Calibration of a Heterogeneous Traffic Scenario in an Agent-Based Simulation Framework

Amit Agarwal*, Dominik Ziemke, Kai Nagel

Technische Universität Berlin, Berlin, 10587, Germany

* Corresponding author: agarwal@vsp.tu-berlin.de

Abstract. By raising the issue of data requirements for the purpose of modal development, validation and application, this study proposes an approach to calibrate travel demand in heterogeneous traffic condition using minimal empirical data. For this, a real-world scenario of Patna, India is chosen. For the calibration, a Bayesian framework based calibration technique (Cadyts: Calibration of dynamic traffic assignment) is used. Commonly available, mode-specific, hourly-classified traffic counts are used to generate full day plans of agents and their initially unknown activity locations. While the proposed approach implements location choice implicitly, the approach can be applied to a variety of other problems. Further, the effect of household income is included in the utility function to filter out inconsistencies in the plans, which originate from the survey data.

1 Introduction

In a transportation system, a wide variety of data (*e.g.*, network data, socio-economic data) is required for the purpose of model development, validation and application. The aim of such models is to simulate and analyse travel demand and test the policies, which can help transport planners to understand the decision making process of individual travellers. A model should be causal, flexible, transferable, efficient, and sensitive to policy objectives [1]. Most travel demand models minimally require information about the trip origin, trip destination, and trip mode. The information about origin and destination (OD) can come in different forms and at different level of aggregation, *e.g.*, as an OD matrix, as daily plans, *etc.* The traditional way to estimate the OD matrix relies on roadsideor household surveys, which are, however, error-prone and likely to be biased [2,3]. As an alternative, there are several approaches to estimate the OD matrix using traffic counts (*e.g.*, see [4–6]).
Given the origin-destination information of an area, static traffic assignment (STA) provides the traffic flow on each highway for every time bin. Dynamic traffic assignment (DTA) is a generalisation of STA, which provides time-dependent traffic flow on each highway segment [7]. From the development perspective, DTA models can be classified in two categories, analytical and simulation-based models. The former are often preferred for large urban agglomeration and for microscopic traffic flow characteristics [7,8]. In the context of the application of such models to large urban transportation networks, at least two problems become apparent: a) microscopic modelling is computationally expensive and b) data requirements are high. Mainly based on the underlying traffic flow model DTA models can be classified as physical-queue models [9,10] and non-physical-queue models [7,11]. One such physical-queue model [12,13] is embedded in the activity-based, multi-agent transport simulation framework MATSim [14]. Due to its simplicity, it is able to handle large urban transportation networks [15] and still resembles to Newell's simplified kinematic wave model [16,17]. The first problem mentioned above regarding the resource-intensive models can be managed by such fast traffic flow models.
Traditionally, in order to gather the required data, different types of data collection techniques are used, which are either manual or automatic. Such approaches include mid-block traffic count surveys, spot-speed surveys, origin-destination surveys, household surveys *etc.* [18]. The use of mid-block traffic counts survey is popular in India for various purposes. However, this information is not sufficient to simulate the travel demand for an urban scenario in order to understand the behaviour of individual travellers. The complexity rises if traffic streams are populated with different vehicle types, which is very common in most developing economies. In this direction, this study proposes an approach to calibrate travel demand in heterogeneous traffic conditions using minimal empirical data.
In contrast to traditional data collection techniques, several studies apply alternative approaches to derive and validate travel demand. Detailed surveys to collect the data (*e.g.*, household surveys), which require origin and destination information, trip modes, trip purposes, start times, end times *etc.*, are often associated with high non-responses and misreporting rates [19,20]. Traffic data collection based on manual or automated traffic counts is usually easier to manage. With the recent technological advances, new approaches are presented , which make use of GPS (Geographical Positioning System) technology in traditional travel

surveys, which is likely to improve the quality and robustness of the data [20–22]. Similarly, in the last couple of years, several other studies proposed different approaches to collect data using CDR (call detail records) from smartphones [23, 24]. A simulation-based approach to construct all-day trip chains using mobile phone data is proposed by Zilske and Nagel [25], which reduces spatio-temporal uncertainties.
In this direction, this study proposes an approach to construct trip diaries in heterogeneous traffic conditions using hourly classified mid-block traffic counts. For this, a real-world scenario of Patna, India, is considered. The data for the scenario is taken from the Comprehensive Mobility Plan (CMP) for Patna [26]. A few inconsistencies in the survey data are observed, which are likely to occur in other scenarios as well. Some of these inconsistencies are repaired in the scenario. The remainder of the paper is structured as follows. Sec. 2 illustrates the calibration process, Sec. 3 exhibits the travel demand for the scenario and construction of an income-dependent utility function. Calibration results are presented and discussed in Sec. 4. The study is concluded in the Sec. 5.

