

25th Euro Working Group on Transportation Meeting (EWGT 2023)

# Environmental benefits of taxi ride-sharing in New York City

Marzieh Afsari<sup>a</sup>, Nicola Ippolito<sup>b</sup>, Lory Michelle Bresciani Mistrice<sup>b</sup>, Guido Gentile<sup>c\*</sup>

<sup>a</sup>MSc. Graduated, Sapienza University, Rome 00184, Italy.

<sup>b</sup>Ph.D. student, Sapienza University, Rome 00184, Italy.

<sup>c\*</sup>Department of Civil Engineering, Sapienza University, Rome 00184, Italy.

---

## Abstract

This paper explores the potential environmental benefits of ride-sharing in New York City by finding a balance between supply and demand; for the supply side considering factors such as distance and emissions while taking into account demand-side factors like waiting time and deviation from ride time (DRT). A heuristic algorithm called ADARTW (Advanced Dial-A-Ride problems with Time Windows) is used for a time-constrained version of the Dial-A-Ride problem. The algorithm creates a "pick-up window" for each request and assigns customers to vehicles by finding feasible customer insertions into the work schedules of vehicles. Furthermore, a cost function is employed to optimize the insertion process to select the best customer insertion within the algorithm. This cost function takes into consideration several key factors. Then employs a nonlinear objective function to guide the insertion process and estimate the potential reduction in the number of vehicles required for transportation. The study reveals that ride-sharing could reduce the number of vehicles by 52% and greenhouse gas emissions by 35% in NYC.

© 2023 The Authors. Published by Elsevier BV.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 25th Euro Working Group on Transportation Meeting (EWGT 2023)

*Keywords:* Ride-sharing, cleaner environment, emission reduction, sustainable transport, efficient mobility.

---

## 1. Introduction

With cities housing more than half of the world's population, urban infrastructure is under increasing strain. Transportation is essential to urban sustainability because it consumes up to 28% of total energy consumption and emits pollution, contributing to global warming and causing over one million yearly deaths. The solution for reducing the environmental effect of transportation is in two levels: vehicle level and system level (Lokhandwala and Cai 2018). Encouraging people to use more sustainable and environmentally friendly transportation systems, which leads to reduce Empty Vehicle kilometer traveled and increase average vehicle occupancy, has a significant effect on air pollution **produced** from the point of view of system solution (Zwick, Kuehnel et al. 2022). However, the rapid application of geo-locating smartphones and the Internet's advent, along with its use of big data, has made the availability of a variety of transportation options by considering different factors, including cost, travel time, flexibility, convenience, reliability, and security perception (Furuhata, Dessouky et al. 2013).

The emergence of the sharing economy concept has positioned "Ride-Sharing" as a compelling solution that

addresses both sustainable transportation and the evolving needs of users. The sharing economy is an economic model frequently defined as a peer-to-peer activity of acquiring, providing, or sharing access to goods and services through a community-based online platform (Rasheed Gaber and Elsamadicy 2021). Ride-sharing is a service in which a vehicle is shared with other riders along a fully or partially overlapping route (Lokhandwala and Cai 2018). The service falls between public and private transportation regarding privacy, comfort, journey times, or service quality (Zwick, Kuehnel et al. 2022). On the one hand, taxi ride-sharing must be able to compete with the most significant advantage of a private car, namely, immediate access to door-to-door transportation for users (Dai, Ding et al. 2022). On the other hand, one of the most important aspects of a ride-sharing platform is how it influences or coexists with traditional transportation systems (Pan 2021). Travel time is identified as the most critical factor influencing willingness to use the share option, followed by availability. On the other hand, users expect a specific discount to prefer ride-sharing (Zwick, Kuehnel et al. 2022). This also provides access to high occupancy vehicle (HOV) lanes and eliminates the need to look for parking. In a case study conducted by (Wang, Zheng et al. 2018) focusing on Singapore, it was observed that introducing ride-sharing services resulted in a significant reduction of nearly 50% in rejected requests for internet taxis. Moreover, ride-sharing platforms catered to 20-25% more taxi requests and effectively mitigated waiting times, particularly during peak hours.

