Transportation Research Part A 192 (2025) 104360

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

The potential of governmental regulation on shared mobility-on-demand systems

Fabian Anzenhofer [a,*], Simon Schmidbaur [b], Robert Klein [a], Claudius Steinhardt [b]

[a] Chair of Analytics & Optimization, University of Augsburg, Universitätsstraße 16, D-86159 Augsburg, Germany
[b] Chair of Business Analytics & Management Science, University of the Bundeswehr Munich, Werner-Heisenberg-Weg 39, D-85577 Neubiberg, Germany

ARTICLE INFO

Keywords:
Shared mobility-on-demand
Selective dial-a-ride problem
Optimization
Feasibility analysis
Impact analysis
Regulation

ABSTRACT

Shared Mobility-on-Demand services have the potential to contribute to a more socially and environmentally sustainable mobility provision. However, this potential may not be fully exploited due to possible conflicts with the objectives of the service providers. Thus, political discourses address regulatory instruments to influence providers' operational planning. This paper analyzes the effects of two currently debated instruments, i.e., the introduction of a minimum pooling rate and a minimum spatial acceptance rate. Our analysis is based on mathematical optimization models that we formulate as generalizations of the selective dial-a-ride problem. More precisely, the problem is first captured by a single-period model formulation and then generalized to a multi-period horizon to implement different regulatory strategies. In a comprehensive computational study, we solve the regulated model formulations to optimality both for artificial and real-world data provided by our industry partner FLEXIBUS. We evaluate different levels of regulation for both instruments regarding their feasibility and their impact on the Shared Mobility-on-Demand system, and thereby discuss key factors, specific issues, and recommendations with regards to the practical application of regulatory instruments for public authorities that can be derived from the results of our study. Overall, our analysis recommends a multi-period application for both regulatory instruments. In this context, a moderate minimum pooling rate can enhance environmental sustainability, though it may create trade-offs with social or service provider objectives. In contrast, a minimum spatial acceptance rate shows minimal conflicts with other objectives, but requires an effective spatial partitioning approach and is sensitive to the provider's supply–demand ratio.

1. Introduction

Recent advances in communication technology have driven the popularity of platform-based Mobility-on-Demand (MOD) services, a transportation concept that enables customers to access flexible, demand-oriented transportation via mobile applications (Curtis et al., 2019; Shaheen and Cohen, 2020). As customer behavior and expectations change, MOD emerges as a promising approach to meet customers' growing desire for fast and convenient mobility options (Liobikienė and Miceikienė, 2022). Shared MOD (SMOD) enables sharing of resources, in our case the sharing of vehicles in the form of ridepooling. In contrast to traditional taxi services, SMOD

* Corresponding author.
E-mail addresses: fabian.anzenhofer@uni-a.de (F. Anzenhofer), simon.schmidbaur@unibw.de (S. Schmidbaur), robert.klein@uni-a.de (R. Klein), claudius.steinhardt@unibw.de (C. Steinhardt).

https://doi.org/10.1016/j.tra.2024.104360

Available online 18 December 2024

providers can bundle multiple unrelated requests in one vehicle at the same time. In this way, customers can benefit from the advantage of lower fares than those charged by taxis as well as the convenience of more demand-orientation compared to traditional scheduled public transport, such as line buses (Bösch et al., 2018).

SMOD is one of the most prospective approaches to drive the transition to sustainable mobility, as highlighted by several examples from the literature. The ability to consolidate demand can lead to increased routing efficiency, which has environmental benefits in terms of reduced greenhouse gas emissions and economic benefits in terms of saved routing costs (Cai et al., 2019; Ma and Koutsopoulos, 2022; Martinez and Viegas, 2017; Yu et al., 2017). In addition, SMOD can contribute to more socially sustainable public transport by improving accessibility, e.g., for peripheral regions (Shaheen and Cohen, 2020).

However, several studies show downsides of SMOD systems. For example, if operated inefficiently, an SMOD system may generate more vehicle kilometers traveled (VKT) than private rides due to deadheading, i.e., vehicle rides without passengers, resulting in more congestion and ultimately more greenhouse gas emissions (Engelhardt et al., 2019; Ennen and Heilker, 2020; Erhardt et al., 2019; Qiu et al., 2018). Furthermore, while SMOD systems have the potential to increase accessibility of public transport, some literature finds disparities in acceptance rates that arise naturally based on characteristics that favor certain requests over others, such as unfavorable origin–destination-(OD-) pairs (Camarero and Oliva, 2019; Sörensen et al., 2021; Zwick and Axhausen, 2022). One approach to address these downsides is to implement clear and concise legal frameworks, using regulatory instruments to influence operational planning towards desirable directions (Shaheen et al., 2022), e.g., reducing total emissions for environmental sustainability or improving accessibility in peripheral regions for social sustainability.

One example of regulatory instruments already in place can be found in Germany. The Passenger Transport Act (Deutscher Bundestag, 2021) proposes regulatory instruments for the operational planning of SMOD systems, e.g., a minimum pooling rate for environmental sustainability. However, despite the legislation being justified qualitatively by the aim of increasing sustainability, there is little quantitative work to evaluate its potential.

