1. Introduction

A variety of measures are needed to reduce greenhouse gas emissions with the target of mitigating climate change. The transportation sector, in particular, is responsible for a large proportion of emissions and therefore offers significant potential for reduction [1, 22]. According to Wiedenhofer et al. [23], densely populated areas are particularly...
well suited to provide low-emission mobility, suggesting that urbanization may present an opportunity. To increase the sustainability of transport systems, especially in urban areas, it is necessary to create attractive alternatives to private cars that are tailored to the local context. Thus, decision makers need to be provided with reliable information on how the existence and design of new and sustainable mobility services, such as sharing systems, influence individual travel behavior. This can be most accurately represented with agent-based travel demand models that allow for microscopic analysis.

Especially bike sharing, as active travel mode, is popular, sustainable and facilitates multi- and inter-modal travel behavior. According to DeMaio et al. [7], there are currently 2023 active bike sharing systems with more than nine million bikes around the world. Europe and Asia account for 85% of active systems, followed by North America. However, there are different types of systems around the world [8]. The most common distinction is between station-based and free-floating systems, also known as dockless. In station-based systems, bikes can exclusively be picked-up and dropped-off at designated physical stations, whereas in free-floating systems, bikes can be found anywhere in the service area. Due to the different bike sharing systems, but also the behavioral and geographical differences, diverse and flexible modeling approaches are required to reflect reality [11]. Therefore, in this paper, we present a station-based bike sharing approach in an agent-based travel demand model and compare different strategies of how users select bike sharing stations on their trips.

The remainder of this paper is structured as follows. In Section 2 an overview of different bike sharing modeling approaches is given, before the implementation of our station-based bike sharing system is carried out in Section 3 and different strategies for how agents can choose their pick-up and drop-off stations are presented. Afterwards, we discuss the results in Section 4 and finally summarize our findings in Section 5.

2. Literature

Shui and Szeto [20] provide an extensive literature review on bike sharing planning problems. They differentiate in strategic, tactical and operational problems, ranging from optimal station location to bike relocation problems. While some modeling approaches focus solely on station-based [5, 10, 21, 24] or free-floating [16] systems, others integrate both in their models [6, 11]. Most research analyzes bike sharing systems exclusively, without considering multi-modal behavior and interactions with the remaining transport system [11]. According to Calderón and Miller [4] as well as Yang et al. [24] re-balancing bikes within the system is also a major challenge when modeling bike sharing services, which is addressed for example by [5, 10, 21]. However, different approaches of modeling the decision-making process by which agents select stations to pick-up or drop-off bikes has received little attention in current research.

Hebenstreit and Fellendorf [11] develop a bike sharing extension for the agent-based travel demand model MATSim, considering both, station-based and free-floating systems including an e-bike sub-fleet. Agents opt for the nearest stations in terms of distance to their start and end locations for bike pick-up and drop-off. Nevertheless, the study does not provide a specific recommendation regarding the optimal radius within which stations should be considered for selection. According to Lin et al. [15] and Lu et al. [16] bikes should not be further away than 500 meters. Coretti Sanchez et al. [6] choose an agent-based approach implemented in Python to model bike sharing. In their model, agents always go to the nearest station. However, in the event that all bikes at the closest station have been rented, agents redirect to an alternative station where bikes are still available. In contrast, Fernández et al. [10] allow agents to reserve bikes, mitigating the need for redirection when bikes are unavailable. But it is not specified which station is selected. In the approach of Soriguera et al. [21] all agents choose the nearest station within their catchment area, too. However, two groups of agents can be differentiated. Agents equipped with an app benefit from enhanced information and only choose stations with bikes available. Agents lacking the app can still use the service but do not receive any availability information. Also e-bike sharing and re-balancing is included in the model. Chemla et al. [5] aim to minimize the total discomfort experienced by agents. Therefore, the access and egress time on foot as well as the leg by bike are weighted with individual prices and are subsequently minimized.

Thus, it can be stated that in most approaches the nearest station is chosen by an agent. As far as the authors are aware, there is no published research that tests and compares different strategies for how agents determine which station to choose.
3. Methodology

3.1. mobiTopp

In our research, we employ and advance the agent- and activity-based travel demand model mobiTopp [14, 17, 18]. The modeling framework consists of two sequential steps. Initially, in a long-term module, a synthetic population is generated, and long-term decisions are modeled. Following this, a short-term module simulates the travel demand of the population over an entire week.