Many introducing researchers focus on developing ride-sharing algorithms to make them more practical; as an example, (Alonso-Mora, Samaranayake et al. 2017) presented a mathematical model for real-time high-capacity ride-sharing that scales to large numbers of travelers and journeys and generates optimal routes based on online demand and vehicle location dynamically. (Liang, Correia et al. 2020) proposed an integer non-linear programming (INLP) model, to investigate the DARP of automated taxis (ATs) under dynamic travel times generated by the ATs themselves. The model's primary goal is to maximize the total daily profit of such a system by deciding on each AT's routing based on real-time information. (Meshkani and Farooq 2022) another research proposed a novel Graph-based Many-to-One Ride-Matching (GMOMatch) algorithm for traffic congestion's dynamic many-to-one matching problem. This two-step iterative method provides high service quality while being computationally efficient.

Some studies are about the impact of ride-sharing on emissions. For example, one research group investigated the environmental benefits of ride-sharing using shared taxis in Beijing as a case study (Cai, Wang et al. 2019), and another did this investigation in Dublin (Guo and Xu 2020). Overall, ride-sharing has the potential to reduce energy consumption and greenhouse gas emissions from the transportation sector and provide a more sustainable and accessible transportation option for urban residents. However, further research is needed to address ride-sharing's challenges and optimize its benefits.

This paper aims to evaluate the environmental benefits of shared taxis by analyzing real-world trip data and comparing the emissions reduction of  $N_2O$ ,  $CH_4$ , and  $CO_2$  resulting from ride-sharing scenarios in New York City. This research assesses these environmental advantages by utilizing real-world trip origins and destinations derived from taxi trajectory data in New York City (NYC). Although the NYC dataset serves as the basis for this study, the framework and methodologies employed can be applied to any urban area. A comprehensive description of the algorithm used is provided in section 2. Subsequently, the dataset was analyzed to identify the intended target group. The following section defines the optimal sharing scenario based on specific criteria. This scenario is then compared to the no-sharing scenario to gain insights into the environmental impact of ride-sharing.

## 2. Methodology

### 2.1. ADARTW algorithm

This study presents a heuristic algorithm for addressing the time-constrained variant of the advance-request, multi-vehicle, many-to-many Dial-A-Ride problem. The term "many-to-many" indicates that each customer possesses a unique origin and destination, while "advance" request implies that all requests are received well in advance of vehicle dispatching. Customers specify their desired pick-up time (DPT) at the origin point. The problem encompasses two distinct constraints, namely, service quality and operative constraints.

Service quality constraints ensure that the ride time of each customer does not exceed a predefined maximum, determined as a function of the direct origin-to-destination ride time. Moreover, the actual pick-up time for a customer must not deviate from the desired time by more than a specified threshold.

On the other hand, operative constraints impose limitations on the capacity of vehicles, implying that vehicles have finite capacities. Additionally, it is essential to note that vehicles carrying passengers are not permitted to remain idle, as this assumption forms a fundamental basis for the problem's logical structure. To tackle this problem, the proposed heuristic algorithm provides an efficient and practical approach for generating near-optimal solutions within a reasonable computational timeframe. The algorithm considers distance, time, capacity constraints, and individual customer requirements to determine the optimal vehicle assignment and route planning. The algorithm progressively incorporates requests into the vehicle routes through iterative iterations while respecting the defined constraints and optimizing the overall objective.

Furthermore, the algorithm incorporates an improvement phase, where the local search or enhancement heuristics are applied to refine the obtained solution. These techniques may involve optimizing the sequence of customer visits within each route or performing exchanges between vehicles to enhance the solution quality further while adhering to the specified constraints.

The algorithm's outcome is a final solution comprising the routes for each vehicle, providing the order of customer visits, and ensuring compliance with service quality and operative constraints. Notably, the algorithm's specific implementation and effectiveness may require customization and further investigation to account for the unique characteristics and requirements of the time-constrained ADARTW problem.

The ADARTW algorithm constructs vehicle tours by sequentially inserting customers, guided by a nonlinear objective function. The initial step involves establishing a "pick-up window" for each request, as illustrated in Fig 1.

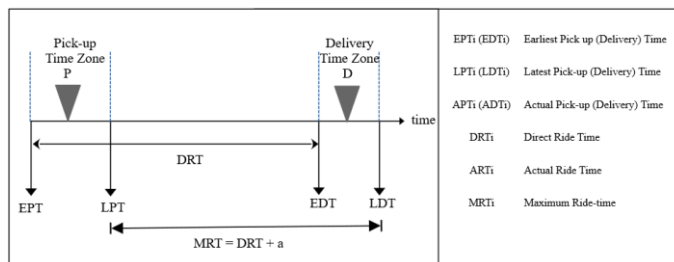


Fig 1. Pickup Time Window.