In this regard, with this paper, we make the following contributions:

- To the best of our knowledge, we are the first to consider an external third-party intervention in the operational planning problem of an SMOD provider, such as regulatory instruments proposed by public authorities. For this purpose, we methodologically capture the operational planning problem, which results for SMOD, as a dial-a-ride problem (DARP) which has been studied intensively in recent years (Cordeau and Laporte, 2007; Ho et al., 2018; Molenbruch et al., 2017). Reflecting the demand management aspect, the so-called selective DARP (SDARP) allows the rejection of requests that are not favorable according to the respective planning objective.
- Building on this foundation, we introduce the SDARP-REG, which further generalizes the SDARP by including regulatory instruments inspired by recent German legislation (Deutscher Bundestag, 2021). More precisely, on the one hand, we define a minimum pooling rate (MPR), which is designed to ensure a minimum degree of consolidation of customer requests to contribute to environmental sustainability by reducing VKT. On the other hand, we define a minimum spatial acceptance rate (MSAR) to make SMOD operations more socially sustainable. By implementing MSAR for predefined subsets of requests with specific desired origin and destination, this regulatory instrument is designed to ensure a minimum level of accessibility for customers regardless of their desired OD-pair. From a methodological point of view, we base the formulations of the SDARP-REG on a recent location-augmented-event-based formulation by Gaul et al. (2025), which is currently one of the most efficient modeling approaches to solve DARPs to optimality.
- Further, we generalize the SDARP-REG to a multi-period horizon (M−SDARP−REG). By considering multiple periods, we can enforce regulatory instruments with more flexibility, i.e., minimum rates must be met on average over a given set of periods, while they can be violated in single periods. With the M−SDARP−REG, we also provide a generic methodological contribution as the M−SDARP−REG's model size can be increased by more periods without suffering from an exploding graph size for larger instances. This is due to the fact that the graphs of the different periods are independent and only connected via the regulatory instruments (see Braekers and Kovacs (2016) for a similar modeling approach enforcing driver consistency over a multi-period time horizon).
- Based on the new mathematical models, we perform an extensive computational study comprising a feasibility analysis and an impact analysis of both regulatory instruments on artificial data from the literature and on real-world data from our German industry partner FLEXIBUS. This study aims to evaluate the proposed regulatory instruments and thereby identify key factors, specific issues, and recommendations for public authorities that can be derived from the numerical results with regard to the practical application of regulation. For the feasibility analysis, we investigate how sensible instances react to the respective regulatory instruments, i.e., the restrictiveness of MPR and MSAR, by determining the maximum applicable level of regulation that providers are able to comply with, meaning that there is a feasible solution to the SDARP-REG or M−SDARP−REG. For the impact analysis, we define metrics that represent goals from a provider, environmental, and social perspective to reflect the multidimensional nature of the problem.
- As a side contribution from the computational study, we introduce distance savings per passenger as a novel performance metric that has not received much attention in the literature, yet. This metric provides a hypothetical quantification of the direct environmental impact of SMOD compared to individual motorized transport. By measuring the difference between the desired OD-pair distances weighted by the number of passengers and the actual vehicle kilometers traveled, it yields a proxy for SMODs' potential to reduce their ecological impact.

Please note that, in practice, an SMOD provider operates under dynamic and stochastic conditions, where requests arrive probabilistically throughout the day. However, due to the curses of dimensionality, there is no adequate solution approach that optimally

solves the respective stochastic-dynamic problem (Powell, 2011). Thus, throughout this work, we focus on the static-deterministic problem, i.e., we formulate a corresponding mixed integer linear program (MILP), capture the regulatory instruments as constraints, and solve it to optimality. By using an exact solution approach, we can isolate and quantify the direct effects of regulatory instruments under ideal conditions, i.e., we can accurately evaluate the MPR and MSAR in single-period and multi-period model variants. The results can reliably be interpreted and compared, while heuristic approaches with potentially suboptimal solutions could lead to misleading interpretations with regard to the regulations' influence.

The paper is structured as follows: In Section 2, we provide an overview of the relevant literature, focusing on existing mathematical models for SDARPs. The problem description and the formulation of the new model variants, i.e., SDARP-REG and M−SDARP−REG, are presented in Section 3. Then, in Section 4, we show the numerical results of our computational study for artificial and real-world data, focusing mainly on the presentation and technical evaluation of each regulatory instrument. In Section 5, we comprehensively discuss key factors, issues for application, as well as recommendations to public authorithies. Finally, in Section 6, we conclude the findings and discuss possible future research areas.

2. Literature review

In this section, we provide an overview of the existing literature. We focus on the mathematical modeling of SDARPs. Thus, we do not include any empirical analysis of SMOD systems (e.g., Chen et al., 2017) or agent-based simulation (e.g., Dandl et al., 2021). First, we address relevant literature on different modeling approaches for the DARP (Section 2.1). We then discuss the generalization to the SDARP from a modeling perspective, in particular regarding alternative objective functions (Section 2.2). Finally, we position our work within these publications (Section 2.3).

2.1. Modeling the Dial-a-Ride problem

In general, vehicle routing problems (VRPs) are classified according to the time at which information is provided and the certainty of the information (Pillac et al., 2013). In the static case, all information is known in advance, while in the dynamic case, information is revealed over time. If the information is known with certainty, the problem is deterministic, otherwise it is stochastic. In our review, we exclude dynamic (e.g., Hungerländer et al., 2021) and stochastic (e.g., Anzenhofer et al., 2024) problems and focus on static-deterministic MILP models for the reasons listed in Section 1. For a comprehensive overview of DARPs, we refer to Molenbruch et al. (2017), Ho et al. (2018), and Cordeau and Laporte (2007). Furthermore, we focus on problems related to passenger transportation and therefore exclude pickup and delivery problems with time windows (e.g., Berbeglia et al., 2007).

