This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

Identifying and Planning for Group Travellers in On-demand Mobility Models

Grace O. Kagho*, Milos Balac * *Institute of Transport Planning and Systems, ETH Zurich

***Abstract*—Understanding group travel is vital for transportation planners and policymakers, especially when modelling emerging on-demand mobility such as ridesharing and shared autonomous vehicles. Existing agent-based simulations of ridesharing services hardly consider group travel, even though these services mainly occur during the weekend and for leisure trips where people are more likely to travel in groups. This is due to the limited availability of group travel data in many travel demand models. This study uses a Swiss synthetic travel demand where car drivers and passengers are modelled separately to identify group travellers. A heuristic approach based on mixed integer linear programming is implemented to create group travellers by matching car drivers and passengers. An agent-based simulation model is set up to simulate ridesharing while considering group travel to reveal the impact on operational policies for ridesharing.**

***Index Terms*—Group travel, travel party size, agent-based models, on-demand mobility, ridesharing.**

I. Introduction

As the field of transport planning evolves, planners are increasingly considering the impacts of emerging on-demand mobility options like ridesharing and Shared Autonomous Vehicles (SAVs). Over the last decade, there has been a surge in on-demand mobility simulation studies, particularly those utilizing agent-based models. Agent-based models are powerful tools for modelling complex problems related to travel behaviour at the microscopic scale. They offer a foundation for crafting policies and operational decisions for on-demand mobility that could directly impact travellers, ultimately allowing for more effective travel demand management.

Typically, in agent-based on-demand simulation models, potential travellers can request a ride at any time, either directly from their doorstep or a designated area, and a vehicle arrives to transport them to their desired destination and activity. Behind this operation is a centralized dispatching system, which balances the management of an available fleet of vehicles against service efficiency and effectiveness to ensure optimal service levels. However, an important omission from most simulation studies is the consideration of group travel. Conventional simulation approaches often model a request as a single person, even though one request might involve multiple individuals travelling together for a specific activity. Group travellers are typically modelled independently, but this can lead to inaccurate estimations of group travel impacts on transport infrastructure, such as fleet size requirements or the formulation of pricing and subsidy schemes.

This paper was produced by the IEEE Publication Technology Group. They are in Piscataway, NJ.

Manuscript received Month XX, 2023; revised Month XX, 2023.

Despite the growing importance of on-demand mobility, existing models for ridesharing services often neglect group travel, even though these services have been shown to occur majorly during the weekends and for leisure trips where persons are more likely to travel in groups [1], [2]. The Swiss 2015 household travel survey (HTS) [3] reveals that most weekend trips have a higher percentage of car passenger trips (about 19%) compared to weekdays (9%). The richer 2008 French National HTS data [4] shows that 64% of group trips are made within households, out of which 80% were made by car, 16% by walking, and 3% by public transport (PT). For non-household group trips, 70% by car, 19% by walking, and 8% by PT. Clearly, families or groups of friends tend to travel together and carpool. This behavioural trend suggests that group travel should be a key consideration when simulating on-demand mobility in travel demand models, especially when modelling weekend travel patterns.

However, modelling group travel has proven challenging due to data limitations and difficulty in identifying group travellers. The methods for identifying group travellers range from traditional discrete choice models to more complex optimization algorithms. Discrete choice models have been used to identify group travellers in households across different travel demand models [5]–[7]. [8] and [9] used transit smart card datasets to identify group travellers that travel with public transit by grouping transit passengers by their time of entry and exit from a train. However, these methods have not yet been replicated in simulation models due to data availability and implementation constraints.

Carpooling, a term typically used when individuals familiar with each other share a private vehicle ride, is a form of group travel, and has been well-studied with a rise in models that match carpool passengers to carpool drivers [10]. Furthermore, matching passengers and drivers to create group travellers can also be seen as a dial-a-ride problem. Carpooling and ridesharing services have been studied with various similar approaches to match groups of travellers and drivers, typically by solving a vehicle routing problem using optimization techniques and heuristic approaches such as local search, hybrid genetic algorithms, and simulated annealing [11]–[14]. Some have improved these matching algorithms through the use of generalized cost [15]. Usually, different objective functions exist in these optimization approaches with a focus on reducing trip distances between drivers and passengers, system-wide vehicle kilometres travelled, travel cost, or maximizing the number of participants in a vehicle.

Despite these advancements, realistically simulating car passenger trips in travel demand models remains challenging.

Car passenger trips can occur within different contexts such as households, personal social networks, or even between strangers. Furthermore, the availability of a passenger mode option depends on the behaviour of the driver offering a ride. Thus, car passenger trips are often modelled simplistically. While there have been a few attempts to model group travel in travel demand simulation models [16], they often resemble a ride-hailing or carpooling system rather than generating group travellers that travel from the same households or locations.

In the context of this study, ”group travel” refers to multiple individuals, whether from the same household or not but known to each other, who travel together from one or more origins to one or more destinations. This differs from ”ridesharing” (commonly associated with shared taxi services like UberPool) and aligns more closely with ”carpooling”. In contrast to typical carpooling studies, which focus on matching passengers to drivers and minimizing time, distance, and travel costs, this study aims to identify individuals who travel together from the same household or location.

Overall, the primary goal of this paper is to highlight the importance of integrating group travel into on-demand mobility simulations. As not doing so, can create a false impression of the impacts of these services on the transport system. Furthermore, this paper seeks to show how the impact of group travel varies between weekdays and weekends and to quantify these differences.

The subsequent sections of this paper present an approach for identifying group travellers using synthetic travel demand data. The study further includes a simulation case study to demonstrate the influence of considering group travel in on-demand mobility simulations. The Swiss synthetic travel demand data is used to identify potential group travellers, with a heuristic approach based on mixed-integer linear programming implemented to match car drivers and passengers, who are grouped into travel groups. Such a group could comprise members of the same household or a circle of friends. The matching process, which ensures that matched group travellers have similar departure times, origins, and destinations, is necessary to obtain groups as they are not available directly from the data. The approach also considers vehicle capacity constraints and the overall vehicle occupancy distribution of the region, ensuring that the essence of group travel is preserved while optimizing the driver-passenger matching based on time and distance considerations. Finally, an agent-based simulation model is set up to simulate ridesharing while considering group travel.

II. Methodology

Group travel can be classified based on [17]'s four patterns of ridesharing: identical, inclusive, partial, and detour ridesharing. For example, in identical ridesharing, both the origin and destination of the driver and passenger are the same. In partial ridesharing, the pick-up and drop-off locations of the passenger are on the way of the driver's original route, but either the origin or the destination is not on the way. In this study, we first explore the problem of identifying group travellers as travellers with the same origin and destination and similar departure time, using a car as their mode of travel, i.e., identical group travellers.

A. Matching

Matching problems are not new, as they occur wherever there is a need for a proper allocation of resources. Online matchmaking, network flows, and Uber/taxi pickup problems are a few examples. They are sometimes described as classical assignment problems or graph problems. In this context of driver-passenger matching to identify group travellers, the problem follows a minimum weight matching in a bipartite graph problem where we can represent each driver and passenger as a node and draw an edge between a driver and a passenger if they can be matched. The weight of each edge represents the cost of the matching, which could be based on factors such as the distance between the origins and destinations of the driver and passenger or the departure time required for the trip. The following are assumptions for the matching:

- A travel group is made up of two or more persons, depending on the vehicle capacity, with only one person as a driver
- Persons in a travel group may be within the same household or among friends
- A travel group consist of persons whose trips have approximately the same departure time, point of origin, and destination

Therefore, given two subsets, N and M, representing car passengers and drivers, respectively, the problem instance is described by an $n \times m$ matrix C, where each $C[i, j]$ represents the cost of matching node i of the first subset, car passengers, and node j of the second subset of car drivers. The cost is made up of two components: time and distance. The time cost, t_{ij}, is defined as the absolute difference in departure times between passenger i and driver j. The distance cost, d_{ij}, is computed as the sum of two components: the distance from the driver's origin to the passenger's origin, and the distance from the passenger's destination to the driver's destination. These two cost components are combined into a single measure by applying a weighted sum. The cost of matching a passenger with a driver is thus defined as in follows in Equation (1):

$$c_{ij} = \sum(\alpha \times t_{ij} + (1 - \alpha) \times d_{ij}) \quad (1)$$

where α is a parameter that determines the relative importance of time versus distance. A higher value of α would give more importance to the time cost in the matching decision, while a lower value would emphasize the distance cost.

