values (in yellow) and simulated values (in blue) for the main transit route stops in hourly bins. General occupancy analysis indicates the mean relative error (red line in last sub-figure) for the whole transit line that fluctuates around 50% and 70% before any calibration attempt.

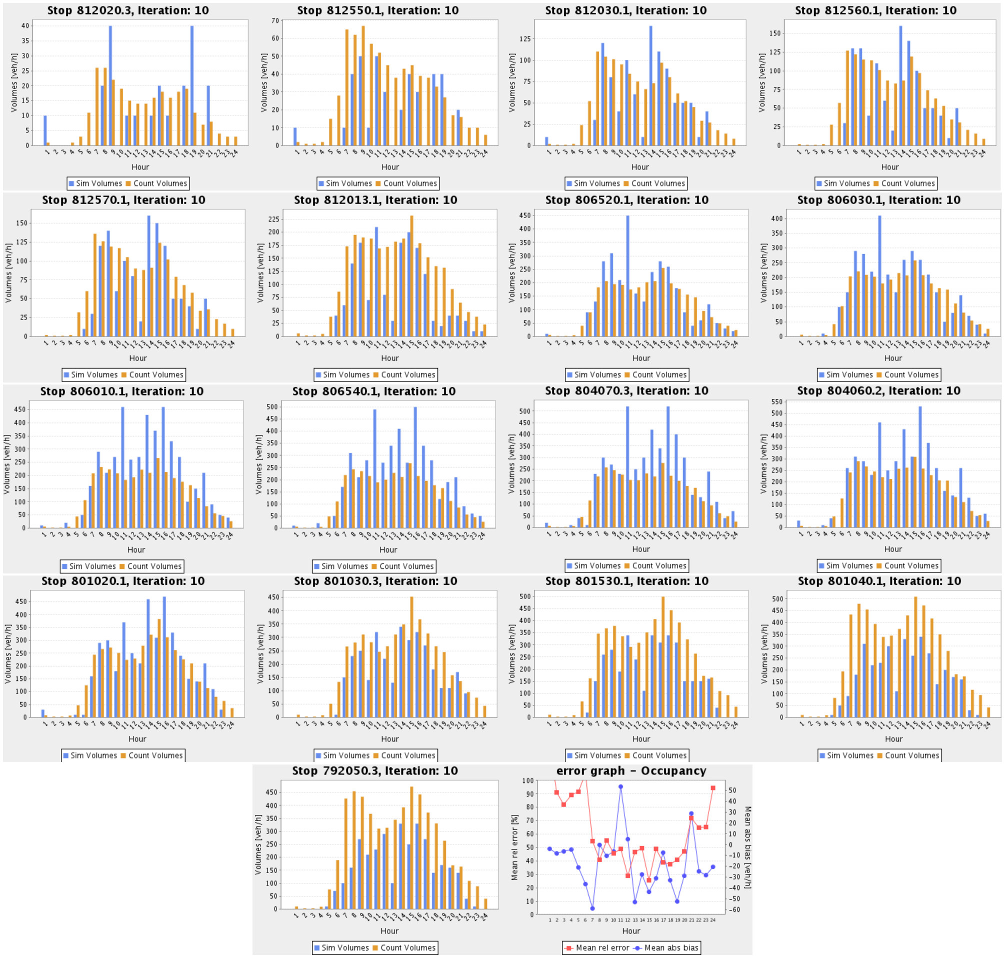


FIGURE 3 Per stop counts data-simulation comparison and general error graphs after manual calibration (5x expanded population)

Manual calibration of the utility function

A first task was to find an acceptable set of parameter values that may produce close to reality occupancy simulation values. Weight variations on 3 cost variables were tested as follows:

- walking time (MUTTW from $-1/3600s$ to $-10/3600s$ in increments of $-1/3600s$)
- transit travel distance (MUTDT from $-0/1000m$ to $-1.4/1000m$ in increments of $-0.1/1000m$) and
- utility of line switch (ULS from $0 * \text{MUTTT}$ to $1200 * \text{MUTTT}$ in increments of $60 * \text{MUTTT}$)

The Marginal Utility of Transit Travel Time (MUTTT) remained constant with its default (dis)utility value of $-6/3600s$.

An exhaustive search of combination of different parameters values was done according to the range of values for each variable. That is, 3,150 parameter combinations were obtained from the number of variations of each parameter (10*15*21 = 3150). Clearly, a strongly negative MUTTW value represents a high resistance to walk to, between or from stops. A strongly negative ULS value represent a high resistance to change vehicle, a strongly negative MUTDT represents a high resistance to choose long distance connections. In every case, initial and end values were set such that the plausibility of the routing results was already obviously impaired.