2 Calibration procedure

In this study, the multi-agent based transport simulation framework MATSim is used (see Sec. 2.1), which is able to handle large-scale scenarios because of its fast network loading algorithm (see [27] for details) and ability to handle mixed traffic conditions [28, 29]. Together with this, the calibrator Cadyts ("Calibration of dynamic traffic assignment"; see Sec. 2.2) is used. It has been used previously to adjust traffic demand of car traffic [30, 31] and to calibrate the demand for public transit [32]. It has also been applied to solve the problem of location choice [31], which was applied in the creation of an open scenario for Berlin [33]. In these approaches, however, Cadyts was used for homogeneous traffic conditions, while the present study extends the approach for heterogeneous traffic conditions.

2.1 Travel Simulator: MATSim

In this study, the MATSim transport simulation framework [14] is used for all simulation experiments.The minimal inputs for a simulation run are the physical boundary conditions (i.e., the road network), daily plans of individual travellers and scenario-specific parameters. MATSim is composed of an iterative cycle in which every individual traveller is considered as an agent. The cycle consists of following three parts: (1) Plans execution: In this step, the plans of all individual travellers are executed simultaneously on the network using a mobility simulation. In this study, a time-step-based queue simulation approach [12, 34] is used. This can also simulate heterogeneous traffic conditions [27–29]. (2) Plans evaluation: The executed plans are evaluated using a utility (scoring) function. In this study, the default 'Charypar-Nagel' scoring function [35] is used and further modified to include the effect of household income (see Sec. 3.2.2). (3) Re-planning: A new plan is generated for some agents by modifying an existing plan's attribute (departure time, route, mode) using so-called innovative strategies. The old plans are kept in the agents' memories and can be selected by so-called non-innovative strategies later on. The new plan is executed in the next iteration. The above steps are repeated in an iterative process. Innovation is used until a certain iteration. Finally, a number of additional iterations are run only with non-innovative strategies enables (i.e., plan selection), which finally results in stabilized simulation outputs.

2.2 Calibrator: Cadyts

In an activity-based simulation framework, traffic counts are insufficient to generate whole day plans of individual travellers. To address this issue, a calibrator called "Cadyts" is used [36, 37], which is based within a Bayesian framework. Together with simulation framework, this is integrated to the utility function such that probability of selecting a plan i from the j plans is given by Eq. (1). In this, y_{lt} and q_{lt} are the measurement and simulation values for spatial location l and time bin t. σ_{lt}^2 is variance of measurement. V_i is the utility of the plan and ω is weight parameter for correction ΔV_{lt} (Eq. (2)).

$$P(i|y) = \frac{exp(V_i + \omega \cdot \sum_{lt} \Delta V_{lt})}{\sum_j exp(V_i + \omega \cdot \sum_{lt} \Delta V_{lt})} \quad (1)$$

$$\Delta V_{lt} = \frac{y_{lt} - q_{lt}}{\sigma_{lt}^2} \quad (2)$$

In this study, hourly classified traffic counts are available, which are used to generate whole day plan for the travellers. From the Eqs. (1) and (2), one can observe that a plan, in which, an agent traverses a link whose simulated counts are underestimated, is more likely to be chosen. For heterogeneous traffic conditions, Eq. (2) is modified as shown in Eq. (3); where m is the mode for which measured traffic counts at link l, time bin t are available.

Revisiting Eqs. (1) and (3), it can be observed that, if the choice set of an agent contains plans with different modes, the correction is likely to fix the modal share as well. In this study, Cadyts is used to generate full day plans of agents and its initially unknown activity locations.The choices for the different activities are provided by creating multiple plans corresponding to each plausible activity location (see Fig. 1). The approach can be applied to a variety of problems.

$$\Delta V_{ltm} = \frac{y_{ltm} - q_{ltm}}{\sigma_{ltm}^2} \qquad (3)$$

3 Real-world case study: Patna, India

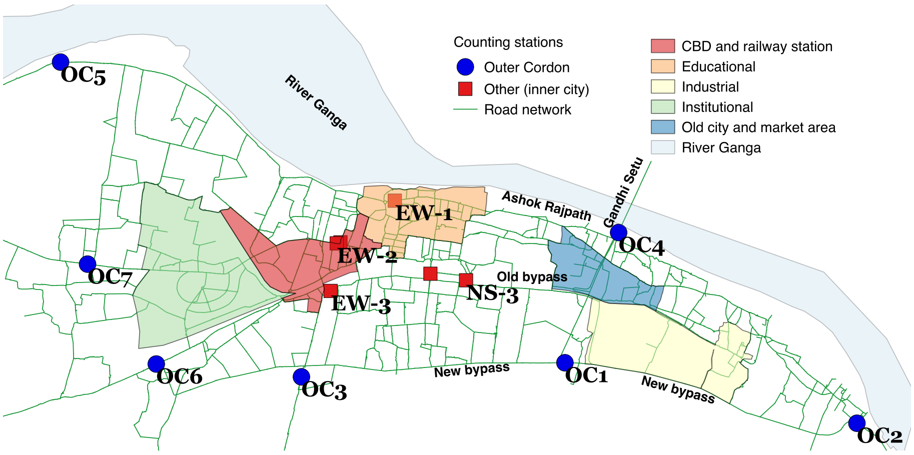


Figure 1: Patna road network, survey locations and land-use pattern.