The aim of the matching algorithm is to find a complete matching of the drivers and passengers by minimizing the total cost of matching passengers with drivers, subject to certain constraints. These include limiting the number of passengers that can be matched with a single driver, and requirements on the vehicle occupancy distribution in the region. Figure 1 presents the matching algorithm. The linear problem is thus formulated as follows:

Minimize

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} (c_{ij} \times x_{ij}) \tag{2}$$

where $x_{ij} = 1$, if passenger i is matched with driver j otherwise 0

Subject to:

Each passenger is assigned one driver

$$\forall i \in \{1, \ldots, n\} \sum_{j=1}^{m} x_{ij} = 1 \tag{3}$$

Each driver is assigned at most k passenger

$$\forall j \in 1, \ldots, m \sum_{i=1}^{n} x_{ij} \leq k \tag{4}$$

The occupancy distribution of all drivers is close to the desired defined distribution

$$\forall l \in \{0, \ldots, k\} Z_l - \epsilon \leq \sum_{j=1}^{m} Y_{j,l} \leq Z_l + \epsilon \tag{5}$$

whereby l represents the occupancy level of a vehicle, i.e., the number of passengers in a vehicle. It ranges from 0 (vehicle is empty) to k (vehicle is at maximum capacity).

Z_l represents the desired occupancy distribution of vehicles, i.e., the number of vehicles with an occupancy level of l. For instance, Z_0 would represent the desired number of empty vehicles, Z_1 would represent the desired number of vehicles with one passenger, and so on up to Z_k, which represents the desired number of vehicles at maximum capacity.

The problem can be seen as a variant of the classical assignment problem studied by researchers using various methods, including mixed-integer linear programming (MILP) and solvers such as Gurobi or CPLEX. However, the problem formulated above is not a standard matching problem, given that N and M can be as large as 500,000 elements; solving this problem using a MILP solver can be computationally expensive and may not scale well. Therefore, a heuristic method is proposed to solve the problem. The algorithm implements a greedy approach to find the minimum cost pairing between drivers and passengers iteratively. The motivation behind a greedy algorithm is that it can provide near-optimal solutions more efficiently than traditional optimization techniques, especially for large-scale problems.

The algorithm identifies the next 'best' match for a passenger at each iteration under the following conditions:

1) The driver has not reached the maximum passenger limit.
2) The cost of pairing with the current driver is lower than with any other driver.
3) The number of matches for the current driver is within the specified passenger occupancy distribution range.
4) The number of drivers with the current number of matches is less than the occupancy distribution value for that number of matches.

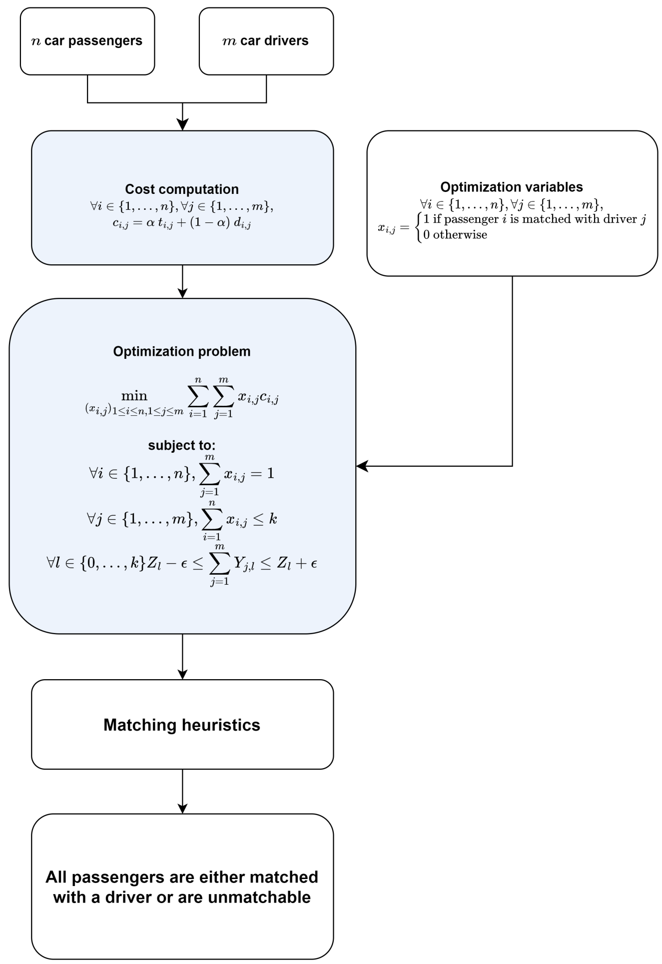

Fig. 1. Matching Algorithm

The algorithm, in its original form, has a time complexity of $O(n^2)$, where n represents the number of passengers and drivers, with the code's efficiency depending on the size of the cost matrix. However, to enhance the performance of the algorithm, the strategy is to reduce the number of potential driver candidates. This is achieved by segmenting the driver pool into time bins and only considering drivers within a specific time bin of a passenger as potential matches. This allows the algorithm to focus on a manageable subset of drivers, thereby reducing computation time. A minimum threshold for the number of candidates is set to ensure that there are adequate candidates for efficient matching. If the number of candidates within a time bin falls below a minimum threshold, the time bin is expanded incrementally in both directions until either the threshold is met or a pre-defined maximum number of attempts is reached. If these attempts are unsuccessful, the passenger is considered unmatchable. To ensure realistic pairings, a constraint is added to ensure that driver candidates have trip lengths greater or equal to the passenger's considered trip length. Further, a maximum acceptable difference for both time cost and distance cost can be defined to filter out unlikely candidates during the matching. Both scenarios, with and without the filtering constraints, are

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

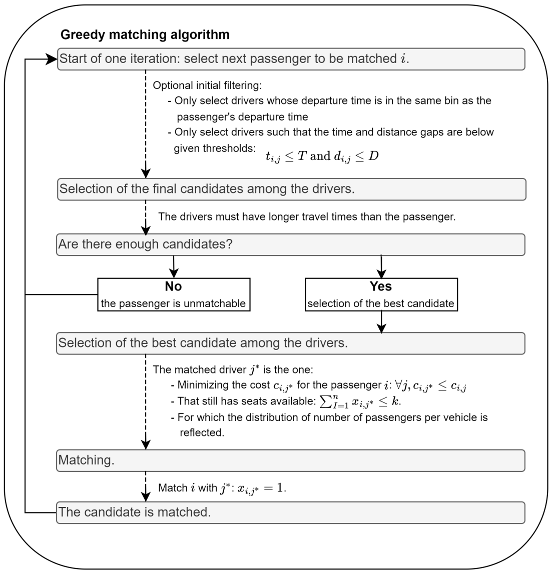

Fig. 2. Greedy Matching Heuristic

explored in this study to understand their impact on the algorithm's performance and the resultant matchings. This greedy matching heuristic approach shown in Figure 2, while it may not guarantee a globally optimal solution, provides a practical approach to solving a large-scale group travel matching and offers a balance between computational efficiency and solution quality (See Appendix A for a small scale comparison between a standard MILP algorithm and the heuristic approach).

B. Simulation Framework

The multi-agent transport simulation framework MATSim [18] is used for this study and extended to include the capability for ridesharing and group travel. MATSim provides the ability to model large-scale transport systems realistically. It requires the following scenario data to emulate the transportation system: a network of links and nodes and travel demand in the form of a synthetic population of agents with their travel plans and other transport elements such as facility location and transit schedules. Agents in MATSim are iteratively routed and simulated through the network, using their chosen modes of transportation to get from one place to another while engaging in various activities. As the agents interact with each other in the network, congestion occurs, affecting their decisions in the next iteration. Agents optimize their plans in each iteration until the system converges to a steady state. This enables MATSim to simulate emerging behaviours that drive travellers' decision-making. In this study, only the network routing functionality of MATSim is used, which serves as the basis for future agent-based research. This is mainly because this work focuses solely on comparing ridesharing group travel patterns without the need to capture the entire travel behaviour of the transportation system.

To represent a fleet managing system for shared autonomous vehicles (SAVs) that allows for pooling and group travel, MATSim's demand-responsive transit (DRT) extension is used [19]. The DRT extension was developed to enable MATSim to simulate a dynamic ridesharing service where vehicles can pick up and drop off agents at their request. A central dispatching system that manages vehicles is responsible for scheduling or rejecting incoming requests. The dispatcher is presented with a list of available vehicles. It traverses the list and assigns each request to the closest vehicle after ensuring that constraints on wait time or detour times for passengers are not violated. These constraints are that: (1) the overall time spent on travelling for the passengers currently in a vehicle or waiting for the vehicle and that of the new customer does not increase beyond defined thresholds and (2) the expected boarding times for the awaiting customers and the new one are within a requested time frame. Should no suitable vehicle be available, the request is rejected. This is described in detail in [19]. Only individual passenger trips are considered for each request, and as such, group travel is not captured in the existing system. In reality, a single request might involve a group of travellers, and this will require more than one seat, which must be considered when assigning an appropriate vehicle for such a request. Therefore, in this study, a constraint for group travel is added to the dispatching algorithm to check the number of passengers in the vehicles to ensure enough seats are available to accommodate group travel requests.