Passenger routes resulting from high resistance to walk (more strongly negative MUTTW value) and also from high resistance to transfer (more strongly negative ULS value) produced simulated values closer to actual counts data. In the case of ULS, the best output was achieved with values more strongly negative than $240*$MUTTT (which means an equivalent penalty of 4 minutes per transfer) and in MUTTW with values lower than $-6/3600s$. Fig. 3 shows an example of error comparison of simulation counts data reached just by this approach. It can be seen that the general error percentage without calibration fluctuates around 50% and 30%. The best combination of travel parameter values from this manual calibration is:

- Marginal Utility of Travel Time Walk (MUTTW): $-10/3600s$ (compared to $-6/3600s$ in the original router)
- Marginal Utility of Travel Distance Transit (MUTDT): $0.0/1000m$ (same as in the original router)
- Utility of Line Switch (ULS): $240 *$ MUTTT (compared to $60 *$ MITTT in the original router)

Parameter	matsim old	this section	Florida	Commuter	Toronto	San Francisco	Santiago
in-vehicle time [min]	0.1	0.1	0.02	0.025	2.0	0.023	0.119
walk [min]	0.1	0.17	0.045	0.047	1.0	0.029	0.240
line switch	0.1	0.4	0.045	./.	./.	./.	0.449
wait time [min]	0.1	0.4	0.045	0.046	2.733	0.044	0.111
walk/in-veh	1	1.7	2.25	1.88	0.5	1.26	2.02
switch/in-veh	1	4	2.25	./.	./.	./.	3.77
wait/in-veh	1	1	2.25	1.84	1.37	1.91	0.93

TABLE 1 Coefficient values comparison. The four top rows show the absolute values. The three bottom rows show the values for walk times, switch occurrences, and wait times divided by the value for in-vehicle times.

Compared with the original routing parameters, travelers attempted to reduce their amount of walking and the number of interchanges. This suggests that in-vehicle times and in-vehicle distances should increase. And indeed, with from the original routing parameters (Fig. 2) to the calibrated ones, the average in-vehicle travel distance for all M44 users increased from 1,804 to 2,441 meters.

As an attempt to validate these tendencies, individual transit connections requests were compared with the BVG journey planner (19). It turned out that similar connections were suggested also by the BVG site.

Table 1 compares these values with the ones found by other mode choice and transit assignment studies in different scenarios: Florida (24), the averages for a number of american cities with COMMUTER v.2(25), Toronto (26), San Francisco Bay Area(27) and Santiago(28). The three bottom rows divide the values for walk time, switch occurrences, and wait time by the in-vehicle time, resulting in more meaningful numbers. All values seem to be in a similar range. Given the importance of the penalty for line switching in Berlin, a comparison with those models that also include a penalty for line switching seems most meaningful. Out of those, the Santiago model comes out strikingly similar to ours, while the Florida model has relatively more penalty on waiting and relatively less penalty on line switches. Overall, our result seems to be in line with others.

Coupling microsimulation and calibration

As the next step, an integration code was written in Java to work as bridge between Cadyts and MATSim transit simulation.

The Cadyts generic network link type was originally meant to represent network links with traffic count stations. For the estimation of passenger travel behavior, it was adapted to represent transit stop facilities with available passenger occupancy counts instead. Thus, variables y and q of Eq. (2) acquire these meanings:

y is the actual occupancy count at transit facilities after unloading and loading, and
q is the simulated occupancy count after unloading and loading
$V(i)$ is set to zero for the purposes of this paper, in order to score plans just by their consistency with real counts.

Automatic calibration with Cadyts

In order to apply Cadyts, the route choice generation was separated from the simulation process. That is, routes were pre-calculated in independent routing calls before the MATSim iterations, with the calibrator enabled, started. The calibrator would thus select between the pre-computed plans, but not add new plans to the choice set.

The criteria to create the different plans were: variety of routes, and the search of connections with minimal number of interchanges and minimal walk distances

In the end, three different transit plans per synthetic traveler were generated. The parameter values used for the three different public transit plans are:

- Combination 1: MUTTW= -6/3600, MUTDT= -0.0/1000, ULS= 1200*MUTTT, i.e. strong transfer penalty.
- Combination 2: MUTTW= -10/3600, MUTDT= -0.0/1000, ULS= 240*MUTTT, i.e. strong walk penalty.
- Combination 3: MUTTW= -8/3600, MUTDT= -0.5/1000, ULS= 720*MUTTT, i.e. moderate walk and transfer penalties.

In addition, in order to obtain an elastic demand, the following was done:

- All synthetic travelers (of the “5x” sample) were duplicated.
- All synthetic travelers got an extra plan in which they stayed at home.

The result is that the calibrator will not only affect the transit routing, but also the overall level of demand, which can be increased or decreased by decreasing or increasing the fraction of “stay-home” plans.