Cordeau (2006) formulates a three-index MILP for the DARP, which is considered a seminal formulation in the literature (Ho et al., 2018). The underlying graph is complete and directed, and reflects a location-based formulation (Gaul et al., 2025), i.e., nodes correspond to the depot and customers' desired origin and destination, while arcs represent the direct rides that vehicles travel between pairs of nodes. The objective function minimizes the total routing costs, while all requests must be served and all constraints regarding time, capacity and routing need to be satisfied. Cordeau (2006) also provides artificial data, which has been used extensively to date (Ho et al., 2018; Molenbruch et al., 2017). The DARP model of Cordeau (2006) is the basis for several other location-based MILP formulations: Ropke et al. (2007) adapt the model and formulate a tighter two-index MILP. For extended artificial data from Cordeau (2006), they show faster solution times compared to the three-index MILP. Braekers and Kovacs (2016) generalize the models by Cordeau (2006) and Ropke et al. (2007) to multi-period DARPs by including an additional index with the aim of enforcing driver consistency over a multi-period time horizon. More recently, a novel modeling approach is introduced by Rist and Forbes (2021), which uses route fragments, i.e., certain fixed parts of a route plan. Their approach uses binary variables to decide which route fragments to utilize. Their study shows a promising reduction in computation time for artificial data from Cordeau (2006).

Gaul et al. (2022) introduce a new model for the DARP, which relies on an event-based graph. Their computational study shows that event-based formulations substantially reduce the solution time on artificial data from Ropke et al. (2007) compared to existing location-based formulations. In an event-based graph, nodes do not represent locations, but events, i.e., they describe a state of a vehicle with information about its current location and all customers on board. Arcs denote the transition between these events. The event-based graph structure has the advantage of a reduced number of constraints compared to the location-based MILP formulation due to its implicit consideration of pairing, precedence, and capacity constraints. In more recent work, Gaul et al. (2025) propose a location-augmented-event-based MILP, which exploits the advantages of both location-based and event-based MILPs, leading to a simplification of the time constraints, and thus, further improvement in solution times for artificial data.

2.2. Objectives in Selective Dial-a-Ride problems

In practice, an SMOD provider does not necessarily serve all requests that arrive during a service day, as required in standard DARPs. It is common for an SMOD provider to reject requests due to infeasibility or to apply some sort of demand management, i.e., to decide whether to accept or reject requests based on the respective objective, e.g., maximization of profit (Cauchi and Scerri, 2020; Hosni et al., 2014; Parragh et al., 2015) or maximization of the number of served requests (Riedler and Raidl, 2018). Therefore, we consider the SDARP as a generalization of the DARP, in which a selected subset of requests is served. For example, Riedler and Raidl (2018) test exact solution approaches on modified artificial data from Ropke et al. (2007), Berbeglia et al. (2011), as well as on their own artificial data based on Cordeau (2006) with an increased number of requests. As already stated, an SMOD provider usually operates in a multidimensional environment, so authors may define a multi-objective function including complementary or conflicting

objectives for SDARPs. Some authors propose weighted objective functions integrating criteria such as routing costs, additional travel time due to detours, or penalties on rejected requests (Gaul et al., 2022; Kirchler and Wolfler Calvo, 2013; Reinhardt et al., 2013; Wolfler Calvo and Colorni, 2007; Zhang et al., 2015). For example, Gaul et al. (2022) found in their computational study on real-world data that it may be worth finding a reasonable tradeoff between service-related metrics and routing costs by adjusting the weights in the objective function.

2.3. Positioning of our work

To the best of our knowledge, we are the first to consider regulatory instruments in the context of SDARPs. In contrast to the existing multi-objective considerations discussed above, our work focuses on a completely different perspective, which is in line with a common real-world situation: The models we consider include the maximization of the number of served passengers as the SMOD provider's main objective, while an external party implements regulatory tools that are considered by constraints, e.g., by MPR or MSAR. This way, the decisions of an SMOD provider are influenced implicitly.

From a methodological point of view, we base the formulations of our regulated modeling variants on the location-augmented-event-based formulation by Gaul et al. (2025), because, to the best of our knowledge, it is currently the most efficient modeling approach for many settings, and our problems are indeed more complex to solve than standard DARP MILPs. Further, inspired by Braekers and Kovacs (2016), we extend our formulation to a multi-period variant to examine regulatory instruments in a multi-period horizon. We use an exact solution approach by solving the corresponding MILPs to optimality and test them on artificial data by Ropke et al. (2007) as well as real-world data from our industry partner FLEXIBUS (Flexibus, 2024). An overview of the relevant literature and our positioning can be found in Table 1.

3. Problem definition and mathematical formulations

In this section, we present the problem definitions and the mathematical formulations underlying our study. In Section 3.1, we propose a new SDARP generalization, which incorporates specific regulatory instruments, which we term SDARP-REG. In Section 3.2, we then generalize the model to a multi-period variant, termed M−SDARP−REG.

3.1. Selective Dial-a-Ride problem with regulatory instruments (SDARP-REG)

This section starts with a problem description in Section 3.1.1. In Section 3.1.2, we then present the notation for the corresponding MILP formulation, which follows the location-augmented-event-based modeling approach introduced by Gaul et al. (2025). More precisely, first, we focus on the event-based part of the model; second, we introduce the notation derived from location-based models necessary for the location-augmented-event-based SDARP; and third, we provide notation for the MPR and MSAR constraints. Finally, we present the resulting SDARP-REG in Section 3.1.3.