III. Case Study

In this study, we enrich a subset of the Swiss synthetic travel demand with information about group travellers based on the Swiss HTS. The focus is on Zurich city consisting of the 12 districts of the city of Zurich, extended by a buffer of 5 km as shown in Figure 3. The Swiss synthetic travel demand model, often referred to as the Switzerland Baseline scenario [20], consists of a synthetic agent population that reproduces Switzerland's sociodemographic characteristics and travel behaviour.

The Swiss HTS surveys only one person per household and only captures when a person has taken a trip within a group, either as an accompanying person or as a car driver, including the number of people in the car. Information on whether they are from the same household or just friends is not captured. Using this information, car passenger trips have been identified and treated as their own distinct modes when generating the Swiss synthetic population for travel demand. This scenario initially developed for the average work day has been extended to represent weekend travel behaviour accounting for the differences in trip patterns between a typical workday and the weekend.

The modal shares from the cut-out Zurich travel demand for the modelled days were calibrated using MATSim and validated against the Swiss HTS modal shares for the same region. Out of the suite of available modes: bike, walk, public transit, car (driver) and car passenger, the share of trips where travellers drove cars or were car passengers are shown in Table I, highlighting increased group travelling during the weekends, as a substantial share of car passenger trips happens on Saturday and Sunday compared to an average workday. "Sim" refers to the simulated modal shares.

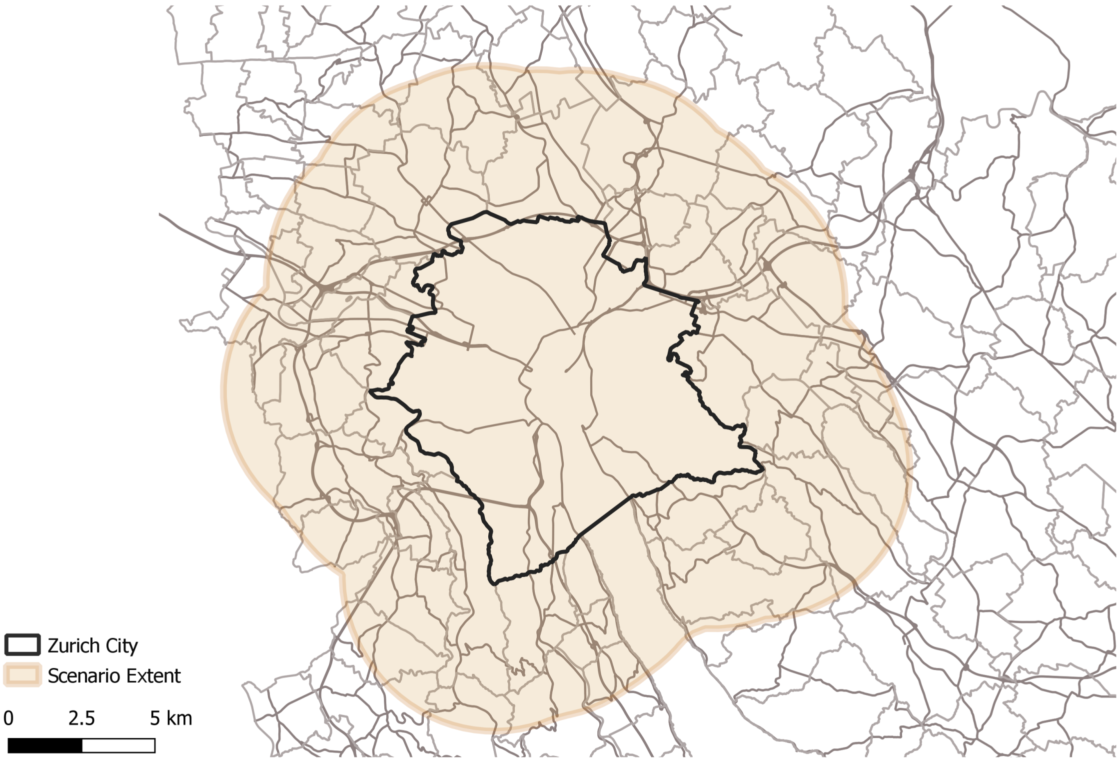

Fig. 3. Analyzed study area with Zurich's city limit indicated with black line)

TABLE I
MODAL SHARES

Mode	Saturday HTS (%)	Saturday Sim (%)	Sunday HTS (%)	Sunday Sim (%)	Avgday HTS (%)	Avgday Sim (%)
car (driver)	28.74	28.85	30.45	29.50	29.19	29.27
car passenger	17.73	17.67	18.9	21.0	9.14	8.46

With the aim to identify group travellers by matching car drivers and passengers using the Swiss synthetic travel demand, the travel demand models for Zurich for an average working day, Saturday and Sunday, are converted to trip-based models. Only car driver and car passenger trips that start and end within the region are considered, leaving 456 148, 539 629, 371 382 trips with an average vehicle occupancy of 1.51, 2.02, and 2.14, for an average workday, Saturday, and Sunday respectively, based on the modal shares highlighted in Table I. These car drivers and passengers' trips are used to identify and generate group travellers using the matching algorithm described above. After the matching is done, a new population file of all agents is generated, whereby all agents are assigned the SAV service as their mode of travel and agents identified as group travellers are given an attribute to specify that they perform trips with other corresponding agents. Unmatched trips are represented as single traveller's trips that use the SAV service individually.

A. *Scenario Definition*

For this study, 63 scenarios were simulated, consisting of the three days of interest (average workday, Saturday, and Sunday), a case where SAV service requests are made by groups of travellers, single travellers that have individual trips or single travellers that have individual trips synched with their matched trips, and a set of 7 fleet sizes. These scenarios are briefly described below and summarized in Table II.

- Group travel: For each simulated day, three scenarios were generally defined with a difference in terms of the presence or absence of group travel and the characteristics of the trips. These scenarios include:
  Individual scenario
  - This represents the baseline scenario

TABLE II
SIMULATION SCENARIOS

Scenario Attributes	Attribute Levels
Days	Workday, Saturday and Sunday
Grouping Scenario	Individual, Group and Synched Schedules
Fleet Sizes	4 000 - 10 000

  - All agents have their own unique, uncoordinated activity chain
  - All agents request SAV individually and independently

Group scenario:

  - Car drivers and car passengers are "matched", meaning that a car driver's trip represents both their trip and its matched car passenger(s) forming a group of travellers. Consequently, the corresponding matched car passengers are removed from the scenario.
  - Matched car drivers make the SAV requests, representing their travelling group
  - The request made for these matched trips contains the travel party size

Synched Schedule scenario:

  - Car drivers and car passengers are "matched", but a car passenger is a separate agent, albeit with the same activity chain as their matched car driver.
  - Matched car drivers and car passenger(s) request SAVs independently and as individuals

For example, the scenarios designed to represent an average workday are named "AvgdayIndividual", "AvgdayGroup" and "AvgdaySynched". In the "AvgdayIndividual" scenario, matched pairs of car drivers and car passengers use SAVs for their trips individually and independently. This is the default way these trips are generally simulated in most studies and thus serves as the baseline scenario in this study. In the "AvgdayGroup" scenario, matched passenger trips are removed because their trips have been merged with the corresponding car drivers. In the "AvgdaySynched" scenario, however, matched travellers have the same travel plans as the original car drivers to whom they are matched, i.e., the same trip origin and destination locations and departure time. This was done similarly for the other two days.

- SAV Fleet Size: Fleet sizes between 4 000 and 10 000 are simulated for all scenarios. Fleet sizes are varied because this is one of the most important operational measures for operators and planners to consider when offering such services, as it affects overall service levels and customer experience. The range of fleet sizes was selected based on results of related studies [21], [22], with assumptions made about the number of vehicles that could serve the proportion of car trips in the study region.