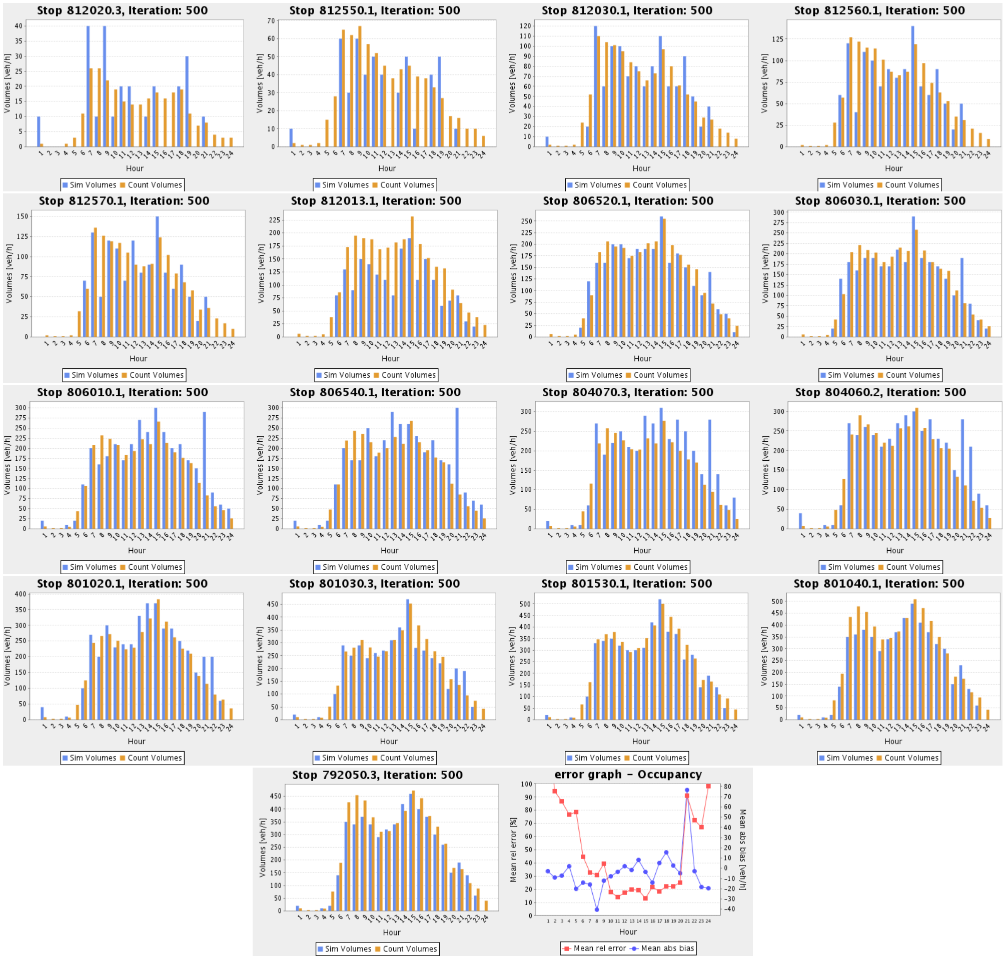


FIGURE 4 Per stop counts data-simulation comparison plots and general error graph after automatic calibration (5x expanded population)

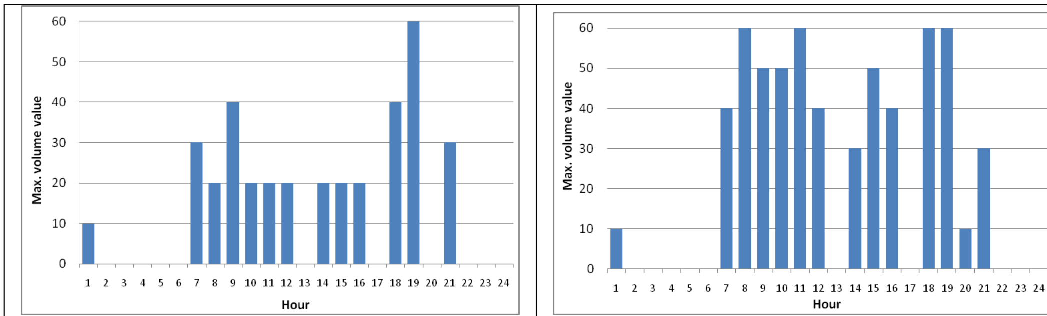
Now, using the Cadyts utility modification as the basis for plan selection, a calibrations run was done loading agents with those 3 different public transit plans plus the "stay-home" plan, and calibrating the period from 06:00 to 20:00 hours.

The comparison of Cadyts-enabled simulation results with real counts data are shown in Fig. 4. The general error was reduced by around 20% in comparison with the manual calibration. Simulated and actual counts reached a suitable comparison at most stops where morning and afternoon peak hours can be identified in both counts types.

One should note, though, that the manual calibration and the Cadyts calibration attempt different things:

- The manual calibration attempted to find *one* set of behavioral parameters that would lead to realistic occupancies.
- The automatic calibration picks *one out of four* different passenger route plans (one of them being the stay-home plan) in the attempt to generate realistic occupancies.

It is clear that the second approach has more degrees of freedom and thus achieves a better fit.