This section exhibits the set-up for a real-world scenario of Patna, India. The road network, survey locations, and the land-use patterns of Patna are shown in Fig. 1 [27].

3.1 Travel Demand

The travel demand of the region can be categorized in two groups, urban and external travel demand.

Urban travel demand Urban travel demand is generated directly from a trip diary survey [26]. Parts of the data in the household survey were unavailable; for such cases the required data were imputed randomly based on other available data in the Patna CMP. This results in 13,278 records. Every such record is translated into one agent with one plan. In absence of other data, two trips for each plan are synthetically generated. In order to get significant number of plans for commuters and through traffic in various categories, 10% sample is used. Therefore, urban plans are cloned and modified by randomizing activity locations (origin, destination) within the zone and departure times.

External travel demand External travel demand is categorized in through traffic and commuters. The former is the traffic which passes through Patna and consists in at most one trip per day, whereasthe latter consists in agents who commute between Patna and nearbyareas. To include the congestion effect of external traffic in the activity-based transport simulation framework, the whole day plans of the external traffic are required. These are generated as follows.

1. The Patna CMP provides hourly classified counts for 7 outer cordon stations (see Fig. 1) in both directions and directional split factors. The directional split provides the share of commuters and through traffic from each counting station.

2. For through traffic, an OD matrix is given, which provides the origins and destinations.In absence of separate factors, factors from the matrix are used for all modes (bicycle, car, motorbike and truck)

and in all time bins; this provides the mode and departure times for the trips. Consequently, a 10% sample is created from the counts such that each plan has one trip only.

3. For commuters, exact locations of the trip destinations are initially unknown. They are calibrated in this study based on the given traffic counts in a similar way as done by Ziemke et al. [31] for car traffic. A few potential activity locations are identified based on the land-use pattern (see Fig. 1). A random point inside any of these probable activity location areas is taken as the trip destination. Thus, for every agent, 5 plans are generated corresponding to each plausible destination and added to the choice set of the agent. From Eq. (1) recall that, a plan is favoured if the agent travels via one of the counting station, which is underestimated in the simulation. In other words, within the simulation framework, location choice is available to the agents.

3.2 Scenario preparation

The calibration of the scenario is performed for the following reasons. a) Trip destinations (activity locations) of the commuters are unknown. b) A few trip diaries do not have mode and income information which is randomly assigned based on the modal distribution from Patna CMP [26]. c) A few trip diaries are inconsistent (see Fig. 3a). For instance, a) persons from very low income group (8-11 *USDct*) make trips by car, b) person from high income group make a 10 *km* long trips using bicycle or walk modes. Such situations are very unlikely and assumed as reporting errors. d) The mode-specific utility parameters are taken from other sources. Alternative (or mode) specific constants (ASCs) for all modes are unknown.

3.2.1 Travel modes

In this study, car, motorbike, bicycle, and truck modes are physically simulated on network (so called main modes or congested modes), whereas walk and public transit (PT) are teleported between origin and destination (so-called uncongested modes). The main difference between the two is that main modes consume flow and storage capacities on the link and thus affect the route choice decision making process of the individual travellers. Tab. 1 provides the maximum speeds for all modes and PCU for congested modes. In the traffic mix, shares of bicycle and motorbike modes are high, therefore, the PCU of bicycle and motorbike is assumed as 0.15 [38].

Table 1: Modal attributes for Patna scenario.

	bicycle	**car**	**motorbike**	**truck**	**PT**	**walk**
Speed (km/h)	15	60	60	30	20	5
PCU	0.15	1	0.15	3	–	–

3.2.2 Utility function

Utility parameters: To evaluate a plan, a scoring function is used which requires explicit values for utility parameters. In order to determine the utility parameters, the value of time and vehicle operating costs is taken from [39] and converted to USD (1 USD ≈ 66.6 INR on 8 June 2016) for a common interpretation (see Tab. 2). The average trip cost per *km* for PT is taken from [40] and shown in Eq. (4). The value are on the lower side, however, seems appropriate due to significant share of low cost “tuk-tuks” in Patna.

Table 2: Values of time and vehicle operating costs [39].

travel mode	**vehicle operating costs ($USDct/km$)**	**value of time ($USDct/h$)**
car	3.75	93.84
motorbike	1.55	48.05
PT	–	59.31

$$\text{PT trip costs} = \begin{cases} 0.045, & \text{if } d \leq 4\ km \\ 0.045 + (d-4)\cdot 0.0047, & \text{if } d > 4\ km \end{cases} \tag{4}$$

Dependency on household income: In general, the value of time is the opportunity cost of time an individual traveller spends on the trip; this is highly dependent on the income level of individual. In order to incorporate the high income differentiation across different modes, the perception of income is added to behavioural decision making process of individual by modifying the utility function.