B. Matching configuration

The matching algorithm utilized in this study requires several key configurations to ensure accurate and reliable results. These configurations include:

- Trips data frame: The validated Zurich travel demand model is used to provide this data. A data frame consisting of trip information such as departure time, trip length, origin and destination coordinates, and mode of travel (i.e., driver or passenger).
- Passenger occupancy distribution: Data from the 2015 Swiss HTS is used to determine the probability distribution of passenger occupancy for the study region and simulated days. This probability distribution indicates the likelihood of each driver having a given number of passengers. It is used to randomly assign a different number of passengers to each driver to create a passenger occupancy distribution that allows all passengers to be matched. See Table III for the occupancy distribution.
- Maximum number of passenger seats: A maximum of five seats is defined based on typical car sizes that seat up to five people, e.g. a small sport utility vehicle (SUV). This value has also been selected while keeping the integrity of the average vehicle occupancies in the HTS.
- Cost weighting value: The cost function in Equation (1) is the sum of the weighted differences in departure time and distance. The α value weights the relative importance of time cost versus distance cost. Here α = 0.2 is used, penalizing distance cost more and placing greater weight on time cost. See Appendix B for sensitivity analysis of α values from 0.1 to 0.9. The departure time filtering constraints described next also make the departure time difference more constrained before matching.
- Time binning value and threshold: A time bin of 5 minutes is used, including drivers with departure times within 5 minutes before or after a passenger's time bin. A minimum threshold of 20 candidate drivers per time bin is required for matching. If this threshold is not met within five draws, the time bin is expanded to a maximum of 45 minutes. If still unmatched, the passenger remains unmatched.
- Filtering values: To ensure that passengers matched have reasonable values for differences in the departure time as well as their origin and destination, a maximum difference in departure time and the maximum difference in origin and destination coordinates can be set. Here, a maximum value of 10 minutes is defined for the difference in departure time, following the recommendation by [7]. For distance, there is more allowance with a 4km bound. This choice is based on initial sensitivity analysis, which showed that about 70% of the matched agents were within the 4km bound while about 95% were within 10km. The assumption here is that perhaps about 30% of the unmatched passengers may have been passengers that were picked up along the way and cannot be identically matched.

The matching configuration values are summarised in Table III

TABLE III
MATCHING CONFIGURATION VALUES

	Average workday	Saturday	Sunday
No. Drivers	302598	343779	231257
No. Passengers	153350	195850	140125
Average Occupancy	1.507	2.027	2.144
Probability distribution			
0	0.73	0.39	0.35
1	0.20	0.36	0.39
2	0.04	0.12	0.11
3	0.02	0.09	0.10
4	0.01	0.02	0.04
5	0.001	0.014	0.010
Occupancy distribution			
0	194262	135138	80218
1	79149	125007	89549
2	17078	40780	24524
3	8917	29734	24176
4	2440	8173	10446
5	752	4947	2344
No. passenger seats	5		
Cost weighting value	0.2		
Time bin (max. expanded time) [min]	5 (45)		
Filtering time cost [min]	10		
Filtering distance cost [km]	4		

C. SAV Simulation Configurations

The SAV service in this study operates door-to-door for 30 hours with vehicles that can accommodate up to 5 passengers. The passengers can either be a group of travellers or single travellers. Although the DRT extension provides a possibility to set rejection constraints for a maximum wait time and maximum detour time (the time an in-vehicle passenger is willing to lose while the vehicle picks up another passenger), this is not considered. Rather, a constraint is set for rejection when there is no vehicle available with a suitable number of seats that can pick up agents travelling in groups. This means very high wait times or detour times could happen during peak demand periods. The SAV vehicles are initially placed based on the population density of the network. When a vehicle is empty, the vehicle remains at its last drop-off location until it needs to serve another request. The simulation goes through four iterations to reach an equilibrium state in terms of the number of SAV trips served, average travel times, and waiting times. This is because the system uses a dynamic vehicle routing problem for its dispatching algorithm, where the free speeds (i.e., speed without congestion) of the network are used to calculate the travel times in the first iteration and updated later until the times are stable.

MATSim version 15.0-SNAPSHOT (last accessed on 5th March, 2023) has been used with the corresponding DRT version. All car drivers' and passengers' trips from Zurich's 100% population is simulated. Although this requires more computational resources due to the dispatching algorithm used by the SAV service, the negative effects of using a sample population can be avoided [23]. The simulations were performed using the ETH High-Performance Computing Cluster (Euler) with a maximum resource request of 8 cores and 192 GB memory.

IV. Results and Discussion

This section is in two parts. The first part describes the results of the matching algorithm, while the second part presents findings from the simulation of the ridesharing scenario with group travel behaviour.

A. Matching results

Matching stats

The results from the matching process are summarized in Table IV. The results differ depending on the day being modelled and whether filtering constraints are applied to the matching algorithm, i.e., "With filtering" or "Without Filtering". The results of the matching algorithm show that without filtering, the percentage of successful matching is 94% for the workday and 99% for Saturday and Sunday scenarios. However, the matching algorithm was able to successfully match around 78%, 79% and 78% of the passengers with car drivers when filtering for matches whose time costs are within 10 minutes and distance costs are within 4 km (about 75% share of trip distances), for the workday, Saturday and Sunday respectively. Even though the average time cost was generally less than 2 minutes, and 99% of trips with about 7 minutes in departure time difference across the different days, the distance costs on average were almost 3 km on average, with about 25% of the matches having very large differences up to 35 km for the average workday. Many possible reasons exist for these differences, ranging from specific constraints used in the matching, specific characteristics of the population or even differences from how the synthetic population was generated or, perhaps, a certain share of group trips having unidentical origins and destinations with drivers picking up passengers along the way. It makes sense that Saturday and Sunday show higher matches as these days have more likelihood of group travel, even from the data. Further improvements beyond the scope of this paper could be made to the algorithm in order to increase the rate of successful matches.

Comparing the results for the workday and the weekend, one can see that the percentage of successful matching increases for both the case of with and without filtering. This suggests that there are fewer potential passengers and drivers available on workdays than on the weekend and one could easily point to the idea that the matching result only reflects the higher tendency for group travel on the weekends. However, looking at the relative change in the decrease, there is only a slight difference in how the different days perform. This indicates that the differences are related to the characteristics of the trips based on the synthetic travel demand, whereby a workday reflects less potential for group travel.

Matching Validation

To validate the matching algorithm, the distribution of passenger occupancy and trip distances from the algorithm's results were compared to those from the Swiss HTS data. Figure 4 compares the occupancy distribution used as input to the algorithm (based on multinomial logit probabilities) to the matching results, with and without filtering. Figure 5 presents the validation of the matching algorithm in terms of how well vehicle occupancies are represented in their trip distances for

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

TABLE IV
MATCHING STATISTICS

	Average Workday		Saturday		Sunday	
	Without filtering	With filtering	Without filtering	With filtering	Without filtering	With filtering
No. Unmatched passengers	8413	33428	569	39658	233	29835
% of successful matching	94.52	78.23	99.71	79.75	99.83	78.71
Time cost stats [min]						
Mean	1.06	0.87	0.99	0.88	1.00	0.91
Median	0.62	0.57	0.62	0.60	0.65	0.63
75%	1.27	1.15	1.27	1.20	1.30	1.25
95%	3.35	2.73	2.98	2.67	3.00	2.75
99%	6.78	4.82	5.67	4.32	5.33	4.35
max	24.82	9.90	24.37	9.88	23.97	9.82
Distance cost stats [km]						
Mean	2.74	2.02	2.89	2.15	2.96	2.20
Median	2.21	1.93	2.43	2.08	2.52	2.15
75%	3.37	2.67	3.62	2.83	3.73	2.89
95%	6.35	3.62	6.46	3.69	6.48	3.72
99%	11.07	3.92	10.04	3.93	9.75	3.94
max	32.99	3.99	33.78	3.99	34.24	3.99
Occupancy distribution						
1	79149	68925	75626	73613	55092	53556
2	17078	15168	25727	22627	19629	16618
3	6104	4605	9659	7281	7046	4958
4	2440	1409	3986	2558	3171	1576
5	752	282	4656	1050	2344	464
Computation time [min]	134	-	188	-	106	-

the workday. The observations here for the average workday are similar to the weekend days.

Figure 4 shows that the matching results slightly overestimate the number of drivers with 1-2 passengers by 3-6% for the different days. However, the relative difference increases for higher occupancies, particularly on Saturdays and Sundays. Similarly, this is reflected in Figure 5 whereby the distance distribution diverges for occupancies of 4 or more, as there are fewer observations to match at higher occupancies and longer distances.

Furthermore, in Figure 5, while the matching algorithm adequately represents lower occupancies and distances, the differences between the algorithm's and HTS' distributions increase for higher occupancies. Nonetheless, the algorithm shows promising results for the scope of this study, and the matched results with filtering constraints were then used for the SAV simulation of group travel.

B. Simulation results

The simulation results demonstrate that the operational performance and efficiency of a ridesharing system vary across different scenarios. Notably, factors such as vehicle occupancy, vehicle empty distance, wait time, and detour time are affected. Vehicle occupancy refers to the number of passengers per vehicle. The vehicle empty distance is the distance driven by an SAV without any passengers onboard. Wait time represents the duration of time a passenger must wait until an SAV arrives to pick them up. Lastly, detour time denotes the extra time required for a passenger to reach their destination when the vehicle takes a route deviation to pick up or drop off another passenger. Detour time is calculated as the detour factor, which is the ratio of the total travelled distance to the estimated direct travel distance.