Stuthirtenweg	Ringslebenstr./Mollnerweg

FIGURE 5 Maximum volumes for the first two bus stops for any given hour during "manual calibration" (5x expanded population)

Investigation of missing demand segments

The first two stops of the presented transit route showed a lower consistency with the real data than the rest of the stops, even after the calibration runs. It is quite clear that a synthetic population that is based on a simple "5x" expansion of a 2% sample may have gaps that cannot be filled by the adjustment process. The problem can already be visually taken from "Stop 812020.3" in Fig. 4 where one notices that the simulation can provide passengers only in increments of "10", corresponding to the 10% sample where every passenger also stands for 9 others. That is, for some hours of the day there may simply be no demand available that can be shifted to match those counts.

To investigate, occupancy counts were reviewed along the complete set of 3,150 parameter combinations to find which ones may supply higher volumes or any volumes at all. However, it was not found any combination that could be able to provide any volume for hours 2, 5, 6, 13, 17, 20, 22, 23, 24 neither at the first "Stop 812020.3" not at the second "Stop 812550.1" for hours 2, 3, 4, 5, 6, 13, 17, 22, 23, 24, as it can be seen in their maximum volumes graph in Fig. 5. It means in general that the original population sample is not enough at those stops to reproduce satisfactorily the occupancy counts.

As a way to settle the insufficient demand at the first stops, a second version of the population with agents allocated at different hours was tested. It was also originated from the same 2% basis sample and prepared in the same way, but for the expansion, 9 copies instead of 4 were created. Moreover, time mutation was applied on the activities of those new agents with a random range of 7,200 seconds. To compare its effects, the same procedures of data preparation, routing and calibration were done with the new synthetic population version. The results are shown in Fig. 6.

It can be seen that with the time mutation of agents' activities, the calibration is able to improve the reproduction of occupancy volumes even at the bus stops with less demand. The general error also is placed between 10% and 20% for most of the calibrated hours.

Investigation of residuals

Previous figures with counts comparisons help to recognize the individual contribution of each stop to the general error. However, it is tangible that some stops are more representative in terms of the

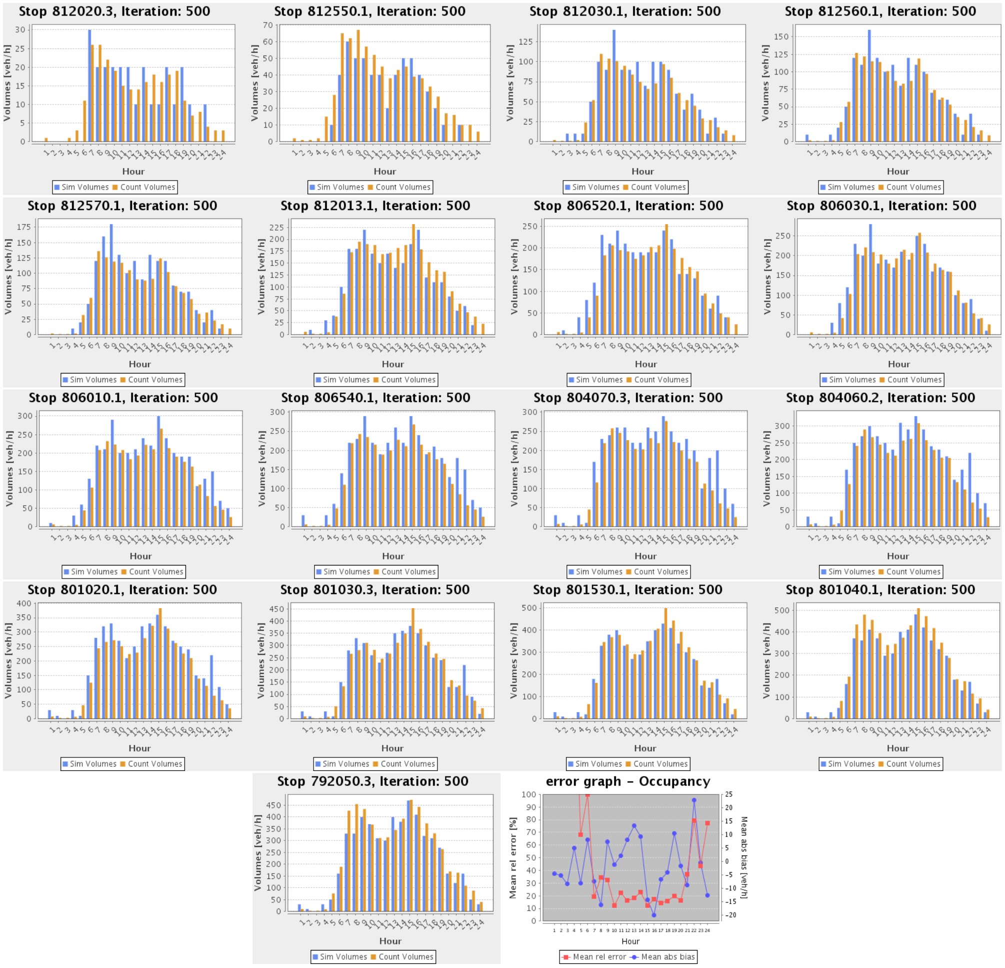


FIGURE 6 stop comparison and general error after calibration of 10x expanded synthetic population (with time mutation)

error reduction than others due to their occupancy volumes magnitude. Specially in the examined route, last stops are presented with higher values than those of the first stops.