3.1.1. Problem description

The standard SDARP can be described as follows: An SMOD provider receives a number of requests and must make decisions for each one whether it can and should be served. For each request, the pick-up and drop-off location, and the number of passengers are given. We assume a uniformly defined service time for each request which reflects the time span needed for the boarding or alighting of passengers at the pick-up as well as the drop-off location. Each request has a maximum ride time, which must not be exceeded. This maximum ride time is calculated by multiplying the direct ride time from the pick-up to the drop-off location by a factor, which is identical for all requests and represents the relative maximum additional ride time due to detours that is allowed to result from pooling requests. Now, for each request, time windows for both pick-up and drop-off locations are defined, based on a given, request-specific

Table 1
Literature overview, PAF: Profit after Fullfiment, RC: Routing Costs, SR: Served Requests, SP: Served Passenger, WS: Weighted Sum; A: Artificial Data, O: Original Data.

Authors	Selective	Objective	Period	Type of Graph	Data	Regulatory constraints
Braekers and Kovacs (2016)	N	RC	Multi	Location-based	A	N
Cauchi and Scerri (2020)	Y	PAF	Single	Location-based	D	N
Cordeau (2006)	N	RC	Single	Location-based	A	N
Gaul et al. (2022)	Y	WS	Single	Event-based	A, O	N
Gaul et al. (2025)	N	RC	Single	Location- Event-based	A, O	N
Hosni et al. (2014)	Y	PAF	Single	Location-based	A	N
Kirchler and Wolfler Calvo (2013)	Y	WS	Single	Location-based	A	N
Parragh et al. (2015)	Y	PAF	Single	Location-based	A	N
Reinhardt et al. (2013)	Y	WS	Single	Location-based	O	N
Riedler and Raidl (2018)	Y	SR	Single	Location-based	A	N
Rist and Forbes (2021)	N	RC	Single	Fragment-based	A	N
Ropke et al. (2007)	N	RC	Single	Location-based	A	N
Wolfer Calvo and Colorni (2007)	Y	WS	Single	Location-based	A	N
Our work	Y	SP	Single, Multi	Location-Event-based	A, O	Y

desired time within the service horizon. For inbound requests (e.g., for customers intending to travel back home), the desired time is the earliest possible pick-up time. The resulting pick-up time window closes with a uniformly defined waiting time added to the desired time. The respective feasible drop-off time window is then derived by implicitly considering the direct ride time and the maximum ride time given the fixed pick-up time window. For outbound requests (e.g., for traveling from home to an appointment), the opposite applies: The drop-off time window begins with a uniformly defined waiting time before the desired time and ends exactly at the desired time. The pick-up window for outbound requests is determined similarly to the drop-off window for inbound requests, by implicitly considering the direct ride time and maximum ride time. For the detailed calculations of each time window, we refer to Appendix A. To serve requests, the SMOD provider uses a fleet of vehicles that start from the same depot. For the depot, we also specify a time window representing the start and end of the service horizon.

Now, we generalize the SDARP to include regulatory instruments. As motivated in Section 1, public authorities can use regulatory instruments such as an MPR (or MSAR) for SMOD systems to implicitly influence providers to act in a more environmentally (socially) sustainable way. By imposing an MPR, SMOD providers are constrained to consolidate requests and thus improve environmental efficiency. We define the MPR according to the German Passenger Transport Act (Deutscher Bundestag, 2021), in which the pooling rate equals the ratio of the passenger kilometers traveled (PKT) to the vehicle kilometers traveled (VKT). Note that empty vehicle rides, e.g., due to deadheading, are included in the VKT. The PKT comprises the total distance traveled, weighted by the number of passengers who are in the vehicle. By definition, the MPR favors requests with more passengers, which is reasonable, since it is desirable for a provider to transport as many passengers as possible. The MSAR constrains SMOD providers to avoid spatial discrimination of requests, i.e., rejecting requests because of their unfavorable desired pick-up or drop-off locations. To derive the corresponding additional model constraints, first, a descriptive analysis of spatial customer demand patterns must be conducted. Second, based on this analysis, the set of all requests must be divided into disjoint subsets that contain spatially similar requests with particular pick-up-drop-off combinations, e.g., requests with desired pick-up and drop-off locations within a city hub. The MSAR covers all subsets, i.e., a minimum ratio of accepted requests that must be satisfied for all subsets. In this way, a minimum degree of accessibility is ensured regardless of spatial request properties.

3.1.2. Notation

First, we introduce the notation for the event-based graph, which largely follows Gaul et al. (2022). Let the set of requests be $R = \{1, \cdots, n\}$ and the set of homogenous vehicles be $\mathscr{K} = \{1, \cdots, K\}$. Now, an event-based graph is defined as $G = (V, A)$ with nodes V denoting the set of events and arcs A denoting the transition between events.