The simulation results for the study are presented in Table V and Figures 6 and 7. Table V provides an overview of the simulation results for scenarios with a fleet size of 6 000. Figures 6 and 7 depict the operational performance of the SAV services. The blue lines represent the "group" scenarios, where group travellers can request a ride as a group, and the red lines represent the "individual" or "synched" scenarios. Figure 7 differs from Figure 6 such that in the "synched" scenarios, matched travellers have the same travel plans as the original car drivers to whom they are matched, i.e., the same trip origin and destination locations, and departure time. In contrast, Figure 6 retains the original travel plans for the matched travellers. The majority of the results' discussion focuses on the "individual" scenarios depicted in Figure 6, which can be considered the baseline.

Average vehicle occupancy: From Figure 6, one can immediately see some differences between the "group" scenarios (blue lines) and "individual" scenarios (red lines). The average vehicle occupancy highlighted in Figure 6a, shows that there are differences between "group" and "individual" scenarios, with the former pooling more passengers on average. The relative difference between the average occupancy of the "individual" scenarios and the "group" scenarios in this study is up to 100%. For Figure 7a, where the travellers in "group" scenarios are the single travellers in the "synched" scenarios,

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

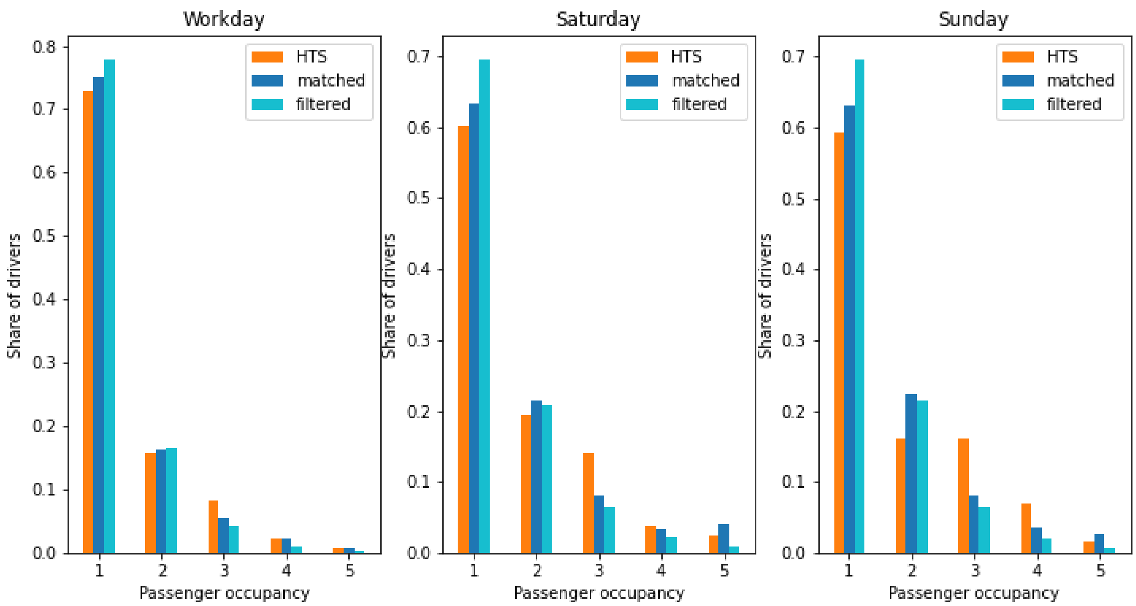

Fig. 4. Share of drivers for different passenger occupancy

the relative difference is quite smaller, only up to about 15%, demonstrating how well the dispatching algorithm of the ridesharing system can perform when group travellers are represented in the scenario even if singular requests are made by each member of the group. Even in this case, still not considering group travel in both the trip characteristics and the ridesharing system can result in an underestimation of the pooling capability of the system. The study thus highlights the need for ridesharing simulation studies to consider group travel to avoid underestimating its key benefit.

The average vehicle occupancy strongly varies by the day modelled, specifically for the SaturdayIndividual scenario, where the fleet size of 4000 in Figure 6a shows an outlier due to insufficient vehicles for the demand. This is possibly due to the fact that the SAV system is not allowed to reject passengers' requests for single trips unless there is no vehicle in the system able to serve its request within the operation time (based on the study design). Consequently, for the requests in the "individual" scenario, one can observe very high unrealistic detour times and even higher empty distances (see Figure 6b and d) as vehicles try to meet up the high demand.

Differences in fleet sizing: Fleet sizing is also impacted by group travel. The average empty distance in Figure 6b notably shows that the trends for the fleets are quite different when the party size of group travellers is taken into account. Taking as an example the Saturday scenario, which highlights the most obvious difference, for small fleets, the empty distances are higher for group travel compared to large fleets where group travel has lower empty distances. This means that a small fleet could result in an incorrect estimate of the empty distance, which is important as empty distances contribute to externalities. Furthermore, Figure 6a and d further highlights the impact of group travel on fleet sizing. In these figures, one can infer that in ridesharing simulations where the party size of group travellers is not taken into account, one could overestimate the number of vehicles needed to serve the demand. Here one can see that for the "group" scenarios, a recommended fleet size to serve the different days would be 5 000 vehicles compared to 6 000 vehicles for the "individual" scenarios.

Wait time and detour time: Wait time is not impacted much by group travel except for a smaller fleet size of 4 000, where it is slightly longer, as shown in Table V. It is worth noting that the wait time results are subject to limitations imposed by the study assumptions, such as the absence of other competing vehicles on the road. Therefore, the wait time outcomes may differ when other traffic elements are taken into account. Detour time, on the other hand, and as expected, reduces by up to 17% when group travel is considered, resulting in a more efficient system that avoids unnecessary detours. As a result, the average driven distance is lower for "group" scenarios, as illustrated in Figure 6c. This observation highlights the benefits of considering group travellers in ridesharing simulations, which can result in a more efficient system by avoiding unnecessary detours and reducing overall vehicle kilometres travelled. Larger groups have a reduced potential for multiple pickups along the way, which allows for higher occupancy rates and more streamlined routes.

Service rate: Table V provides an overview of the simulation results. As expected in a system where wait time and detour time rejection constraints are not considered, rejections occur when there are no available vehicles to serve passengers. In the "group" scenarios, there are more rejections since there is a higher chance of not finding vehicles that can accommodate the size of the group of travellers who have