On these grounds, another way of analysis was done representing the error proportion for stop. It is based on the mean weighted square error calculation that indicates the average quadratic deviation between real and simulated traffic counts presented in Sec. 4 of (17), but in this case representing all error contributions for stop and hour. Thus, omitting the average calculation, and taking the same variable meanings as in Eq. 2, the weighted square error WSE of a count location a at a given time bin k is estimated like this:

$$WSE_a(k) = \frac{(y_a(k) - q_a(k))^2}{2\sigma_a^2(k)} \tag{3}$$

The weighted error graphs of the time-mutated synthetic population calibration is presented in Fig. 7. The series of graphs shows the bigger impact that middle and last stops have on the error correction in the calibration. That is, it becomes quite comprehensible that Cadyts does not attempt

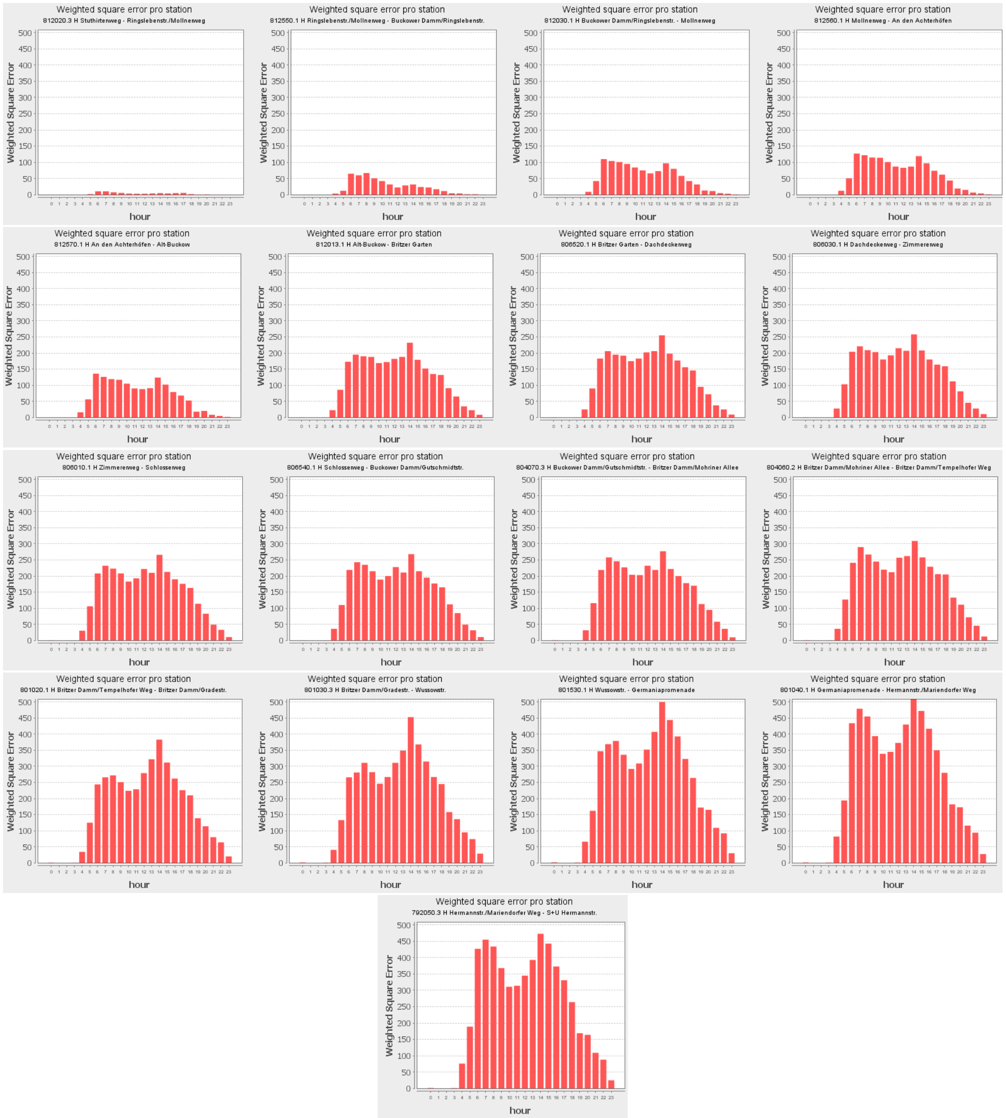


FIGURE 7 Weighted squared error for bus stops for calibration of 10x expanded synthetic population

harder to correct the remaining errors at the first two stops: Those errors are relevant in relative terms, but not in absolute terms.