1. **Utility of travelling:** The utility of travelling is given by:

$$S_{trav,mode} = C_{mode} + \beta_{trav,mode} \cdot t_{trav} + (\beta_{d,mode} + \beta_m \cdot \gamma_{d,mode}) \cdot d_{trav} \tag{5}$$

   where C_{mode} is ASC for mode *mode*, $\beta_{trav,mode}$ is marginal utility of travelling (normally negative or zero), $\beta_{d,mode}$ is marginal utility of distance (normally negative or zero), β_m is marginal utility of money and $\gamma_{d,mode}$ is mode-specific monetary distance rate (normally negative or zero). t_{trav} and d_{trav} is travel time and travel distance between two activity locations.

2. **Marginal utility of travelling:**

   a) As is common (*e.g.*, [41]), it is assumed that the income-dependent marginal utility of money ($\beta_{m,j}$) is indirectly proportional to the income: $\beta_{m,j} = \bar{y}/y_j$ where $\bar{y}$ is the median income for all individuals, and y_j is the income of individual j.

   b) It is assumed that the (dis)utility of travelling by car ($\widetilde{\beta}_{trav,car}$) is the same for every individual although car is predominantly used by persons with a higher income. Therefore, the value of travel time saving for the car mode (VTTS*car*) is given by:

$$\text{VTTS}_{car} \stackrel{!}{=} \frac{-\widetilde{\beta}_{trav,car}}{\beta_{m,highIncome}} \qquad \beta_{m,highIncome} = \frac{\bar{y}}{y_{highIncome}}$$

   where $y_{highIncome}$ is the median income for car users. Thus, the marginal utility of travelling by car will become (VTTS values come from Tab. 2):

$$\widetilde{\beta}_{trav,car} = \text{VTTS}_{car} \cdot \frac{\bar{y}}{y_{highIncome}} = 0.94 \cdot \frac{4000}{20000} = 0.19 \, \frac{\text{util}}{h},$$

   c) Similarly, for motorbike and PT, the marginal utility of travelling will be:

$$\widetilde{\beta}_{trav,mb} = 0.48 \cdot \frac{4000}{6250} = 0.31 \, \frac{\text{util}}{h} \qquad \widetilde{\beta}_{trav,PT} = 0.59 \cdot \frac{4000}{4000} = 0.59 \, \frac{\text{util}}{h}$$

   d) In absence of the values of time for bicycle and walk modes, (dis)utility (or disagreeability) of being (stuck) in traffic for bicycle and walk mode is assumed same as motorbike; i.e., $\widetilde{\beta}_{trav,bicycle} = \widetilde{\beta}_{trav,walk} = \widetilde{\beta}_{trav,mb} = 0.31$ util$/h$

   These values express that car is the most favourable of all available modes, and PT the least favourable. The fact that the VTTS of car in Tab. 2 comes out as the one with the *highest* willingness-to-pay to shorten its duration is explained by the higher income of car users, and not as a general inconvenience of car, which seems to be more plausible.

3. **Utility of performing an activity:** Considering the marginal utility of time as a resource, a unit reduction in travel time (Δt) would not only save the direct (dis)utility of travel $\beta_{trav} \cdot \Delta t$ but also increase the score by the utility of time as a resource, which approximately is $\beta_{dur} \cdot \Delta t$ [42]. The latter is the opportunity cost of time gained by performing the activities for the saved time (Δt). This results in $\widetilde{\beta}_{trav,mode} = \beta_{dur} - \beta_{trav,mode}$ where the sign convention is such that the parameter β_{dur} is typically positive. Following [42], the explicit value of marginal utility of performing (or marginal utility of activity duration) an activity (β_{dur}) is taken as the lowest of marginal utility of travelling for different modes ($\beta_{dur} = \widetilde{\beta}_{trav,car} = 0.19$ util$/h$), and the corresponding direct marginal utility, $\beta_{trav,car}$ is set to zero. All other direct marginal utilities of travelling are set relative to this value, i.e., $\beta_{trav,mode} = 0.19$ util$/h - \widetilde{\beta}_{trav,mode}$ The resulting mode-specific marginal utilities of travelling for MATSim scoring function are shown in Tab. 3.

Further, the ASCs for different modes are calibrated to capture the influence of variables not explicitly included in the scoring function. Along with this, to include the physical effort in bicycle and walk mode, the marginal utilities of distance for bicycle and walk, $\beta_{d,bicycle}$ and $\beta_{d,walk}$, are also calibrated.

In absence of any relevant data utility parameters of bicycle, car, and motorbike from urban and external traffic are assumed to be the same. For trucks, a different behavioural model is required, which is out of the scope of this study. However, for the scenario completion and to include the congestion effects from commercial vehicles, trucks are also included in the simulation with default utility parameters.