Intuitively, an event refers to a pick-up or a drop-off of some request. Further, it contains information about the current state of the corresponding vehicle $k \in \mathscr{K}$, i.e., all requests on board as well as the location of the vehicle. The event is denoted as Q-tuple $(v_1, \cdots, v_Q)$, in which Q refers to the maximum number of requests that could potentially be on board. Each component $v_1, \cdots, v_Q$ corresponds to a request $i \in R$ on board of the vehicle. In case the actual number of requests in a vehicle is less than Q, the remaining components of the tuple are marked by 0. The first element of each Q-tuple also includes information about the current position of the vehicle, with each request i being associated with a pick-up location i^+ and a drop-off location i^-. Thus, e.g., the tuple $(1^+, 2, 0)$ denotes an event with the vehicle at the pick-up location of request 1, i.e., 1^+, while passengers of request 2 are already on board of the vehicle. The requests at positions $v_2, \cdots, v_Q$ are sorted in descending order based on the index to avoid equivalent event states due to different permutations. For example, we use the tuple $(1^+, 3, 2)$ instead of $(1^+, 2, 3)$. One major advantage of the event-based graph structure is that the physical capacity constraints of vehicles are considered implicitly, with no more than $\mathscr{C}$ passengers being carried by a vehicle. More precisely, each request i comprises a number of passengers q_i. We eliminate events for which the number of passengers on board exceeds the capacity, i.e., $\sum_{l=1}^{Q} q_{v_l} > \mathscr{C}$.

The set of events V is split into three types of disjoint subsets $V = V_0 \cup \bigcup_{i=1}^{n} V_{i^+} \cup \bigcup_{i=1}^{n} V_{i^-}$. $V_0 = \{(0, \cdots, 0)\}$ refers to the event in which a vehicle is located at the depot, while V_{i^+} and V_{i^-} contain the events corresponding to a pick-up or drop-off location of a request i. For example, V_{i^+} contains all events in which the respective pick-up location i^+ is the first element of the Q-tuple. Arcs describe transitions between events, representing information about rides between locations and the change of requests on board as customers are picked up or dropped off. For example, the arc $((1^+, 0, 0), (2^+, 1, 0))$ represents the ride from the first customer's pick-up location 1^+ to the second customer's pick-up location while customer 1 is on board. The arc set A is divided into six disjoint subsets, which concern the different feasible transitions. For a formal definition of the event and arc sets, we refer to Gaul et al. (2022). Using the event-based graph, we implicitly consider precedence and pairing constraints and exclude infeasible transitions by definition. The traveled distance between two events $v, w \in V$ is defined as c_a with $a = (v, w) \in A$. For each event $v \in V$, we also define the set of incoming arcs $\delta^{in}(v)$ and the set of outgoing arcs $\delta^{out}(v)$. Depending on the number of vehicles, at most K vehicles can leave the depot. This, combined with both $\delta^{in}(v)$ and $\delta^{out}(v)$, allows us to model the flow conservation for all events, and thus guarantees that a route starts and ends at the depot.

Now, we present the notation necessary for the time constraints, which are based on the location-based formulation. As each request $i \in R$ is associated with a dedicated pick-up location i^+ and a dedicated drop-off location i^-, we can define the overall set of pick-up locations $P = \{1^+, \cdots, n^+\}$ and the overall set of drop-off locations $D = \{1^-, \cdots, n^-\}$. Then, the set of all locations equals $J = P \cup D \cup \{0^+, 0^-\}$ with $\{0^+, 0^-\}$ denoting the start and end depot, which we set to a single location 0. In a pure location-based formulation (e.g., Ropke et al., 2007), a complete and directed graph is taken as a basis, in which each arc (i, j) represents the vehicle traveling from location $i \in J$ to $j \in J$ with non-negative travel times $t_{i,j} \geq 0 \ \forall i, j \in J$ and travel distances $c_{i,j} \geq 0 \ \forall i, j \in J$, which fulfill the triangle inequality. Now, for the location-augmented-event-based model, the distance between two events v and w can be defined analogously to the distance between the respective locations in the location-based model, while sticking to the event-based graph, i.e., $c_a = c_{i,j}, a =$

$(v,w) \in A, i,j \in J : v_1 = i \wedge w_1 = j$. The time windows are denoted as $[e_j, l_j]$ for each location $j \in J$. The time windows corresponding to a pick-up or drop-off location are calculated as explained in Section 3.1.1 and in Appendix A. For the depot, we define $e_{0^+} = e_{0^-} = 0 = e_0$ as the start of the service horizon and $l_{0^+} = l_{0^-} = e_0 + T = l_0$ with $T \geq 0$ as the end of the service horizon. For each request, we denote the service time as $s_i \geq 0$. The maximum ride time for a request is defined as L_i, and is calculated by using the maximum additional ride time factor ϵ and multiplying it by the direct ride time t_{i^+,i^-}, thus $L_i = t_{i^+,i^-} + \epsilon \cdot t_{i^+,i^-}$.

Finally, we introduce the notation for the two regulatory instruments, MPR and MSAR, motivated in the introduction and in the problem description. We define an MPR$_{\nearrow}$ to specify a ratio of PKT to VKT that must be at least met. To compute the PKT, we determine the number of transported passengers ω_a for each arc $a \in A$. Similar to our procedure for implicitly considering capacity constraints, we can exploit information in the events about passengers on board, i.e., we can precompute the number of transported passengers with $\omega_a = \sum_{l=1}^{Q} q_{v_l}$ for each arc $a = (v,w)$. Note that we also consider deadheading, e.g., an empty vehicle driving from a drop-off location to a pickup location of another passenger, resulting in $\omega_a = 0$. For MSAR, we divide the set of requests into designated subsets $\sigma \in \mathscr{O} = \{1, \cdots, \Sigma\}$ and assign each request to exactly one subset such that they are mutually exclusive and collectively exhaustive, i.e., $\bigcup_{\sigma \in \mathscr{O}} R_\sigma = R$ and $\bigcap_{\sigma \in \mathscr{O}} R_\sigma = \varnothing$. Each subset $\sigma \in \mathscr{O}$ comprises requests with the same spatial type defined by the underlying OD-pair. Finally, we denote a general MSAR β for each subset $\sigma \in \mathscr{O}$.