DISCUSSION

As stated earlier, it is no wonder that the calibrator is able to achieve a better result than the manual calibration, since it does the equivalent of modifying each individual traveler's behavioral parameters in order to reproduce the real-world counts. Future work will have to show how this can be made behaviorally more plausible, e.g. by including taste variations into the synthetic travelers and then calibrating the taste coefficients.

In the meantime, it should be pointed out that also the current method has its applications. For example, it is planned to look at the interaction between schedule stability measures and demand for a single line in much more detail. For such an investigation, it is useful to have a demand that is as close as possible to the actual counts. Clearly, for this is it possible to just use the boarding and alighting counts directly as demand (see, e.g., (29)). Yet, for many investigations it is desirable to have that demand embedded in the remainder of the system in order to investigate interactions such as, say, demand shocks from subway lines. For such investigations, the presented approach seems very appropriate.

SUMMARY AND FUTURE WORKS

The integration of MATSim simulation and Cadyts for transit demand estimation was presented here. The objective of the experiments on the Berlin scenario was to reproduce the actual counts data inside the simulation, first with the search of suitable travel parameter combinations during the manual calibration and then, its use for the automatic calibration runs. Route diversity was achieved with high walk resistance, high transfer resistance and medium values with special focus on stops with problematic counts reproduction. At the end of all calibration experiments, general error was reduced by 35% from about 50% to about 15%.

The calibration effects were tested only on one bus line. A natural following step is the inclusion of more transit lines (including subway and tramway). Some studies suggest that passengers show some preference to rail based vehicles, and it could be included inside the route choice and probed with calibration.

A more appropriate method of calibration should include the scoring function working together with Cadyts as replanning strategy. Modifying also its parameters to find best count matches might help to reach a more complete description of passengers travel behavior. Up to now the route choice has been separated as an initial and independent step from simulation, a future task is its integration in the same replanning process with a route diversity dynamic creation.

ACKNOWLEDGMENTS

This work is founded in part by the National Council on Science and Technology of Mexico (CONACyT) and the German Academic Exchange Service (DAAD). The computing calculations are possible with the help of Prof. H. Schwandt's group from the Institute of Mathematics at TU Berlin. Gunnar Flötteröd provided very useful suggestions and support for this work. Andreas Neumann provided the initial demand. The Berlin public transit company BVG provided the passenger occupancy counts data.

References

[1] Criden, M., *The Stranded Poor: Recognizing the Importance of Public Transportation for Low-Income Households*. National Association for State Community Services Programs. CSBG Issue Brief. Washington D.C., 2008.

[2] MATSim, *Multi-Agent Transportation Simulation*. http://www.matsim.org, accessed 2011.

[3] Cadyts, *Calibration of dynamic traffic simulations*. http://transp-or2.ep.ch/cadyts, accessed 2011.

[4] Chu, L., H. Liu, J.-S. Oh, and W. Recker, A Calibration Procedure for Microscopic Traffic Simulation. In *Proceedings of the Transportation Research Board Annual Meeting (TRB)*, Washington D.C., 2004.

[5] Vaze, V., *Calibration of Dynamic Traffic Assignment Models with Point-to-Point Traffic Surveillance*. Master's thesis, Massachusetts Institute of Technology, 2007.

[6] Zhang, M., J. Ma, and H. Dong, *Developing Calibration Tools for Microscopic Traffic Simulation Final Report Part II: Calibration Framework and Calibration of Local/Global Driving Behavior and Departure/Route Choice Model Parameters*. California PATH Research Report. University of California Davis, 2008.

[7] Yaldi, G., M. A. P. Taylor, and W. L. Yue, Developing a Fuzzy-Neuro Travel Demand Model (Trip Distribution and Mode Choice). 30th Conference of Australian Institutes of Transport Research, 2008.

[8] Tamin, O. Z. and R. Sulistyorini, Public Transport Demand Estimation by Calibrating the Combined Trip Distribution-Mode Choice (TDMC) Model From Passenger Counts. *World Academy of Science, Engineering and Technology*, Vol. 54, 2009.

[9] Li, Y. and M. J. Cassidy, A generalized and efficient algorithm for estimating transit route ODs from passenger counts. In *Transportation Research Part B 41*, 2007, pp. 114–125.

[10] Parveen, M., A. Shalaby, and M. Wahba, G-EMME/2: Automatic Calibration Tool of the EMME/2 Transit Assignment Using Genetic Algorithms. *Journal of Transportation Enginnering*, Vol. 133, No. 10, 2007, pp. 549–555.

[11] Wahba, M. and A. Shalaby, Large-scale application of MILATRAS: case study of the Toronto transit network. *Transportation*, Vol. 38, 2011, pp. 889–908, 10.1007/s11116-011-9358-5.