In the long-term module, person agents are generated and grouped to households. These agents and households are provided with socio-demographic attributes based on a representative national household travel survey [19]. As a result, each agent is affiliated with a workplace or educational site [12], has a household income, and is equipped with an individual set of mobility tools. These include personal bikes, a driver’s license, shared household vehicles, transit passes, and a range of memberships with shared mobility service providers, such as bike sharing.

In the subsequent short-term module, all agents are modeled throughout an entire week at a time resolution of one minute. Different discrete choice models are applied to model agents decision making process of their destination and travel mode choice. Impedance values used in these choices (e.g., travel time and travel cost) are pre-calculated on a traffic analysis zone level, no route assignment is performed inside of mobiTopp.

In the following, we use a mobiTopp instance of the city of Hamburg for our research. In total, 1.9 million agents are modeled in their mobility behavior. Although our model is capable of modeling an entire week, we show our results for a typical Monday. The bike sharing infrastructure of the model mirrors StadtRAD, the largest bike sharing service in Hamburg. The station-based service features 301 stations distributed across the city with a total of 3,700 bikes [9]. Historical information about the stations and their respective usage behavior was obtained from Hamburg’s Urban Data Platform [2].

3.2. Bike Sharing Algorithm

We implement a station-based bike sharing algorithm capable of adding, removing and relocating stations. Additionally, the initial number of bikes available per station can be adjusted. Thus, this algorithm may help decision makers in expanding and enhancing their bike sharing service. In the following, we explain the underlying methodology and distinguish three different scenarios for how agents select their pick-up and drop-off station.

The implemented bike sharing algorithm has components in both, the long- and short-term module of mobiTopp. Within the long-term module, multinomial logit models are used to allocate bike sharing memberships to agents. In the membership model, we differentiate between people living in the city or the surrounding area and those who are tourists. In the subsequent simulation, therefore, only agents possessing a membership can choose bike sharing as a means of transportation. The actual mode choice is implemented in a cross-nested logit model. Among others, both travel times as well as travel costs are explanatory variables in the model. When using bike sharing, the travel time consists of three components: the travel time by bike in between the two stations, along with additional time allocated for access to and egress from the bike sharing station by foot.

The locations where activities are performed are referenced with precise geocoordinates, they are obtained using the method proposed by Klinkhardt et al. [13]. Since bike sharing stations are also georeferenced, access and egress travel times by foot are calculated from the stations and can be accurately calculated. The travel time considered for the actual bike sharing leg is identical to the one of private bikes. Since most other components of mobiTopp are based on traffic analysis zones, the travel times are also approximated at zone level, to ensure consistency. Instead of using the same travel times as for private bikes, a separate travel time matrix for bike sharing could also be integrated without additional effort and thus reflect slightly slower speeds observed in reality [26]. The price for using bike sharing for a specific trip is based on StadtRAD’s standard fare [9]. As a result, no costs are incurred for trips of less than 30 minutes, while a fee of ten cents is charged for each additional minute. Theoretically, additional fare options could also be included.

For each trip, in the mode choice decision model it is first determined, whether a zone with a bike sharing station is within a catchment area. Our algorithm allows to specify the catchment area of bike sharing stations by providing a radius for zones included. For this paper, we applied a 500 meter radius as suggested in previous research [15, 16]. As the walking distance to the chosen station has a negative influence on the travel time and thus on the attractiveness
of the means of transportation, even larger catchment areas can be suitable to allow agents a more flexible behavior. Furthermore, the subset of stations is reduced to only those having bikes available at the time of departure. In times of mobility apps and the possibility to reserve bikes, the assumption of well informed agents seems reasonable. Consequently, only stations within the subset of zones with bikes available can be selected by an agent. The determined catchment area also applies to the subset of return stations for the bikes. However, there is no limit to the number of bikes per station in the city of Hamburg [9]. Hence, we did not include a limitation in our model either. The only constraint is that pick-up and drop-off stations cannot be the same, to avoid infeasible solutions, as an activity must occur in between two trips.

Particularly in the city center, it is likely that several stations have to be included in an agent’s decision-making process, as they are all located in the determined catchment area and bikes are available. A decision must therefore be made on which station to choose. As described in Section 2, most studies simply use the closest station. However, there are different strategies for selecting pick-up and drop-off stations for the rented bikes. We distinguish and implement the following in our algorithm:

- **Scenario 1:**
  In the first scenario, for each bike sharing trip, the nearest station to the start and end of the planned trip is selected. The proximity is determined by the walking distance.

- **Scenario 2:**
  In this scenario, reliability is taken into account. Thus, the pick-up station within the catchment area having most bikes available is chosen. Therefore, this model can be understood as a gravity model. However, bikes are returned at the nearest station to the destination, as done in Scenario 1.