For the model formulation we present in the next section, we define the following decision variables:

- x_a is a binary decision variable indicating whether an arc $a \in A$ occurs ($x_a = 1$), otherwise $x_a = 0$.
- p_i is a binary decision variable which equals 1 if request $i \in R$ is served, otherwise $p_i = 0$.
- B_j is a continuous non-negative decision variable, which specifies the start time of service at location $j \in J$.

We refer to Appendix B for a summary of the notation.

3.1.3. Model formulation

Based on the notation introduced in Section 3.1.2, the SDARP-REG can be formulated by the following MILP:

$$Max \sum_{i \in R} p_i q_i \tag{1}$$

subject to

$$\sum_{a \in \delta^{in}(v): v \in V_{i^+}} x_a = p_i \quad \forall i \in R \tag{2}$$

$$\sum_{a \in \delta^{in}(v)} x_a - \sum_{a \in \delta^{out}(v)} x_a = 0 \quad \forall v \in V \tag{3}$$

$$\sum_{a \in \delta^{out}((0,\ldots,0))} x_a \leq K \tag{4}$$

$$e_j \leq B_j \leq l_j \quad \forall j \in J \tag{5}$$

$$B_{i^-} - B_{i^+} - s_i \leq L_i \quad \forall i \in R \tag{6}$$

$$B_j \geq B_i + s_i + t_{ij} - M_{ij}\left(1 - \sum_{v,w \in V: v_1 = i \wedge w_1 = j} x_{(v,w)}\right) \quad \forall i,j \in J \tag{7}$$

$$\sum_{a \in A} x_a c_a \omega_a \geq_{\nearrow} \cdot \sum_{a \in A} x_a c_a \tag{8}$$

$$\frac{\sum_{i \in R_\sigma} p_i q_i}{\sum_{i \in R_\sigma} q_i} \geq \beta \quad \forall \sigma \in \mathscr{O} \tag{9}$$

$$x_a \in \{0,1\} \quad \forall a \in A \tag{10}$$

$$B_j \geq 0 \quad \forall j \in J \tag{11}$$

$$p_i \leq 1 \quad \forall i \in R \tag{12}$$

The objective function (1) maximizes the number of served passengers. Constraints (2) ensure that if a request is served, exactly one incoming arc $a = (w,v)$ of the pick-up event v is used. If a request is rejected, no incoming arc for the pick-up event of a customer is selected. Constraints (3) provide flow conservation for the events, i.e., they ensure that if there is a transition to a particular event, there must also be a corresponding transition out of that event. Constraint (4) guarantees that at most K vehicles are used in the

solution. Constraints (5) and (6) ensure that each time window for every location and the maximum ride time for every request are respected. Constraints (7) preserve the time flow for every location $j \in J$, i.e., the start of service B_j is dependent on the start of service B_i, the travel time $t_{i,j}$, and the service time s_i if location j is visited directly after location i. Here, we utilize the interplay between events and their respective locations. If location j is visited after location i we know that one arc from an event with location i to an event with location j is used, i.e., $x_{v,w} = 1, v, w \in V : v_1 = i \wedge w_1 = j$. The big-M parameter is chosen as a sufficiently large number and can be set to $M_{i,j} = l_i + s_i + t_{i,j} - e_j \ \forall i, j \in J$ (Gaul et al., 2025). Due to the time flow, constraints (7) also prevent subtours, i.e., the route plan includes closed loops that do not consider the depot as a start or end point. Constraint (8) and constraints (9) concern the MPR and MSAR, respectively. To implement the MPR in constraint (8), we impose the pooling rate ν as the minimum factor by which the PKT must at least exceed the VKT. For constraints (9), we propose a general minimum acceptance rate β, i.e., the ratio of served passengers to all passengers in the respective subset $\sigma \in \mathscr{O}$. Finally, constraints (10–12) specify the domains of the decision variables. Note that, due to (2) and (10), p_i is binary. That is, p_i does not have be defined as a binary variable, and it is sufficient to require $p_i \leq 1$, which reduces the number of binary variables.

3.2. Generalization to a Multi-Period problem

So far, we considered the SDARP-REG in a single-period variant. In this section, we generalize the SDARP-REG to a multi-period horizon. In this context, regulatory instruments do not have to be strictly adhered to a single period, e.g., a day. Instead, it is possible to observe multiple periods, e.g., a few days, for which the respective regulatory instrument must be met on average. Since customer demand can be volatile, there may be days at which it is more difficult to comply with a certain level of MPR or MSAR. Therefore, a multi-period approach allows SMOD providers to violate the minimum rates on one day, if the violation is overcompensated on another, ultimately increasing the solution space of the SDARP-REG for individual periods under consideration.

Note that the underlying graph is constructed separately for each day. The M−SDARP−REG is more difficult to solve because we have to consider constraints (8) and (9) that must be satisfied on average over the multi-period horizon. However, as already emphasized in Section 1, the increase in complexity is rather moderate, since the model's size can be expanded by multiple periods without suffering from an exploding graph size.