3.2.3 Simulation setup

Table 3: Utility parameters converted to MATSim format.

travel mode	car	motorbike	PT	bicycle	walk
monetary distance rate (γ_d)	$-3.7\cdot 10^{-5}$	$-1.6\cdot 10^{-5}$	Eq. (4)	–	–
marginal utility of travelling (β_{trav}) [util/h]	-0.0	-0.12	-0.40	-0.12	-0.12
marginal utility of performing (β_{dur}) [util/h]			0.19		

The modal splits of the urban travellers from reference study and initial plans are shown in Tab. 5. In order to replicate this modal split, mode choice is allowed for urban travellers and the ASCs are calibrated. The calibration is performed over 200 iterations together with Cadyts in order to generate the synthetic plans for the external demand (see Sec. 2.2) and find destinations for commuters. For the calibration process, the maximum limit of plans in the choice set of an agent is set to 10. After calibrating with Cadyts, only the best plans for each agent and in consequence only the destinations best matching the traffic counts are kept. The simulation is then continued for another 1000 iterations to stabilize the urban and external demand in absence of Cadyts.
Different so-called innovative modules are used for different sub-populations (urban and external). a) Urban: In a given iteration, 15% of the urban travellers are allowed to change their route, 10% are allowed to change mode and 5% are allowed to mutate the departure time of the activity. The mutation of the departure time of the activity is performed randomly between -2 to $+2$ h. The time mutation is turned off after Cadyts calibration, i.e., the departure times of the urban travellers are then fixed. b) External: In a given iteration, 15% of the agents from external traffic are allowed to change routes until innovation is turned off. After 200 iterations, the origin-destination pairs of the external demand are fixed. Innovation is used until 80% of iteration (i.e., initially for 1-160 iterations and then 201-1000 iterations). The remaining agents until 80% of the iterations and all agents afterwards chose a plan from their generated choice sets.

4 Calibration results

4.1 Calibrated utility parameters

Table 4: Calibrated utility parameters.

parameter	bicycle	car	motorbike	PT	walk
ASC (util)	0.0	-0.6	-0.58	-0.545	0.0
$\beta_{d,mode(q)}$ (util/m)	-0.00011	–	–	–	-0.00012

The (manually) calibrated ASCs for all modes and marginal utility of distance for bicycle and walk modes are shown in Tab. 4. The ASCs for bicycle and walk modes are estimated to zero, which can be interpreted as no initial impedance. Car/motorbike and PT often have some initial overhead either in terms of getting the car out of the garage or in terms of walking to a In this scenario, walking to PT stop is marginally less burdensome as getting the car/motorbike out of the garage/parking location. As a consequence of mode choice, the share of walk mode increases (see Tab. 5), which can be controlled either by a negative ASC or by having marginal utility of distance for walk mode ($\beta_{d,walk}$). The former has less significance for the walk mode and therefore the latter is chosen. In contrast to bicycle, the walk mode is teleported and thus the utility for a person with walk mode is not affected by congestion. The marginal utility of distance for the walk mode

($\beta_{d,walk} = -1.2 \cdot 10^{-4}$ $util/m$) is estimated marginally higher than the marginal utility of distance for the bicycle mode ($\beta_{d,walk} = -1.1 \cdot 10^{-4}$ util$/m$). This means, for walking 1 km, an agent will loose 0.12 util. At a speed of 5 km/h, it will take 12 min which could be used for performing an activity. Thus, the agent will also loose 0.024 util ($= \beta_{trav,walk(q)} \cdot 0.2\ h$) for travelling and 0.038 util ($= \beta_{dur} \cdot 0.2\ h$) opportunity cost of time which could be used for performing an activity.

4.2 Modal split

A comparison of the modal splits at different stages is shown in Tab. 5. It can be observed that the modal share for the walk mode is significantly different in the reference study and in the initial plans. The aim of the calibration is to replicate the modal shares from the reference study. Clearly, the modal split after calibration (column "it.1200" in Tab. 5) has close resemblance with the reference study.

Table 5: Modal splits for urban demand.

mode	reference study [26]	initial urban plans from travel diaries; it.0	after calibration it.1200
bicycle	33%	29.0%	32.3%
car	2%	4.0%	2.7%
motorbike	14%	20.3%	14.7%
PT	22%	26.6%	21.7%
walk	29%	20.1%	28.6%

4.3 Traffic counts

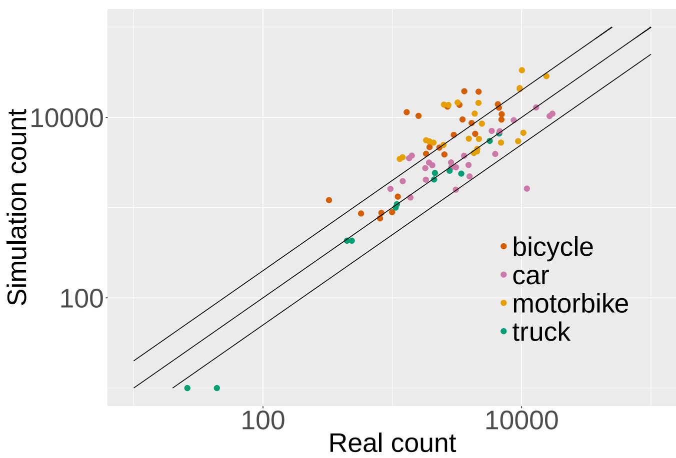


Figure 2: Comparison of 24 h simulation and real traffic counts.