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

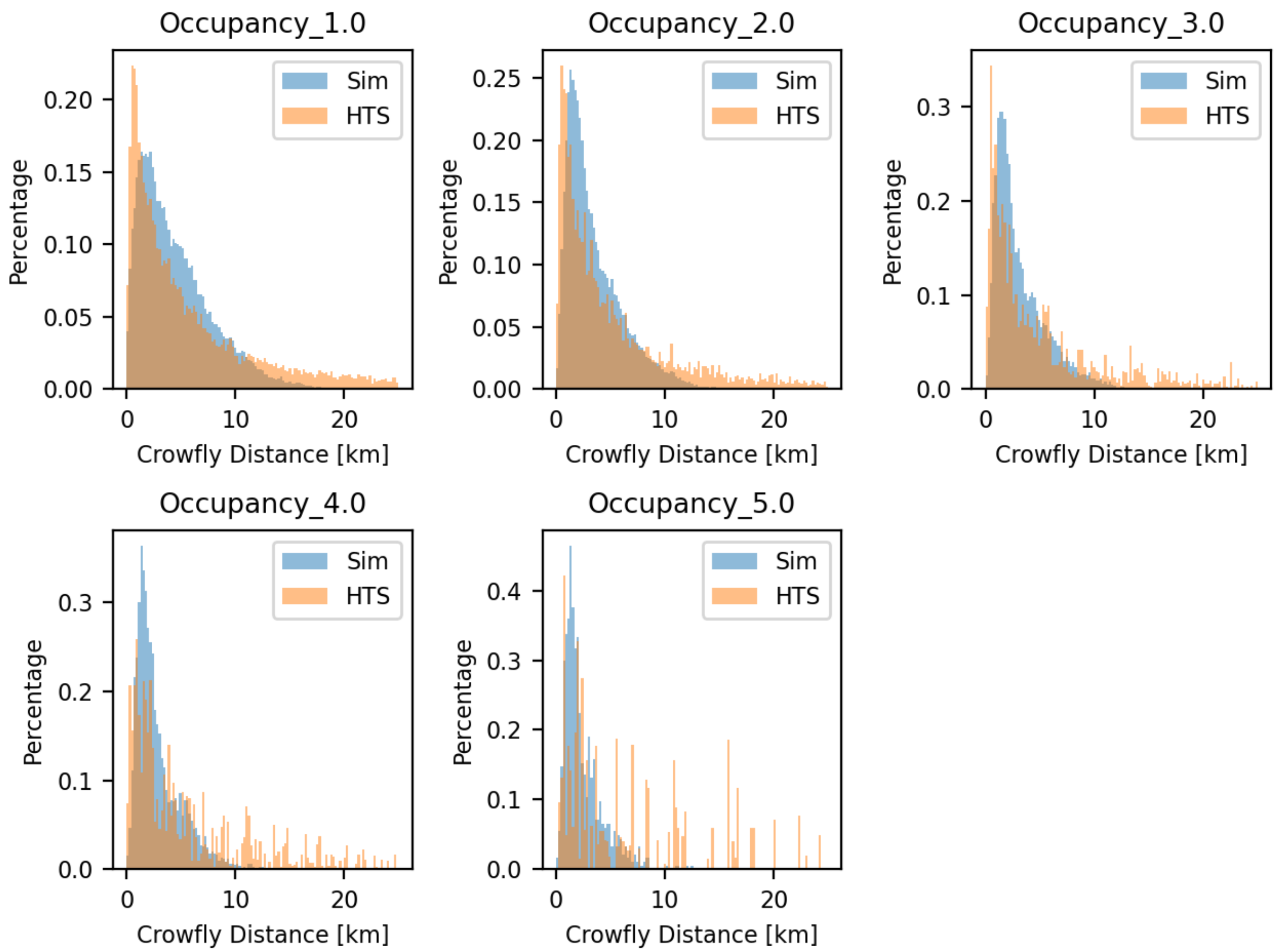

Fig. 5. Distribution of passenger occupancy by distance for a workday (Sim: synthetic population, HTS: Swiss HTS)

TABLE V
RIDESHARING OPERATIONAL PERFORMANCE FOR SCENARIOS WITH FLEET SIZE OF 6 000

	Average Workday		Saturday		Sunday	
	Group	Individual	Group	Individual	Group	Individual
No. Served Requests	331006	428789	371676	521391	243407	349151
Rejections	268	86	1132	234	660	197
Avg Wait time [s]	259.82	320.96	229.55	370.95	173.9	293.80
Avg Detour Factor	1.18	1.24	1.18	1.30	1.15	1.22
Total vehicle driven distance [10^6 km]	1889	2248	1982	2615	1358	1848
Total empty vehicle distance travelled [10^6 km]	117	138	97	111	59	86
Total vehicle occupied distance [10^6 km]	1771	2109	1884	2504	1298	1762
Vehicle occupancy	1.90	1.59	2.12	1.64	2.13	1.56
Avg Travel time [s]	686.26	667.7	655.89	688.25	672.53	677.30

requested a ride. This finding warrants further investigation into different vehicle capacities for group travel, potential tendencies for groups to prefer not sharing with other groups, and how these factors can affect fleet sizing and vehicle operations. However, from this simplistic implementation, one can observe a higher likelihood of rejections when there are fewer vehicles available to accommodate large groups.

In conclusion, the simulation results suggest that ridesharing systems could benefit significantly from including group travel, resulting in greater efficiency and improved outcomes. Incorporating group travellers into the synthetic population demand, and taking account of the travel party size while making requests, allows the dispatching algorithm of the ridesharing system to optimize pooling, reduce detours, and minimize travelled distances, as shown in Figure 6. It is also important to note that it is the proper matching of group travellers in time and space that is essential in achieving optimal results. By ensuring that the requests of group travellers occur simultaneously, even when they are generated independently, one can improve the results, as demonstrated in Figure 7, even if travel party size is not explicitly considered. This is particularly relevant in cases where group trips are generated

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

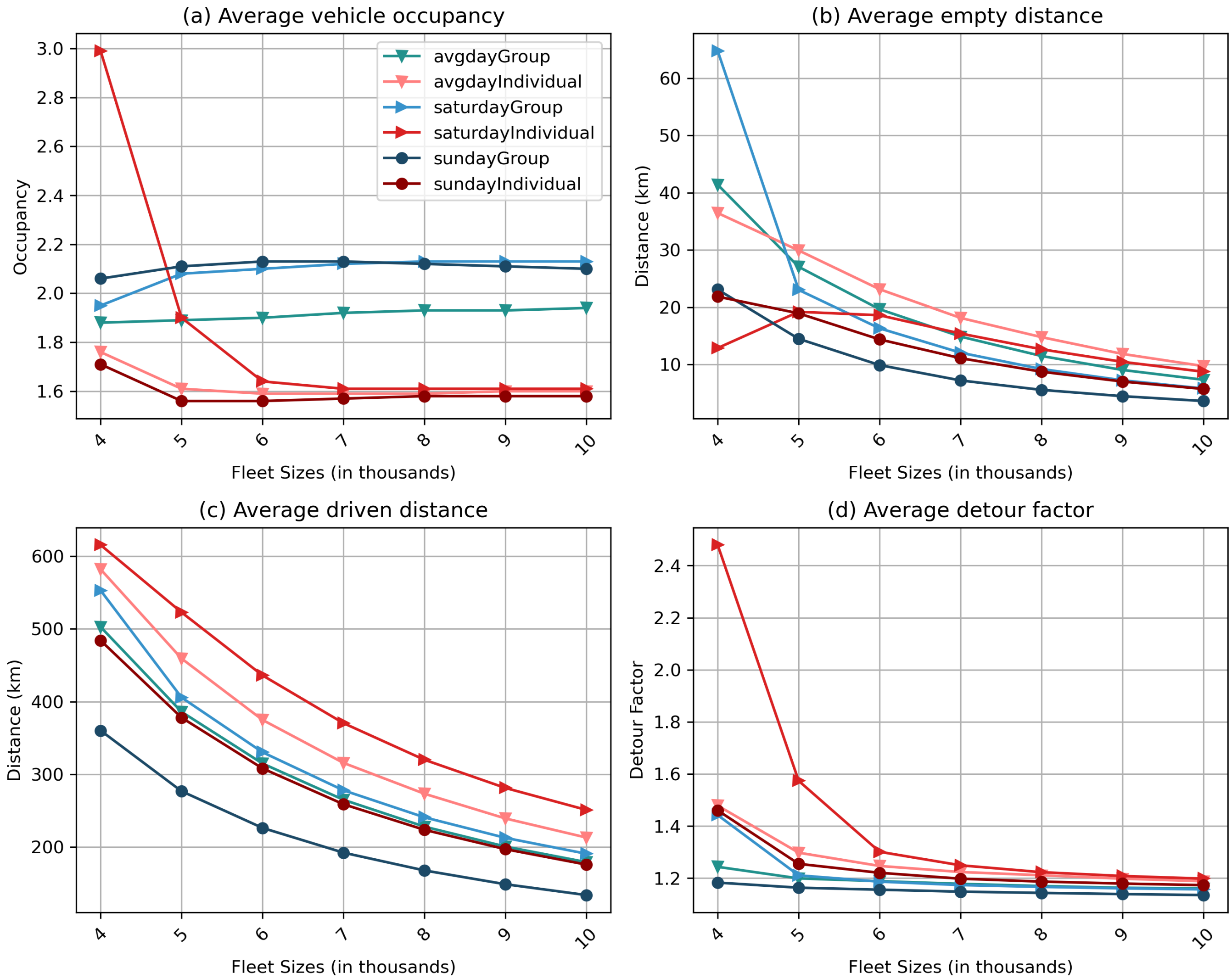

Fig. 6. Operational performance. (blue lines represent group travellers, and red lines represent the individual - single travellers)

independently, without joint modelling of household and non-household travel.

V. Conclusion

The aim of this study is to model group travel in agent-based models of on-demand mobility, which can inform policy decisions related to travel behaviour and the operational performance of ridesharing services. To achieve this goal, the study developed an algorithm to simulate group travel based on the Swiss HTS and used synthetic travel demand data from Switzerland to identify potential group travellers. The algorithm matches car drivers and passengers based on vehicle capacity constraints and overall vehicle occupancy distribution.

Moreover, the study developed an agent-based simulation model to simulate ridesharing and analyze the impact of group travel on ridesharing policies, using SAVs as a case study. The results of the study can provide valuable insights into the potential benefits of promoting group travel in ridesharing services. They can aid in the design of more efficient and effective ridesharing policies.

After applying time and distance constraints, the matching algorithm successfully matched about 78% of passengers to drivers, with an average difference of under 4 km and 10 minutes for the workday, Saturday, and Sunday trips. However, the matching results slightly overestimated the number of drivers with 1-2 passengers by 3-6% for different days. The distance distribution diverged for occupancies of 4 or more, as there were fewer observations to match at higher occupancies and longer distances.