Subsequently, the assignment of customers to vehicles entails a systematic search for feasible insertions within the work schedules of the vehicles, followed by an optimization phase aimed at identifying the most favorable insertion. In ADARTW, we compute and store four statistics for each stop  $r = 1, \dots, d$ , on each schedule block  $p$ , defined as follows:

$$BUP_r = \text{Min}_{1 \leq t \leq r} AT_t - ET_t, SLACK_p \tag{1}$$

$$BDOWN_r = \text{Min}_{1 \leq t \leq r} LT_t - AT_t \tag{2}$$

$$AUP_r = \text{Min}_{r \leq t \leq d} AT_t - ET_t \tag{3}$$

$$ADOWN_r = \text{Min}_{r \leq t \leq d} LT_t - AT_t, SLACK_{p+1} \tag{4}$$

The variables  $BUP_r$  ( $BDOWN_r$ ) denote the maximum allowable time adjustment for each stop, including stop  $r$ , such that the time-window constraints are upheld when advancing (delaying) the stops. Likewise,  $AUP_r$  ( $ADOWN_r$ ) represents the maximum permissible time adjustment for each stop, including stop  $r$ , to preserve the time-window constraints when advancing (delaying) subsequent stops. The process of inserting pick-up and delivery points within a block is visually depicted in (Fig 2).

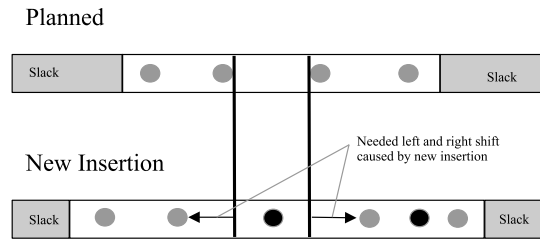


Fig 2. Allowed left and right shifts.

Besides checking for violations of the time-window constraints, it is necessary to check that no maximum-ride-time constraints are violated for the newly inserted customer and for the customers already in the schedule block. This can be done quickly by scanning through the list of these customers and comparing the respective actual ride times and the maximum allowable ride times. Finally, vehicle loads at each stop between the inserted pick-up and delivery of customer  $i$  are checked so that vehicle capacity is not exceeded.

A cost function is used to find the best insertion, a weighted sum of disutility to the customers (due to excess ride times and deviations from the earliest pick-up), and system costs represented by the consumption of available vehicle resources. The disutility to customer  $i$  is given by:

$$DU_i = DUD_i + DUR_i \quad (5)$$

Where  $DUD_i$  is the disutility due to deviation from a desired time, and  $DUR_i$  is the disutility due to excess ride time.

$$DUD_i = C_1 \times x_i + C_2 \times x_i^2, 0 \leq x_i \leq WS_i, x_i = APT_i - DPT_i \quad (6)$$

$$DUR_i = C_3 \times y_i + C_4 \times y_i^2, y_i \geq 0, y_i = ART_i - DRT_i \quad (7)$$

The incremental cost,  $VC_i$  to the system's operator due to inserting customer  $i$ :

$$VC_i = C_5 \times z_i + C_6 \times w_i + U_i \times (C_7 \times z_i + C_8 \times w_i) \quad (8)$$

$z_i$  the additional active vehicle time required to serve the customer  $i$

$w_i$  the change in vehicle slack time due to the insertion

$U_i$  an indicator of system workload

$U_i = (\text{number of customers in interval } T_i) / (\text{number of vehicles available in interval } T_i)$

$$T_i = EPT_i - W_1, EPT_i + W_2 \quad W_1 \text{ and } W_2, \text{ which are externally specified constants} \quad (9)$$

Finally, it is worth mentioning that filters for searching vehicles are considered the capacity.

## 2.2. Inputs and Assumptions

This research does not explicitly consider traffic congestion, waiting at traffic signals, etc., but considering the actual average speed of taxis for simulation helps to consider it implicitly. Furthermore, the capacity of vehicles is assumed to be the same for all vehicles and equal to 4 passengers as a casual vehicle. In addition, the distribution of vehicles in the study area is homogeneous. It is assumed that the shift work duration for all vehicles is equal to 8 hours, and all start before the simulation starts.

### 3. Data Selection and Analyzing

The data sets used in this study were collected by technology providers licensed by the Taxi and Uniform Passenger Improvement Program and provided to the New York City Taxi and Limousine Commission (TLC). The data for the first six months of 2016 of yellow taxis were selected for analysis, this data set contains approximately 72 million trip records, and each trip consists of the Longitude and Latitude of pickup and drop-off points, exact time and date of pickup and drop-off points, number of passengers, and distance.