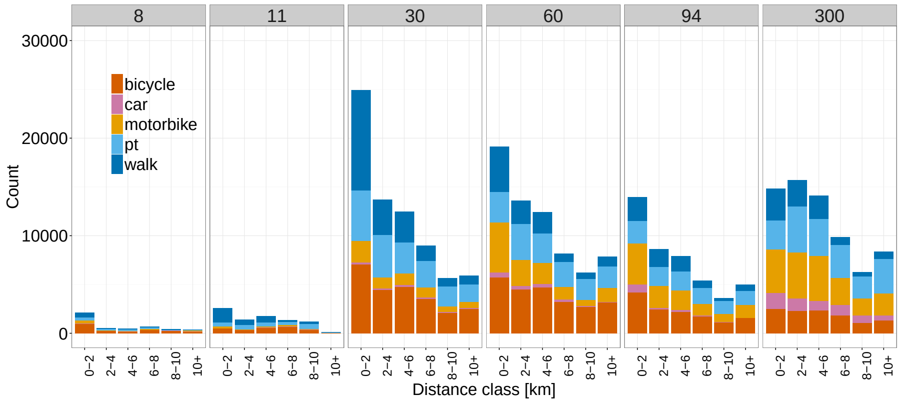
Fig. 2 shows the comparison of average weekday real counts and average weekday simulation counts after 1200 iterations. In the first step, Cadyts pushes agents on the routes by adding a correction factor (Eq. (2)) to the scoring function such that the simulation counts match the measured counts. Afterwards, in absence of the Cadyts correction factor, the simulation counts become higher than the real counts (see Fig. 2), however, the calibration results after 1200 iterations provide a good fit for modal split and synthetic plans for external traffic.

4.4 Income distance distribution

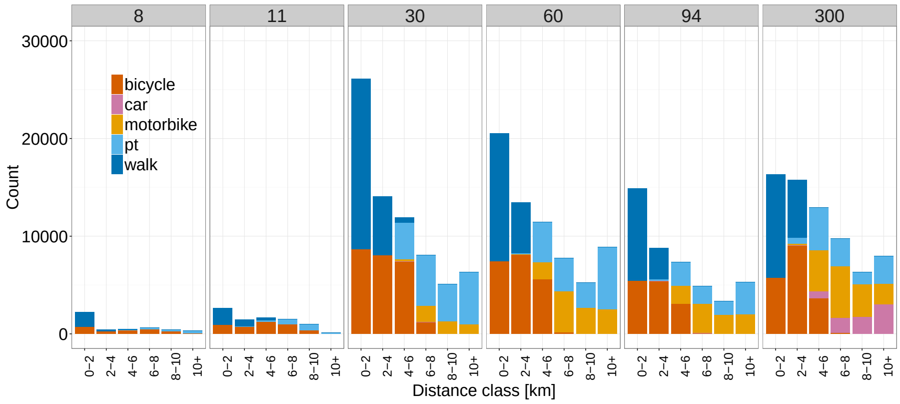
In order to understand the impact of the income-dependent scoring function for different modes, the income-distance distribution is plotted in Fig. 3. The income attributes are taken from the initial trip diaries and trip distances are the direct distances between origin and destination activities. The following observations are made:

a) After the calibration, the car is restricted to high income groups. In contrast to the initial plans, in the base case, the car is used for the longer distances.

b) PT is used mainly for longer distances ($> 4km$), whereas bicycle and walk modes are used for



(a) it.0 (initial plans)



(b) it.1200 (calibrated plans)

Figure 3: Income-dependent distance distributions for initial plans and calibrated plans. The x- and y-axes depict the distance classes (in km) and number of trips respectively. The average income (in USD) is shown at the top of each frame.

relatively shorter distances ($< 6km$). A few longer bicycle trips can also be observed for households with a very low income.

c) To replicate the modal share from the reference study, the scenario is calibrated such that the share of walk trips is about 8% higher after the calibration (see Tab. 5). A higher share of walk trips (relatively shorter distance i.e., $< 4km$) can be noticed in the Fig. 3b. Additionally, the scoring function forces the impractical longer ($> 8km$) walk trips to more plausible modes. A similar effect is also observed for the longer bicycle trips from higher income groups.

Overall one can observe that several irregularities from the travel diaries are fixed in the calibrated plans which is further suitable for policy testing.