[12] Rieser, M. and K. Nagel, Combined agent-based simulation of private car traffic and transit. In *Proceedings of The 12th Conference of the International Association for Travel Behaviour Research (IATBR)*, Jaipur, India, 2009, also VSP WP 09-11; see www.vsp.tu-berlin.de/publications.

[13] Rieser, M., *Adding transit to an agent-based transportation simulation concepts and implementation*. Ph.D. thesis, TU Berlin, 2010, also VSP WP 10-05, see www.vsp.tu-berlin.de/publications.

[14] Dijkstra, E., A note on two problems in connexion with graphs. *Numerische Mathematik*, Vol. 1, 1959, pp. 269–271.

[15] Ben-Akiva, M. and S. R. Lerman, *Discrete choice analysis*. The MIT Press, Cambridge, MA, 1985.

[16] Flötteröd, G., Cadyts a free calibration tool for dynamic traffic simulations. In *Proceedings of the Swiss Transport Research Conference (STRC)*, Monte Verita, CH, 2009, see www.strc.ch.

[17] Flötteröd, G., Y. Chen, and K. Nagel, Behavioral Calibration and Analysis of a Large-Scale Travel Microsimulation. *Networks and Spatial Economics*, in press, pp. 1–22.

[18] Flötteröd, G., M. Bierlaire, and K. Nagel, Bayesian demand calibration for dynamic traffic simulations. *Transportation Science*, in press, also VSP WP 08-19; see www.vsp.tu-berlin.de/publications.

[19] BVG, *Berliner Verkehrsbetriebe.-Berlin transportation company site with journey planner page*. http://www.bvg.de (in German), accessed 2011.

[20] Kutter, E., H.-J. Mikota, J. Rümenapp, and I. Steinmeyer, *Untersuchung auf der Basis der Haushaltsbefragung 1998 (Berlin und Umland) zur Aktualisierung des Modells "Pers Verk Berlin / RPlan", sowie speziell der Entwicklung der Verhaltensparameter '86–'98 im Westteil Berlins, der Validierung bisheriger Hypothesen zum Verhalten im Ostteil, der Bestimmung von Verhaltensparametern für das Umland*. Draft of the final report, Sponsored by the "Senatsverwaltung für Stadtentwicklung Berlin", Berlin/Hamburg, 2002.

[21] Scheiner, J., Daily mobility in Berlin: On 'inner unity' and the explanation of travel behaviour. *European Journal of Transport and Infrastructure Research*, Vol. 5, 2005, pp. 159–186.

[22] Rümenapp, J. and I. Steinmeyer, *Activity-based demand generation: Anwendung des Berliner Personenverkehrsmodells zur Erzeugung von Aktivitätenketten als Input für Multi-Agenten-Simulationen*. VSP Working Paper 06-09, TU Berlin, Transport Systems Planning and Transport Telematics, 2006, see www.vsp.tu-berlin.de/publications.

[23] Moyo O., M. and K. Nagel, *Automatic calibration of microscopic, activity-based demand for a public transit line*. VSP Working Paper 11-13, TU Berlin, Transport Systems Planning and Transport Telematics, 2011, see www.vsp.tu-berlin.de/publications.

[24] Abdel-Aty, M. and H. Abdel Wahab, *Calibration of Nested-Logit Mode-Choice Models for Florida*. Ph.D. thesis, Dept. of Civil & Environmental Engineering. University of Central Florida. USA., 2001.

[25] Transportation, R. P. D. O. of Transportation, and A. Quality., *COMMUTER v2.0 Model Coefficients*. U.S. Environmental Protection Agency, 2005.

[26] Miller, E., Generalized Time Transit Assignment in a Multi-Modal/Service Transit Network. Dept. of Civil Engineering, University of Toronto, Presentation to the 20th International EMME Users Conference. Montreal., 2006.

[27] Cambridge Systematics, I. and M. B. R. . Consulting., *Mode Choice Models*. Final working paper, San Francisco Bay Area Water Transportation Authority (WTA), San Francisco. CA., 2002.

[28] Raveau, S., J. C. Munoz, and L. de Grange, A topological route choice model for metro. *Transportation Research Part A: Policy and Practice*, Vol. 45, No. 2, 2011, pp. 138 – 147.

[29] Neumann, A. and K. Nagel, *Avoiding bus bunching phenomena from spreading: A dynamic approach using a multi-agent simulation framework*. VSP Working Paper 10-08, TU Berlin, Transport Systems Planning and Transport Telematics, 2010, see www.vsp.tu-berlin.de/publications.

[30] Stern, R., Passenger Transfer System Review. *Transit Cooperative Research Program (TCRP) Synthesis 19. Transportation Research Board, Washington*, 1996.

[31] O'Sullivan, S. and J. Morrall, Walking Distances to and from Light-Rail Transit Stations. *Transportation Research Record*, Vol. 1538, 1996.