By analyzing the dataset, it is clear that the demand differences between these six months in terms of trip duration are not so dominant (Fig 3) so the decision is to work only on the one-month dataset; since the results can be used for other months. By comparing the average trip duration of six months and each one, it is deduced that “March” has a minor difference among all. In the next step, the average trip duration of all weekdays in March is compared to each other, and “Wednesday” has the best match. So, the Wednesday 16th of March is chosen as a day in the middle of the month.

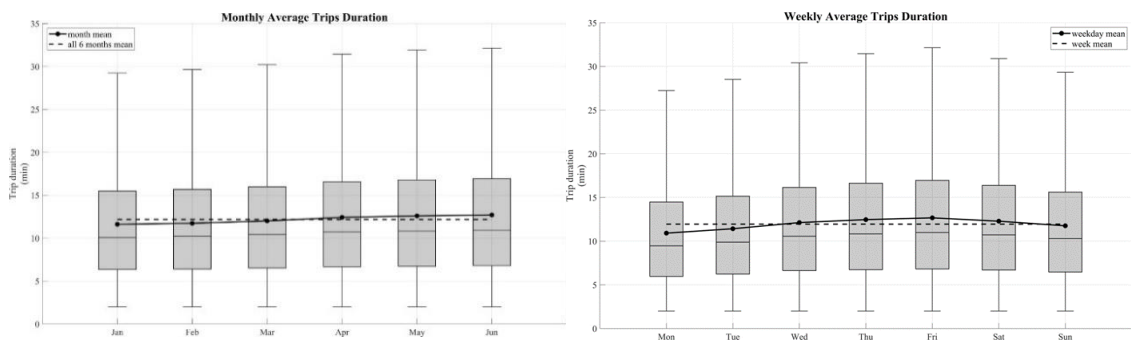


Fig 3. Box plot for the mean trip duration for six months in 2016 and an average of these six months- red line (left hand). Box plot for the mean trip duration for each weekday in March and an average of the week -red line (right hand)

The temporal demand of NYC is shown in Fig 4. This study works on one period: the morning off-peak hour to the morning peak hour from 5 to 8 a.m., demonstrating an increasing demand trend. The reason for choosing this period is that the efficiency of ride-sharing can be studied during the peak and off-peak hours and the transformation during this period can be investigated.

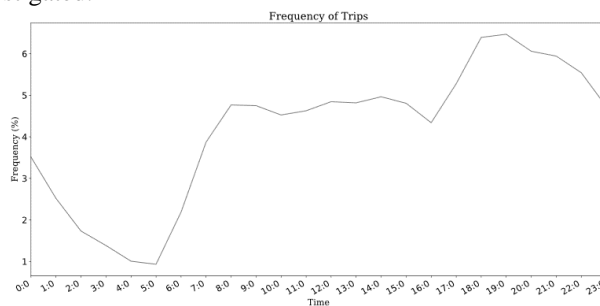


Fig 4. Demand: temporal histogram for pick-ups.

### 4. Simulation and Results

#### 4.1. No-Sharing: Scenario Zero

In scenario zero, there is assumed to be no shared vehicle. First, the no-sharing simulation uses the actual number of vehicles based on the TLC data (Commission 2018). The required vehicles for each category are then determined

through trial and error. No rejection is considered when looking for the required number of vehicles. This means all requests would be accepted by a taxi to proceed. It implies that the number of vehicles is changed at each step until all requests are met, and all vehicles are activated. Table 1 provides a summary of the characteristics of no-sharing simulation. Since the simulation produces a range of outcomes depending on its assumptions and errors, as was mentioned above, these zero scenarios are intended to act as a benchmark against comparing the outcomes of sharing scenarios.