- **Scenario 3:**
  In the third scenario, not necessarily the closest station is taken, but the entire distance to be covered is minimized, leading to more direct connections.

The three main scenarios are illustrated in Figure 1. O and D represent origin and destination of an agent’s trip, while S1 to S5 are the stations available within the respective catchment area C for an agent. The size of the station indicates the number of bikes available. Each bike sharing trip comprises of three legs L. An access leg L1 to the pick-up station and an egress leg L3 from the drop-off station as well as the main leg by bike L2. In Figure 1, Lx,y denotes leg x in scenario y.

Furthermore, our algorithm is tested on two different representations of agents’ attitudes towards station proximity, resulting in a total of six simulations. In the cases where a behavioral parameter was adjusted, the value was set to the estimate of walking to an e-scooter, which is slightly more sensitive in our model but still appears comparable.

Once the most suitable stations for pick-up and drop-off have been selected, the overall travel time and subsequently the travel cost is computed for all three legs. Thus, all required information to calculate the agents utility for bike sharing as means of transportation is available and the utility is calculated. Then, a discrete choice model is applied to determine the mode of transport used by the agent for this trip. If bike sharing is preferred over the other available
modes of transportation, a bike at the desired pick-up station is reserved for the agent. Since all agents are simulated subsequently within a one minute time step, it cannot happen that bikes are overbooked and agents arrive at a station without bikes available. Once a bike has been returned at the drop-off station, it can be included in a new mode choice request of another agent. All steps of the implemented station-based bike sharing algorithm as part of a multi-modal travel demand model are illustrated in Figure 2.

4. Results

In total, we performed six simulations that include each of the three scenarios illustrated in Fig. 1 with two different behavioral attitudes of the agents with respect to station proximity. This allows both between and within-scenario comparisons. Figure 3 illustrates the usage behavior over the simulated day in Scenario 1. The graph shows the total number of bike rentals within a 60-minute period. A typical demand profile with a distinct morning peak and a broader peak in the afternoon can be observed, indicating plausible results of our simulation. Since all agents are at their home location at the beginning of the simulation, there are only a few bookings at the beginning. When modeling more than one day, this would no longer occur from Tuesday onwards. This also explains the difference in the number of rented bikes at the beginning and at the end of the simulation. All other scenarios lead to similar and therefore plausible demand profiles, too.

In Figure 4, each point represents a station. The number of bikes available at each station at the start of the simulation is compared to the relative frequency of use for bike pick-ups at each station. For clarity, the two associated regression lines are also included. In Scenario 1, there is almost no correlation with $cor_1 = 0.06$. In Scenario 2, however, there is a high correlation of $cor_2 = 0.59$. This is reasonable since Scenario 1 only minimizes walking distances, while Scenario 2 considers the number of bikes available, leading to bigger stations being used more frequently. It can also be seen that stations with no bikes at the beginning of the simulation are used during the day anyway in both scenarios, as bikes have been returned at these stations previously. In addition, a positive correlation between the
number of traffic analysis zones in the catchment area of a station and the relative frequency of use of bike pick-ups at the respective station between 0.16 and 0.41 can be determined.

To further prove the plausibility of the results, we carried out a spatial analysis. 98 % of the stations were used, either for bike pick-up or drop-off. In the case of the more sensitive behavioral parameter for access and egress walking legs, 86 % of the stations were used. Stations not selected by agents are neither exclusively in the city center nor in more rural suburbs. Comparing the three scenarios, the observed behavior is identical in Scenario 1 and Scenario 3, while the modal split for bike sharing in Scenario 2 is significantly lower by 12 % ($p < 0.0001$). When agents become more sensitive to walking to or from a station, there is a significant decrease in the modal split for bike sharing, too. This statistical significance remains consistent across all scenarios, as verified by a two-proportions z-test with $p < 0.0001$. Between 95 % and 97 % of all bike sharing trips had a maximum duration of 30 minutes with an average ranging from 12.5 to 13.0 minutes, which is consistent with the behavior observed in Hamburg [3] and can be partly explained by the fact that the first 30 minutes are free [9]. Table 1 compares the average distances observed for all three legs in the six simulations, with the index $\beta$ indicating the changed behavioral parameter.