APPENDIX

TRANSIT ROUTER ADAPTATION

Two modifications were implemented with the goal of increasing the number of found connections and add other realistic elements to the route search.

Simplified transfer link creation

In the search of a transit connection for an agent, a change of transit vehicle is possible thanks to the virtual transfer links created in the transit network as described earlier.

In the first network creation step, nodes stand for the stops along the transit route, and transfer links are meant to join near stops that belong to different transit lines. In the original implementation, transfer links are created between every pair of nodes within the transfer distance that

- *either* belong to different transit lines,
- *or* belong to different stop facilities.

The adaption consisted in dropping the additional condition of linking nodes of different stop facilities, thus joining nodes with the only condition that they should belong to different transit lines.

The goal was to avoid the creation of unnecessary transfer links between consecutive stops of the same transit route that are inside the transfer distance. The elimination of that requirement had the effect of reducing of transfer links in the transit network of the test scenario described in this paper from 106 059 to 83 838 (almost -21%).

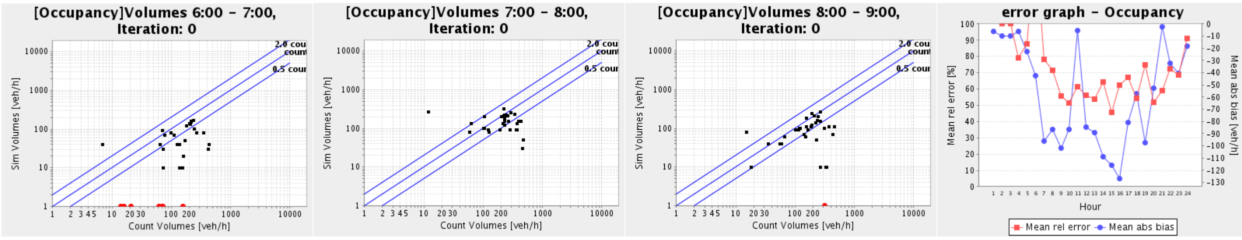
Moreover, in order to have a more realistic implementation of transfers, the original distance of 100 meters for the search of near stop facilities was tripled. This is in accordance with studies that suggest transfer walk distances around 300 meters or even longer (30, 31). This radial distance expansion increased back from 83 838 to 143 154 (almost $+71\%$) the number of transfer links.

Stop search with progressive radius extension.

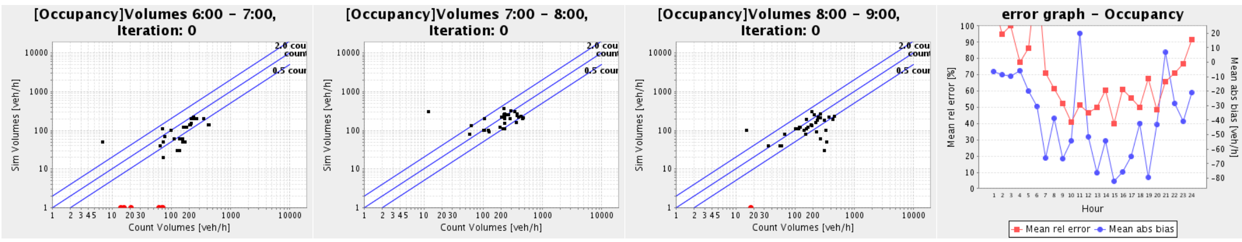
When a transit connection search is requested, stop facilities are to be found around origin and destination points. Originally the router searches for stop facilities inside an initial given radius, but in case that the number of found stops is less than two, the radius is enlarged to the distance of the nearest stop plus an extension radius distance of 200 meters. A modification was done to guarantee a configurable minimum number of stop facilities to start the transit connection, independently of their distance to the activity location. It starts with a predefined initial radius but this is enlarged progressively by the extension radius distance so many times as needed until at least the minimum number of stop facilities are found. For all runs in this paper, that minimum number was set to “2”. Also, instead of the standard 1000 meters distance for initial search, only 600 meters were used; this reduces the problem size in dense urban environment.

Results

A simple comparison was done with the Berlin scenario before before applying any calibration attempt, using the transit router default values presented before. Adapting the progressive stop search and the simple transfer link creation produced altogether more connections (almost $+12\%$)



(a) before adaptations



(b) after adaptations

FIGURE 8 Passenger occupancy results at early hours before (a) and after (b) router adaptations

and reduced the travel time, but it increased the walk distance and time. The sum of these values is shown in Tab. 2.

	before adaptations	**after adaptations**
Number of connections	86 739	97 202
Travel time in seconds	4.49E+12	3.53E+12
Number of transfers	153 644	143 022
Walk time in seconds	1.50E+10	1.72E+11
Walk distance in meters	72443	83 198

TABLE 2 Results of transit router adaptations

In the same way, regarding the the occupancy counts, a distinct improvement was achieved in the occupancy counts, as it can be seen on a comparison of line M44 occupancy at early morning hours as it can be seen in Fig. A.3.