Table 1. No-sharing scenario

Day	Hour	No. of requests	No. of Vehicles
16/03/2016	05:00-06:00	2458	401
	06:00-07:00	7902	1092
	07:00-08:00	15152	1300

#### 4.2. Sharing Scenarios and Results

The waiting time and deviation from ride time are two critical factors for customers to choose the ride-share system. On the other hand, these are fundamental elements that affect the number of vehicles, fuel consumption, and emissions. For considering the effect of waiting time on the whole system, the sharing simulation with an upper bound of 3 minutes for waiting time is done. Then, in the following simulation, this upper boundary increases to 5, 7, and 9 minutes; in all scenarios, the deviation from ride time's upper boundary is assumed constant and equal to 4 minutes. Fig 5 **Error! Reference source not found.** compares the cumulative percentage of waiting time for all four scenarios. In the first one, almost 65% of requests experience no waiting time, while others have a waiting time of less than 30 seconds on average. When the waiting time upper limit expands to 5 minutes, the customer with no waiting experience reduces to 45%. This amount for the 7- and 9-minute boundary is almost the same and equal to 30%, while others experience waiting time averagely equal to 117 and 138 seconds, respectively.

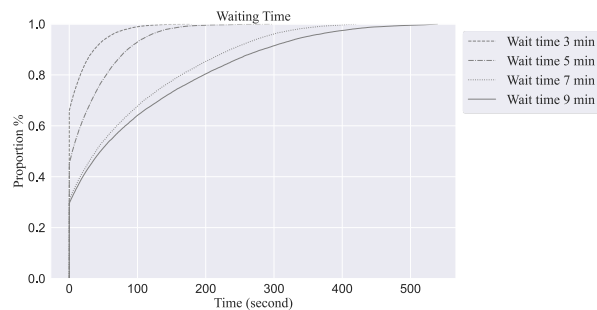


Fig 5. Cumulative percentage for waiting time scenarios.

Table 2 shows how the number of vehicles and total distances changes through all four scenarios. By increasing the waiting time boundary, both factors decreased. However, customers need to wait more, but the system changes are not considered. Increasing waiting time from 3 to 5 minutes decreases the number of vehicles by 22% and distance by 15%. These parameters for changing to 7 minutes are 21% and 12%, respectively. Increasing waiting time by more than 7 minutes has no prominent effect on the system. So, from the point of view of companies, 7 minutes is an acceptable waiting time which increases the profit and reduces the environmental effect of the system.

Table 2. Total distance, No. of vehicles and slack time changes through waiting time scenarios

Waiting Time (minute)	Distance (km)	Different in Distance	No. of Vehicles	Different in No. of Vehicles	Slack Time	Different in Slack
3	74,707	-	1000	-	107	-
5	63,814	15%	780	22%	76.4	29%
7	55,879	25%	620	38%	55	49%
9	55,021	26%	600	40%	52.48	51%

Then, the waiting time is fixed at 7 minutes, and the maximum tolerated deviation from ride time by passengers is changed to evaluate the results. Starting from 2 minutes to increasing it to the point that no increasing changes are seen (2, 4, and 6 minutes). Table 3 show how the number of vehicles and total distances changes through the scenarios. By increasing the DRT from 2 to 4 minutes, the total passed distance reduces 18%, the number of vehicles by 28%, and the slack time by 37%. Then after that, there are no changes.

Table 3. Total distance, No. of vehicles and slack time changes through deviation from ride time scenarios

Deviation from ride time (minute)	Distance (km)	Different in Distance	No. of Vehicles	Different in No. of Vehicles	Slack Time	Different in Slack
2	67,978	-	860	-	87.37	-
4	55,879	18%	620	28%	55.17	37%
6	55,887	18%	620	28%	55.34	37%

Fig 6 **Error! Reference source not found.** demonstrates the cumulative percentage of deviation from ride time for all three scenarios, as it shows the scenarios with DRT equal to 4 and 6 minutes have the same trend, which means the demand of New York City can be served with the maximum DRT 4 minutes, and no more than this is needed. On the other hand, almost 42 percent of customers experience no DRT, while the rest arrive at their destination 36 seconds late on average.

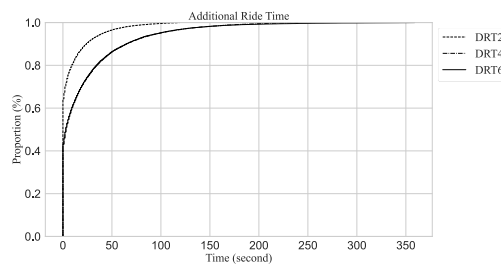


Fig 6. Cumulative percentage for deviation from ride time scenarios.