Despite these results, the algorithm shows promise for simulating group travel at the scope of this study, but further refinement may be needed to improve accuracy. For example, integrating the matching approach prior to generating the synthetic population or even controlling for or incorporating other trip attributes, sociodemographic attributes or household attributes in the matching algorithm could improve accuracy. Future work can also focus on developing more accurate methods for matching car drivers and passengers, such as using

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3328492

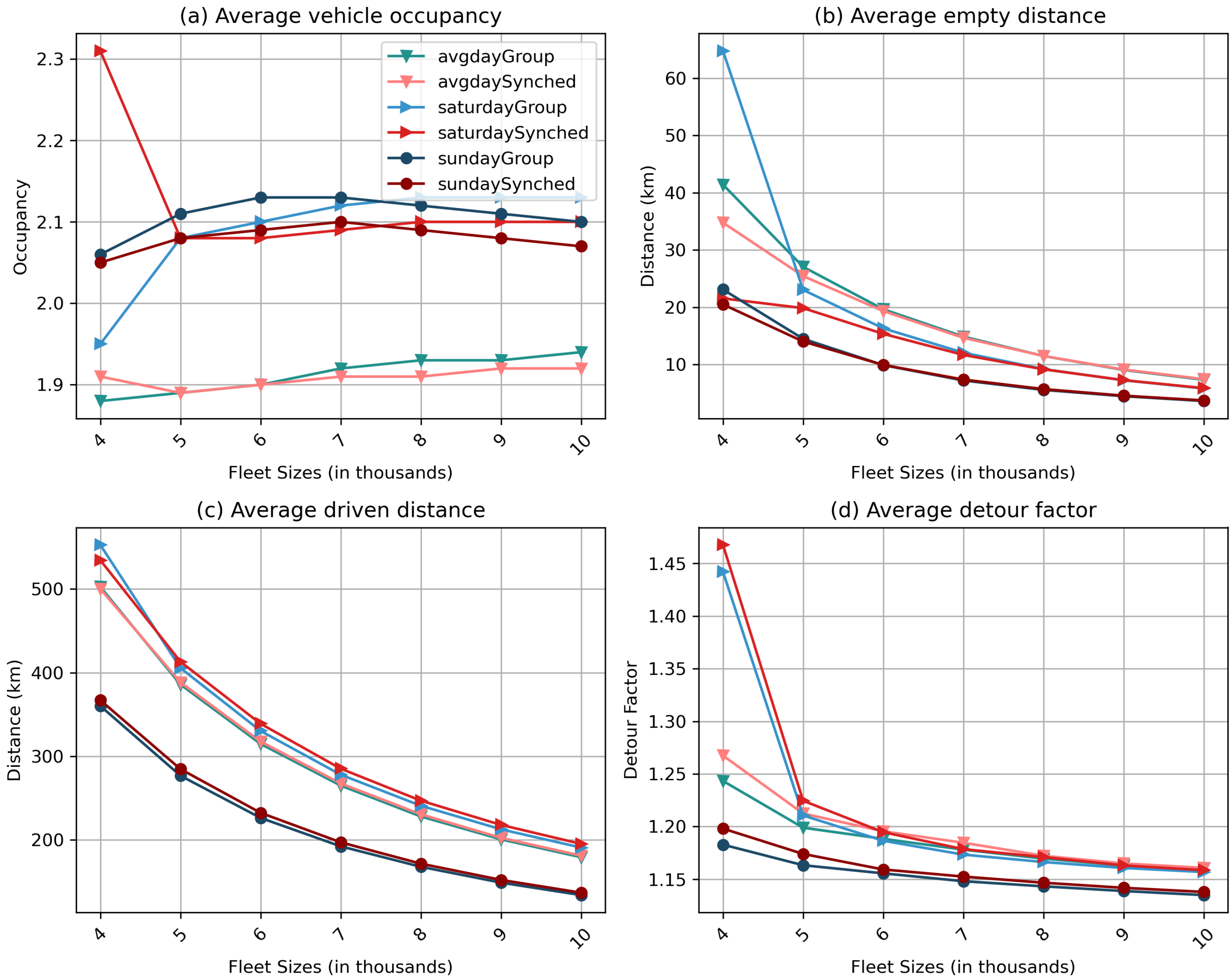

Fig. 7. Operational performance with matched passengers having the same trip characteristics. (blue lines represent group travellers, and red lines represent the synched - single travellers)

machine learning algorithms to predict group travel behaviour. Additionally, a future study can include group travellers who use other mobility modes, such as public transport and active modes, with consideration of how they can be represented as group travellers in an agent-based model.

The simulation results highlight the importance of considering group travel when modelling on-demand mobility services, and policies should factor in the needs and preferences of group travellers. For instance, the relative difference between scenarios where group travel is accounted for when requesting an SAV service and scenarios where it is not can be up to 100% for average vehicle occupancy, which is a significant measure for policies aimed at promoting pooling. However, even if group travellers are only properly matched in time and space within the synthetic population's travel demand, without considering the travel party size in the request, the dispatching algorithm of the ridesharing system can perform better and reduce the relative differences to around 15%. Therefore, it's crucial to ensure that the synthetic population's travel demand includes appropriately matched group travellers in time and space, even when group travellers are generated independently and there is no joint modelling of household and non-household group trips.

This study is an essential first step towards opening up discussions and research on incorporating group travel behaviour in ridesharing and SAV modelling. It provides several opportunities for future research and is especially valuable in regions where data on joint trip creation in travel demand generation is not readily available. The proposed methodology for representing group travellers in simulations offers researchers a useful tool for future studies.

However, this study's scope is limited as it only focuses on car drivers and passengers and does not consider other forms of group travel that may arise from other modes, most notably public transport. This limitation also extends to our simulation model, which doesn't fully capture the complex behaviour of the agents, nor does it account for the competitive influence of alternative transport modes on the on-demand service.

Recognizing these constraints, the inclusion of competing modes in the analysis would offer a richer understanding of the on-demand system's operational performance, as well as provide a more comprehensive depiction of the decisions made by group travellers regarding their mode of transport, and would make the proposed model more beneficial for practitioners. Additionally, the assumption that the trips have approximately the same departure time, point of origin, and destination may not always hold true in reality. Furthermore, with the post-COVID world and the potential increase in work-from-home policies, on-demand mobility modes may see more variability in weekday usage intensity, necessitating separate modelling of different days.

Therefore while this study provides a foundation, future research should aim to integrate these components for a more holistic representation of group travel dynamics, identifying other forms of group travel and relaxing or increasing some of the assumptions made, thus bolstering the study's practical relevance. Furthermore, there is a need to replicate this study in different regions to standardize results and match parameters, enabling comparison across regions.

In conclusion, this study's findings offer important insights into designing and implementing on-demand mobility services, with valuable implications for researchers and policymakers. Therefore, policymakers need to consider the findings in this study, which emphasize the role of group travel dynamics in on-demand mobility services. For example, the observed rise in average vehicle occupancy when accounting for group travel suggests that planners should actively integrate group travel considerations into policy scenarios promoting pooling. These policies could be pricing incentives for group travellers, discounts for shared rides, or dynamic pricing strategies that factor in the party size of travellers. Particularly, as emerging mobility, such as autonomous vehicles, would disrupt how people currently travel together, policymakers should anticipate and accommodate new dynamics that may form for group travel. For example, situations where a person previously required an escort might change, or friends may find it much more convenient to use an SAV. Overall, proactively incorporating group travel considerations into on-demand mobility policy planning will be necessary from now on.

APPENDIX A
COMPARING A STANDARD MILP MODEL WITH THE GREEDY HEURISTIC ALGORITHM

The formulated MILP algorithm was implemented to compare with the heuristic algorithm to check that the heuristic algorithm approximates the results if a standard optimisation model is used. Gurobi was selected as the optimisation solver of choice and Python Gurobi implementation is used for the problem. The decision variables x_{ij} and $Y_{j,l}$ are defined. For which x_{ij} is a binary variable 1 if passenger i is assigned to the driver j and 0 otherwise. $Y_{j,l}$ is a binary variable which is 1 if driver j has occupancy level l, and 0 otherwise.The objective function is defined to minimize the total cost. The following constraints are added:

- Each passenger is assigned to exactly one driver.

TABLE VI
MATCHING CONFIGURATION VALUES

No. Drivers	50
No. Passengers	80
Occupancy distribution	
0	14
1	13
2	9
3	8
4	4
5	2
No. passenger seats	5
Cost weighting value	0.2

- Each driver is assigned at most k passengers.
- Each driver has only one occupancy level.
- The total number of passengers in a vehicle equals its occupancy level.
- The occupancy distribution is close to the desired distribution. The total number of vehicles at each occupancy level is within ϵ of the desired number Z_l.

The constraints guarantee that the solutions will be feasible for the problem, i.e., each passenger is assigned to precisely one driver, each driver is assigned at most k passengers, each driver has exactly one occupancy level, the total number of passengers in a vehicle equals its occupancy level, and the occupancy distribution is close to the desired distribution.