3.2.1. Notation

To generalize the SDARP-REG to a multi-period problem, we define disjoint monitoring periods $h \in \mathscr{H} = \{1, \cdots, H\}$. From now on, the set of requests for period h is denoted as R^h, where each request $i \in R^h$ contains a number of passengers p_i^h, a service time s_i^h and a maximum ride time L_i^h. The event set V^h, the arc set A^h, the incoming arcs $\delta^{in,h}(v)$ and outgoing arcs $\delta^{out,h}(v)$, and the travel distance c_a^h have an additional index for the planning period. Furthermore, the sets $J^h = P^h \cup D^h \cup \{0^+, 0^-\}^h$ denote the sets of pick-up, drop-off, and depot locations for period h. Similarly, the index for the period is also added to all other location-based parameters such as travel time $t_{i,j}^h$, travel costs $c_{i,j}^h$, and time windows $\left[e_j^h, l_j^h\right]$. For the sake of clarity, we assume that the set of vehicles K and its capacity remains constant over all periods.

To incorporate the MPR in the multi-period formulation, we define the number of passengers ω_a^h separately for each planning period. Analogously, the subsets of requests R_σ have to be considered with a multi-period index, i.e., R_σ^h.

3.2.2. Model formulation

Now, we formulate the M−SDARP−REG as follows:

$$Max \sum_{h \in \mathscr{H}} \sum_{i \in R^h} p_i^h q_i^h \tag{13}$$

subject to

$$\sum_{a \in \delta^{in,h}(v), v \in V_{i^+}^h} x_a^h = p_i^h \quad \forall i \in R^h, h \in \mathscr{H} \tag{14}$$

$$\sum_{a \in \delta^{in,h}(v)} x_a^h - \sum_{a \in \delta^{out,h}(v)} x_a^h = 0 \quad \forall v \in V^h, h \in \mathscr{H} \tag{15}$$

$$\sum_{a \in \delta^{out,h}((0,\ldots,0))} x_a^h \leq K \quad \forall h \in \mathscr{H} \tag{16}$$

$$e_j^h \leq B_j^h \leq l_j^h \quad \forall j \in J^h, h \in \mathscr{H} \tag{17}$$

$$B_{i^-}^h - B_{i^+}^h - s_i^h \leq L_i^h \quad \forall i \in R^h, h \in \mathscr{H} \tag{18}$$

$$B_j^h \geq B_i^h + s_i^h + t_{ij}^h - M_{i,j}^h \left(1 - \sum_{v,w \in V^h : v_1 = i \wedge w_1 = j} x_{(v,w)}^h\right) \quad \forall i, j \in J^h, h \in \mathscr{H} \tag{19}$$

$$\sum_{h\in\mathcal{H}}\sum_{a\in A^h} x_a^h c_a^h \omega_a^h \geq \rho \cdot \sum_{h\in\mathcal{H}}\sum_{a\in A^h} x_a^h c_a^h \tag{20}$$

$$\frac{\sum_{h\in\mathcal{H}}\sum_{i\in R_\sigma^h} p_i^h q_i^h}{\sum_{h\in\mathcal{H}}\sum_{i\in R_\sigma^h} q_i^h} \geq \beta \quad \forall \sigma \in \mathcal{O} \tag{21}$$

$$x_a^h \in \{0,1\} \quad \forall a \in A^h, h \in \mathcal{H} \tag{22}$$

$$B_j^h \geq 0 \quad \forall j \in J^h, h \in \mathcal{H} \tag{23}$$

$$p_i^h \leq 1 \quad \forall i \in R^h, h \in \mathcal{H} \tag{24}$$

The objective function (13) and constraints (14–19, 22–24) are analogous to the single-period SDARP-REG (1, 2–7, 10–12), with the additional index for the monitoring period $h \in \mathcal{H}$ respectively. Constraint (20) ensures the MPR in a multi-period context. Here, the total sum of PKT over all periods $h \in \mathcal{H}$ must at least met the total sum of VKT over all periods $h \in \mathcal{H}$, weighted with the pooling factor ρ. For the MSAR, constraints (21) enforce an average minimum acceptance rate β for each subset σ over the sum of all monitoring periods $h \in \mathcal{H}$.

4. Computational study

Our computational study is divided into three parts: First, we describe the experimental setup in Section 4.1, where we explain the conducted feasibility and impact analyses and define the performance metrics we base our study on. We then present our numerical results of the SDARP-REG for artificial data in Section 4.2, in which we evaluate different supply–demand settings of SMOD providers. In Section 4.3, we show the results of the SDARP-REG and M−SDARP−REG for real-world data and thus, explore the potential advantages of multi-period variants over single-period variants, as we can reasonably compare multiple days. In addition, we can draw practical managerial insights for MPR and MSAR for an SMOD system that is close to the real world. We implemented the mathematical models with Gurobi (Gurobi, 2024) in Python (Python, 2024) and performed all computations on an Intel(R) Core (TM) i7-8700 processor with 6 cores, 3.20 GHz, and 32 GB RAM.

Importantly, note that Section 4 focuses mainly on the presentation and technical evaluation of the computational study. It is complemented by Section 5, in which we interpret the numerical results in more context, putting them in relation to each other in order to gain overarching managerial insights with regard to the practical application of regulatory instruments.

4.1. Experimental setup

In our study, we solve the SDARP-REG and M−SDARP−REG to optimality for varying levels of regulation. To reduce computational complexity, we adopted preprocessing methods of Gaul et al. (2025) to eliminate impossible event nodes and arcs due to capacity, time window, and ride time constraints. We consider both regulatory instruments separately so that we can make statements without the two influencing each other. Accordingly, we perform a feasibility analysis and an impact analysis for MPR and MSAR independently, starting from the unregulated case ($\rho = 0, \beta = 0$) and incrementally increasing the level of regulation by a fixed step size.