5 Conclusions

This study addresses the difficulties in the model development and validation due to limited availability of the data. The overall objective of the study was to estimate the alternative specific constants (ASCs) in order to replicate the modal split in the reference study. In this direction, this study extended an approach to generate full day activity plans in heterogeneous traffic conditions. To simulate travel demand, an agent-based travel simulator was used, while for calibration, a Bayesian framework based calibration technique was used. A real-world scenario of Patna was used for this purpose. Diverse income levels were included in the utility function to filter out the errors in the survey data. In this approach, location choice was implicitly implemented to identify the initially unknown destinations based on the land use pattern. The calibrated ASCs show plausible values. With the help of income-based distance distributions, it was shown that the calibrated plans are feasible plans and free from the errors originated from the survey. In future, the authors wish to replace the manual calibration with an automatic calibration process using some optimization techniques [43].

Acknowledgments

The support given by DAAD (German Academic Exchange Service) to first author for his PhD studies at Technische Universität Berlin is greatly acknowledged. This paper is built up from the content of the first author's PhD dissertation.

References

[1] T. Domencich and D. L. McFadden, *Urban travel demand: a behavioral analysis*. North-Holland Publishing Company, 1996.

[2] M. Kuwahara and E. C. Sullivan, "Estimating origin-destination matrices from roadside survey data," *Transportation Research Part B: Methodological*, vol. 21, no. 3, pp. 233–248, 1987.

[3] R. M. Groves, "Nonresponse rates and nonresponse bias in household surveys," *The Public Opinion Quarterly*, vol. 70, no. 5, pp. 646–675, 2006.

[4] M. G. H. Bell, "The estimation of an origin-destination matrix from traffic counts," *Transportation Science*, vol. 17, no. 2, pp. 198–217, 1983.

[5] E. Cascetta, D. Inaudi, and G. Marquis, "Dynamic estimators of origin-destination matrices using traffic counts," *Transportation Science*, vol. 27, no. 4, pp. 363–373, 1993.

[6] H. van Zuylen and L. Willumsen, "The most likely trip matrix estimated from traffic counts," *Transportation Research*, vol. 14B, pp. 281–293, 1980.

[7] W. Szeto and S. Wong, "Dynamic traffic assignment: model classifications and recent advances in travel choice principles," *Open Engineering*, vol. 2, no. 1, pp. 1–18, 2012.

[8] M. C. J. Bliemer, "Dynamic queuing and spillback in analytical multiclass dynamic network loading model," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2029, pp. 14–21, 2007.

[9] W. Szeto, "Enhanced lagged cell-transmission model for dynamic traffic assignment," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2085, pp. 76–85, Dec. 2008.

[10] W. Szeto, Y. Jiang, and A. Sumalee, "A cell-based model for multi-class doubly stochastic dynamic traffic assignment," *Computer-Aided Civil and Infrastructure Engineering*, vol. 26, no. 8, pp. 595–611, 2011.

[11] W. H. Lam and H.-J. Huang, "Dynamic user optimal traffic assignment model for many to one travel demand," *Transportation Research Part B: Methodological*, vol. 29, pp. 243–259, Aug. 1995.

[12] C. Gawron, "An iterative algorithm to determine the dynamic user equilibrium in a traffic simulation model," *International Journal of Modern Physics C*, vol. 9, no. 3, pp. 393–407, 1998.

[13] N. Cetin, A. Burri, and K. Nagel, "A large-scale agent-based traffic microsimulation based on queue model," in *Swiss Transport Research Conference (STRC)*, (Monte Verita, Switzerland), 2003. See http://www.strc.ch.

[14] A. Horni, K. Nagel, and K. W. Axhausen, eds., *The Multi-Agent Transport Simulation MATSim*. Ubiquity, London, 2016.

[15] M. Balmer, K. Meister, M. Rieser, K. Nagel, and K. Axhausen, "Agent-based simulation of travel demand: Structure and computational performance of MATSim-T," in *Innovations in Travel Modeling (ITM) '08*, (Portland, OR), June 2008. Also VSP WP 08-07, see http://www.vsp.tu-berlin.de/publications.

[16] A. Agarwal, G. Lämmel, and K. Nagel, "Modelling of backward travelling holes in mixed traffic conditions," in *Traffic and Granular Flow '15* (V. L. Knoop and W. Daamen, eds.), ch. 53, pp. 419–426, Delft, NL: Springer International Publishing, 1 ed., 2016.

[17] A. Agarwal, G. Lämmel, and K. Nagel, "Incorporating within link dynamics in an agent-based computationally faster and scalable queue model," *Transportmetrica A: Transport Science*, 2017.

[18] T. R. Currin, *Introduction to traffic engineering: a manual for data collection and analysis*. Cengage Learning, 2012.

[19] M. Zimowski, R. Tourangeau, R. Ghadialy, and S. Pedlow, "Nonresponse in household travel surveys," tech. rep., Federal Highway Administration, 1997.

[20] J. Wolf, *Using GPS data loggers to replace travel diaries in the collection of travel data*. PhD thesis, Georgia Institute of Technology, 2000.