The costs c_{ij}, a 2D matrix that holds the cost for each driver-passenger pair, is generated using the same cost definition as in the heuristic algorithm. See Equation 1 in the Methods section. Z_l, a list that holds the desired vehicle occupancy distribution, is generated. Table VI presents the following sample data generated for the comparison. Random origin and destination coordinates are generated for drivers and passengers using a Numpy uniform random number generator. This is the same for departure times.

The results of the heuristic algorithm and the Gurobi model are compared using key metrics: total cost, the number of unmatched passengers, and the distribution of vehicle occupancy. These results are concisely presented in Table VII. A quick look at the results reveals some immediate differences. Though the heuristic algorithm gives a lower total cost of 32.45 compared to the Gurobi model's 88.53, the Gurobi model does better than the heuristic in specific essential performance metrics. For example, it achieves perfect passenger matching, leaving no passenger unattended, unlike the heuristic algorithm, which leaves three passengers unmatched. Also, the Gurobi adhered better to the vehicle occupancy constraints. While the heuristic algorithm offers a similar occupancy distribution, it's worth noting that it does have more vehicles with zero occupancy and fewer vehicles with full occupancy compared to the Gurobi model. This raises questions about the heuristic algorithm's optimality in fully utilizing vehicle capacity. Still, the computational speed of the heuristic algorithm provides a huge advantage which becomes increasingly necessary as the problem size scales up. Furthermore, the heuristic algorithm is more stringent in minimizing the cost. Therefore while the Gurobi model is superior in performance, especially for passenger matching and adhering to the vehicle occupancy

TABLE VII
MATCHING RESULTS

	Gurobi	Heuristic	Input
Unmatched passengers	0	3	-
Total matching cost	88.53	32.45	-
Run time [s]	2.3	0.5	-
Occupancy distribution			
0	14	17	14
1	13	11	13
2	9	8	9
3	8	8	8
4	5	4	4
5	1	2	2

constraints, the heuristic algorithm still provides an adequate solution, particularly for lower costs and faster run time, making it suitable for large scale problem that has been done in this study.

APPENDIX B
SENSITIVITY ANALYSIS OF COST WEIGHTING VALUES

This section presents the result of the sensitivity analysis carried out to decide the α value that has been used for the cost weighting as described in Section II. Based on the results shown in Table VIII, the weight value, α = 0.2, has been selected in this study. This choice penalizes distance cost more and places greater weight on time cost. This helps to account for possible biases in distances that could have been generated during the synthetic population process. In the synthetic population generation, distance distributions have been used to determine secondary activity locations, possibly introducing bias for leisure trips more likely to involve group travel.

TABLE VIII
SATURDAY - SENSITIVITY ANALYSIS OF COST WEIGHTING VALUES

weight		**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
time [min]	mean	1.68	0.99	0.69	0.50	0.38	0.29	0.22	0.16	0.11
	median	1.15	0.62	0.42	0.30	0.22	0.17	0.12	0.08	0.05
	75-percentile	2.28	1.27	0.83	0.60	0.43	0.32	0.23	0.15	0.08
	max	24.37	24.37	24.37	24.37	24.18	24.18	24.18	24.18	24.18
dist [km]	mean	2.48	2.89	3.25	3.59	3.94	4.32	4.77	5.38	6.43
	median	2.03	2.43	2.75	3.04	3.33	3.66	4.05	4.58	5.51
	75-percentile	3.08	3.62	4.09	4.52	4.96	5.42	5.99	6.76	8.12
	max	33.78	33.78	33.78	33.78	35.14	39.23	39.23	39.23	40.60
cost		17855.63	18923.13	18738.05	17820.68	16350.21	14430.73	12081.42	9275.57	5875.74
unmatched passengers		396	383	370	361	358	354	350	354	358

REFERENCES

[1] S. R. Gehrke, M. P. Huff, and T. G. Reardon, "Social and trip-level predictors of pooled ride-hailing service adoption in the greater boston region," *Case Studies on Transport Policy*, vol. 9, no. 3, pp. 1026–1034, 2021.

[2] K. M. Gurumurthy and K. M. Kockelman, "Analyzing the dynamic ride-sharing potential for shared autonomous vehicle fleets using cellphone data from orlando, florida," *Computers, Environment and Urban Systems*, vol. 71, pp. 177–185, 2018.

[3] BFS, "Mikrozensus Mobilität und Verkehr," https://www.bfs.admin.ch/bfs/de/home/statistiken/mobilitaet-verkehr/erhebungen/mzmv.html, 2015, accessed: 2022-05-03.

[4] Ministry of Ecology, "Enquête nationale transports et déplacements (ENTD) 2008," https://www.statistiques.developpement-durable.gouv.fr/enquete-nationale-transports-et-deplacements-entd-2008, 2018, accessed: 2022-11-03.

[5] T. H. Rashidi and T. T. Koo, "An analysis on travel party composition and expenditure: A discrete-continuous model," *Annals of Tourism Research*, vol. 56, pp. 48–64, 2016.

[6] T. A. Arentze and H. J. Timmermans, "A need-based model of multi-day, multi-person activity generation," *Transportation Research Part B: Methodological*, vol. 43, no. 2, pp. 251–265, 2009.

[7] H. Kang and D. M. Scott, "An integrated spatio-temporal GIS toolkit for exploring intra-household interactions," *Transportation*, vol. 35, pp. 253–268, 2008.

[8] L. Sun, K. W. Axhausen, D.-H. Lee, and X. Huang, "Understanding metropolitan patterns of daily encounters," *Proceedings of the National Academy of Sciences*, vol. 110, no. 34, pp. 13 774–13 779, 2013.

[9] Y. Zhang, K. Martens, and Y. Long, "Revealing group travel behavior patterns with public transit smart card data," *Travel Behaviour and Society*, vol. 10, pp. 42–52, Jan. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214367X16301259

[10] X. Liu, H. Titheridge, X. Yan, R. Wang, W. Tan, D. Chen, and J. Zhang, "A passenger-to-driver matching model for commuter carpooling: Case study and sensitivity analysis," *Transportation Research Part C: Emerging Technologies*, vol. 117, p. 102702, Aug. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X20306173

[11] C. Filippi, G. Guastaroba, L. Peirano, and M. G. Speranza, "Trends in passenger transport optimisation," *International Transactions in Operational Research*, 2023.

[12] Q. Sun, S. Chien, D. Hu, G. Chen, and R.-S. Jiang, "Optimizing multi-terminal customized bus service with mixed fleet," *IEEE Access*, vol. 8, pp. 156 456–156 469, 2020.

[13] X. Chen, Y. Wang, and X. Ma, "Integrated optimization for commuting customized bus stop planning, routing design, and timetable development with passenger spatial-temporal accessibility," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2060–2075, 2021.

[14] W. Zhang, R. He, Y. Chen, M. Gao, and C. Ma, "Research on taxi pricing model and optimization for carpooling detour problem," *Journal of Advanced Transportation*, vol. 2019, 2019.

[15] K. Zhang and Y. M. Nie, "To pool or not to pool: Equilibrium, pricing and regulation," *Transportation Research Part B: Methodological*, vol. 151, pp. 59–90, 2021.

[16] N. Mallig and P. Vortisch, "Modeling Car Passenger Trips in mobiTopp," *Procedia Computer Science*, vol. 52, pp. 938–943, Jan. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915009692

[17] M. Furuhata, M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang, and S. Koenig, "Ridesharing: The state-of-the-art and future directions," *Transportation Research Part B: Methodological*, vol. 57, pp. 28–46, Nov. 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261513001483

[18] A. Horni, K. Nagel, and K. W. Axhausen, *Introducing matsim*. Ubiquity Press, 2016.

[19] J. Bischoff, M. Maciejewski, and K. Nagel, "City-wide shared taxis: A simulation study in berlin," in *2017 IEEE 20th international conference on intelligent transportation systems (ITSC)*. IEEE, 2017, pp. 275–280.

[20] C. Tchervenkov, G. O. Kagho, A. Sallard, S. Hörl, M. Balać, and K. W. Axhausen, "The Switzerland agent-based scenario," *Arbeitsberichte Verkehrs-und Raumplanung*, vol. 1802, 2022.

[21] Y. Räth, M. Balac, S. Hörl, and K. Axhausen, "System level policies for pooled autonomous vehicles," in *100th Meeting of the Transportation Research Board (TRB)*, 2021.

[22] S. Hörl, F. Becker, and K. W. Axhausen, "Simulation of price, customer behaviour and system impact for a cost-covering automated taxi system in zurich," *Transportation Research Part C: Emerging Technologies*, vol. 123, p. 102974, 2021.

[23] G. O. Kagho, J. Meli, D. Walser, and M. Balac, "Effects of population sampling on agent-based transport simulation of on-demand services," *Procedia Computer Science*, vol. 201, pp. 305–312, 2022.