As a result, the scenario with a waiting time of 5 minutes and a ride time variance of 7 minutes results in the best reduction in the required number of cars and overall distance traveled. It is compared to the zero scenarios of no shared ride to analyze the effects of this sharing situation. Table 4 compares these two and shows that there has been a 35% decrease in overall distance traveled and a 52% decrease in the number of vehicles. We calculated the reduced greenhouse gases due to ride-sharing based on the distance saved. According to the United States Environmental Protection Agency (EPA), the typical passenger vehicle mainly emits carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), and nitrous oxide (N<sub>2</sub>O) (the United States Environmental Protection Agency (EPA) 2022). The

table shows the reduction of emissions for the simulation period in kg of CO<sub>2</sub>-e and percentage. It is shown that emission reduces 35% in comparison to no sharing system, so if widely adopted, ride-sharing may be a more successful policy in reducing total vehicle VMT, energy use, and pollution.

Table 4. Comparison between no sharing and sharing scenarios

KPIs	Scenarios		Differences	Differences (%)
	No Share	Share		
No. of vehicles	1,300	620	680	52 %
Distance (km)	86,622	55,879	30,743	35 %
CH <sub>4</sub> (Kg of CO <sub>2</sub> -e)	40.6	26.2	14.4	35 %
N <sub>2</sub> O (Kg of CO <sub>2</sub> -e)	308	198.7	109.3	35 %
CO <sub>2</sub> ((Kg of CO <sub>2</sub> -e)	16,096.3	10,383	5,712.7	35 %

## 5. Conclusion

This paper presents an ADARTW algorithm for estimating the environmental benefits of ride-sharing in New York City. The algorithm is designed to solve a time-constrained version of the advance-request, multi-vehicle, many-to-many Dial-A-Ride problem, including service quality and operative constraints. The results show that ride-sharing can significantly reduce the number of trips and vehicle kilometer traveled, leading to reductions in greenhouse gas emissions and air pollution. However, the implementation of ride-sharing also requires careful consideration of operational and regulatory issues and the potential impacts on equity and accessibility. This study provides valuable insights for policymakers and transportation planners seeking to promote sustainable and efficient urban mobility.

## References

- Alonso-Mora, J., S. Samaranyake, A. Wallar, E. Frazzoli and D. Rus (2017). "On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment." *Proceedings of the National Academy of Sciences* **114**(3): 462-467.
- Cai, H., X. Wang, P. Adriaens and M. Xu (2019). "Environmental benefits of taxi ride sharing in Beijing." *Energy* **174**: 503-508.
- Commission, T. T. L. (2018). NYC Taxicab Factbook. Tech. Rep.
- Dai, R., C. Ding, J. Gao, X. Wu and B. Yu (2022). "Optimization and evaluation for autonomous taxi ride-sharing schedule and depot location from the perspective of energy consumption." *Applied Energy* **308**: 118388.
- Furuhata, M., M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang and S. Koenig (2013). "Ridesharing: The state-of-the-art and future directions." *Transportation Research Part B: Methodological* **57**: 28-46.
- Guo, G. and Y. Xu (2020). "A Deep Reinforcement Learning Approach to Ride-sharing Vehicles Dispatching in Autonomous Mobility-on-Demand Systems." *IEEE Intelligent Transportation Systems Magazine* **PP**.
- Liang, X., G. H. d. A. Correia, K. An and B. van Arem (2020). "Automated taxis' dial-a-ride problem with ride-sharing considering congestion-based dynamic travel times." *Transportation Research Part C: Emerging Technologies* **112**: 260-281.
- Lokhandwala, M. and H. Cai (2018). "Dynamic ride sharing using traditional taxis and shared autonomous taxis: A case study of NYC." *Transportation Research Part C: Emerging Technologies* **97**: 45-60.
- Meshkani, S. M. and B. Farooq (2022). "A generalized ride-matching approach for sustainable shared mobility." *Sustainable Cities and Society* **76**: 103383.
- Pan, Y. (2021). "How Ride-Sharing Is Shaping Public Transit System: A Counterfactual Estimator Approach." *Production and Operation Management* **31**(3): 906-927.
- Rasheed Gaber, H. and A. M. Elsamadicy (2021). "What drives customers to continue using ride-sharing apps during the COVID-19 pandemic? The case of Uber in Egypt." *Cogent Business & Management* **8**(1): 1944009.
- the United States Environmental Protection Agency (EPA) (2022). GHG Emission Factors Hub.
- Wang, Y., B. Zheng and E.-P. Lim (2018). "Understanding the effects of taxi ride-sharing — A case study of Singapore." *Computers, Environment and Urban Systems* **69**: 124-132.
- Zwick, F., N. Kuehnel and K. Axhausen (2022). *Review on theoretical assessments and practical implementations of ride-pooling*.