[21] E.-H. Chung and A. Shalaby, "A trip reconstruction tool for GPS-based personal travel surveys," *Transportation Planning and Technology*, vol. 28, pp. 381–401, Oct. 2005.

[22] L. Shen and P. R. Stopher, "Review of GPS travel survey and GPS data-processing methods," *Transport Reviews*, vol. 34, pp. 316–334, Apr. 2014.

[23] S. Iqbal, C. Choudhury, P. Wang, and M. C. González, "Development of origin-destination matrices using mobile phone call data," *Transportation Research Part C*, vol. 40, pp. 63–74, 2014.

[24] J. Chen and M. Bierlaire, "Probabilistic multimodal map matching with rich smartphone data," *Journal of Intelligent Transportation Systems*, vol. 19, pp. 134–148, Sept. 2014.

[25] M. Zilske and K. Nagel, "A simulation-based approach for constructing all-day travel chains from mobile phone data," *Procedia Computer Science*, vol. 52, pp. 468–475, 2015.

[26] TRIPP, iTrans, and VKS, "Comprehensive mobility plan for Patna urban agglomeration area," tech. rep., Department of Urban Development. Government of Bihar, 2009.

[27] A. Agarwal, *Mitigating negative transport externalities in industrialized and industrializing countries*. PhD thesis, TU Berlin, Berlin, 2017.

[28] A. Agarwal, M. Zilske, K. Rao, and K. Nagel, "An elegant and computationally efficient approach for heterogeneous traffic modelling using agent based simulation," *Procedia Computer Science*, vol. 52, no. C, pp. 962–967, 2015.

[29] A. Agarwal and G. Lämmel, "Modeling seepage behavior of smaller vehicles in mixed traffic conditions using an agent based simulation," *Transp. in Dev. Econ.*, vol. 2, no. 2, pp. 1–12, 2016.

[30] G. Flötteröd, Y. Chen, and K. Nagel, "Behavioral calibration and analysis of a large-scale travel microsimulation," *Networks and Spatial Economics*, vol. 12, no. 4, pp. 481–502, 2011.

[31] D. Ziemke, K. Nagel, and C. Bhat, "Integrating CEMDAP and MATSim to increase the transferability of transport demand models," *Transportation Research Record*, vol. 2493, pp. 117–125, 2015.

[32] M. Moyo Oliveros and K. Nagel, "Automatic calibration of microscopic, activity-based demand for a public transit line," Annual Meeting Preprint 12-3279, Transportation Research Board, Washington, D.C., Jan. 2012. Also VSP WP 11-13, see http://www.vsp.tu-berlin.de/publications.

[33] D. Ziemke and K. Nagel, "Development of a fully synthetic and open scenario for agent-based transport simulations – the matsim open berlin scenario," VSP Working Paper 17-12, TU Berlin, Transport Systems Planning and Transport Telematics, 2017. URL http://www.vsp.tu-berlin.de/publications.

[34] P. Simon, J. Esser, and K. Nagel, "Simple queueing model applied to the city of Portland," *International Journal of Modern Physics*, vol. 10, no. 5, pp. 941–960, 1999.

[35] D. Charypar and K. Nagel, "Generating complete all-day activity plans with genetic algorithms," *Transportation*, vol. 32, no. 4, pp. 369–397, 2005.

[36] G. Flötteröd, M. Bierlaire, and K. Nagel, "Bayesian demand calibration for dynamic traffic simulations," *Transportation Science*, vol. 45, no. 4, pp. 541–561, 2011.

[37] G. Flötteröd, *Cadyts – Calibration of dynamic traffic simulations – Version 1.1.0 manual*. Transport and Mobility Laboratory, École Polytechnique Fédérale de Lausanne, Nov. 2010.

[38] S. Chandra and P. K. Sikdar, "Factors affecting pcu in mixed traffic situations on urban roads," *Road and transport research*, vol. 9, no. 3, pp. 40–50, 2000.

[39] IRC:SP:30, *Manual on economic evaluation of highway projects in India*. Indian Roads Congress, New Delhi, India, 2009.

[40] C. V. P. Kumar, D. Baus, and B. Maitra, “Modeling generalized cost of travel for rural bus users: a case study,” *Journal of Public Transportation*, vol. 7, no. 2, pp. 59–72, 2004.

[41] J. Franklin, *The distributional effects of transportation policies: The case of a bridge toll for Seattle*. PhD thesis, University of Washington, Seattle, 2006.

[42] B. Kickhöfer and K. Nagel, “Microeconomic interpretation of MATSim for benefit-cost analysis,” in *The Multi-Agent Transport Simulation MATSim* (A. Horni, K. Nagel, and K. W. Axhausen, eds.), ch. 51, Ubiquity, London, 2016.

[43] A. Agarwal, G. Flötteröd, and K. Nagel, “Calibration of behavioural parameters using optimization technique in an agent-based transport simulation,” in *accepted for 6th Symposium of the European Association for Research in Transportation*, 2017.