Table 1. Simulated mean distances for access, main and egress leg in kilometers

<table>
<thead>
<tr>
<th></th>
<th>Scenario 1</th>
<th>Scenario 1$_{\beta}$</th>
<th>Scenario 2</th>
<th>Scenario 2$_{\beta}$</th>
<th>Scenario 3</th>
<th>Scenario 3$_{\beta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{D}$ Distance $L_1$ (access, foot)</td>
<td>0.654</td>
<td>0.392</td>
<td>0.939</td>
<td>0.604</td>
<td>0.654</td>
<td>0.392</td>
</tr>
<tr>
<td>$\bar{D}$ Distance $L_2$ (main, bike sharing)</td>
<td>3.850</td>
<td>3.715</td>
<td>3.870</td>
<td>3.833</td>
<td>3.850</td>
<td>3.715</td>
</tr>
<tr>
<td>$\bar{D}$ Distance $L_3$ (egress, foot)</td>
<td>0.589</td>
<td>0.410</td>
<td>0.589</td>
<td>0.421</td>
<td>0.589</td>
<td>0.410</td>
</tr>
</tbody>
</table>

Although the differences and similarities between the scenarios seem obvious, we analyzed these statistically. Since the distances shown in Table 1 are not normally distributed (Shapiro-Wilk test $p < 0.0001$), no ANOVA can be performed to compare the scenarios. Therefore, non-parametric Kruskal-Wallis tests and a pairwise post-hoc comparison with Dunn’s test were carried out.

Only the access time in Scenario 2 differs significantly from the two other scenarios. This is reasonable, as agents are not forced to minimize the distance traveled. As stated previously, Scenarios 1 and 3 yield identical results, although they could theoretically be different, as shown in Figure 1. This was also verified by simulations with different initial random numbers. Consequently, there may be differences if the number of bike sharing stations in the model is changed or if the model is transferred to a different planning area.

All scenarios produce reasonable results. Although we could identify differences in the three scenarios implemented, the impact of precise estimates of people’s behavior is of greater influence. Thus it can be sufficient for most...
applications to let agents choose the closest station with bikes available as stated in Section 2. However, we showed that there can be different strategies of agents behavior implemented in agent-based travel demand models depending on the actual circumstances and behavioral attitudes. For a precise representation of real behavior, in our opinion this approach provides a greater flexibility and accuracy in modeling station-based bike sharing.

5. Summary

We presented an approach to model station-based bike sharing in an agent-based travel demand model that allows to easily add and remove bike sharing stations, as well as to manipulate the number of bikes available at the beginning of the simulation at each station. Further, our algorithm allows modeling different strategies for how agents select their pick-up and drop-off bike sharing stations. While there are no differences between minimizing access and egress distances compared to minimizing the total travel distance in our model, a significant difference can be observed when considering bike availability as done in Scenario 2. The authors are aware that this scenario is exaggerated. However, in the case of unreliable systems, it may still make sense to walk a bit further but have a greater number of bikes to choose from. To determine which scenario is most appropriate for a particular use case, behavioral or tracking data is needed to account for regional differences. Depending on the planning area, also further scenarios and influencing factors could be implemented, such as considering topography. However, since our model region has almost no differences in height, we dispensed this factor. In our approach, we minimized distances, as usually done in the literature. With a precise bicycle network as a basis, allowing for different speeds depending on the infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the re-balancing of bikes between stations is of great interest to practitioners. Therefore, it is beneficial to include this in our model to also test different bike re-balancing strategies. Furthermore, we aim to extend our methodology to inter-modal trips, e.g. to study the changing behavior between bike sharing and public transport. Due to the high computation time of agent-based travel demand models [25], it is reasonable to investigate whether our approach results in a significant increase in computation time compared to a heuristic at traffic analysis zone level.

Acknowledgments

This study is part of the project #transmove at the DigiLab of the State Agency for Roads, Bridges, and Waterways (LSBG) for the Free and Hanseatic City of Hamburg. It is funded under the ”Digitalization of Municipal Transport Systems” grant program by the Federal Ministry for Digital and Transport (BMDV) (Grant Reference Number: 16DKV41086). The model is provided by courtesy of MOIA. Further, Deepl and ChatGPT were used to ensure linguistic accuracy.

References

We presented an approach to model station-based bike sharing in an agent-based travel demand model that allows different strategies of agents' behavior implemented in agent-based travel demand models depending on the planning area. Depending on the planning area, also further scenarios and influencing factors could be implemented, such as considering topography. However, since our model region dispensed this factor. In our approach, we minimized distances, as usually differences in height, we used a precise bicycle network as a basis, allowing for different strategies for how agents select their pick-up and drop-off of the simulation at each station. Further, our algorithm allows modeling different bike-rebalancing strategies. Furthermore, we aim to extend our methodology to inter-modal trips, and influencing factors could be implemented, such as considering topography. However, since our model region would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of infrastructure, it would be interesting to minimize travel times instead. As explained in Section 2, the rebalancing of