- *Feasibility analysis*: We use feasibility graphs to illustrate an SMOD system's ability to comply with regulatory instruments, showing the proportion of feasible instances at different levels of regulation. This helps us determine the maximum applicable regulation level (MARL) for MPR and MSAR, where all instances remain feasible.
- *Impact analysis*: We then measure the impact of the respective regulatory instrument up to the MARL on metrics that represent goals from a provider, environmental, and social perspective.

In the following, we define the metrics that we evaluate in our impact analysis. We refer to Table 2 for the formalization of the introduced metrics and Appendix J for a summary of used abbreviations.

Table 2
Definition of performance metrics (impact analysis).

Goal	Metric	Formula
Service Reliability	Global Acceptance Rate (GAR)	$\frac{\sum_{i\in R} p_i q_i}{\sum_{i\in R} q_i}$
Ecological Impact	Distance Savings per Passenger (DSP)	$\frac{\sum_{i\in R} p_i q_i c_{i^+,i^-} - \sum_{a\in A} x_a c_a}{\sum_{i\in R} p_i q_i}$
Social Equitable Mobility	Standard Deviation of Spatial Acceptance Rates (SDSA)	$\frac{1}{\lvert\mathcal{O}\rvert}\sqrt{\sum_{\sigma\in\mathcal{O}}\left(\frac{\sum_{i\in R_\sigma} p_i q_i}{\sum_{i\in R_\sigma} q_i} - \frac{\sum_{i\in R} p_i q_i}{\sum_{i\in R} q_i}\right)^2}$

- *Global acceptance rate (GAR):* This metric represents the percentage of passengers whose ride requests are successfully served. It is calculated by dividing the number of served passengers by the total number of passengers in the SMOD system, taking into account that a single request can include multiple passengers. A high GAR indicates service reliability, suggesting that passengers have a good chance of having their requests fulfilled. This metric is associated with the primary maximization objective of an SMOD provider.
- *Distance savings per passenger (DSP):* DSP is calculated as the difference between the total sum of all booked PKT, i.e., the direct travel distance for a served request c_{i^+,i^-} weighted by the number of passengers, minus the total VKT. This term is divided by the number of served passengers. DSP quantifies the ecological impact of SMOD compared to individual motorized transport and is especially relevant for MPR as a target metrics.
- *Standard deviation of the spatial acceptance rates (SDSA):* SDSA measures the standard deviation of the passenger acceptance rates of the disjoint spatial subsets $\sigma \in \mathscr{O}$, whereby a low dispersion is an indicator of socially sustainable mobility. SDSA is a target metric for evaluating the impact of MSAR, as it reflects the balance in acceptance likelihood across different spatial request types. Unlike GAR, SDSA provides more detailed insights into specific customer groups. Note that the calculation and interpretation of the SDSA depends on the definition of the subsets $\sigma \in \mathscr{O}$.

4.2. Study on artificial data

In this section, we evaluate the SDARP-REG presented in Section 3.1 using artificial data. This section is divided into three parts: Section 4.2.1 introduces the data and instance characteristics, followed by a presentation and technical evaluation of the results for MPR (Section 4.2.2) and MSAR (Section 4.2.3).

4.2.1. Problem instances

We use the *aK_n* instances provided by Ropke et al. (2007), which are based on Cordeau (2006) with a denoting the so-called "a-instances", K the number of vehicles, and n the number of requests. The service area comprises a Cartesian coordinate system ($[-10, 10] \times [-10, 10]$) with the depot at the center $(0, 0)$ and uniformly distributed pickup and drop-off locations for n requests. The travel distance between locations is calculated using the Euclidean Distance and equals the travel time. Each request $i \in R$ has a desired time drawn from a uniform distribution over the service horizon $[0, T]$ which defines time windows with a fixed maximum waiting time of $w = 15$, a uniform absolute maximum ride time $L_i = 30$, and service time $s_i = 1$. Note, that in contrast to our notation in Section 3.1.2, Cordeau (2006) computes the time windows for each pick-up and drop-off location by using a fixed absolute maximum ride time L_i, which is constant for each request and independent of t_{i^+,i^-} and ϵ. The service horizon T varies among instances and ranges from 240 to 720. The vehicle fleet is homogeneous with a passenger capacity of $\mathscr{C} = 3$ and the number of passengers is identically set to $q_i = 1$ for $i \in R$, i.e., the maximum number of requests on board of each vehicle is $Q = 3$.

To perform experiments on the MSAR, we need to spatially partition the set of requests into subsets. For simplicity, we divide the set of requests R into four subsets corresponding to the four quadrants of the Cartesian coordinate system, i.e., $\mathscr{O} = \{1, \cdots, 4\}$. A request i is assigned to the subset R_σ depending on the pick-up location. Thus, we can apply an MSAR to each subset, which should ensure that each request has a comparable chance of being served, regardless of the desired pick-up location.

The a-instances were originally designed for non-selective DARPs. Inspired by Berbeglia et al. (2011), we aim to create realistic supply–demand settings for real-world SMOD systems, as described in Section 2.2. Thus, we modified the a-instances with regard to a varying fleet size K and solved the respective unregulated SDARPs in preliminary experiments. The maximum GAR gives us an indication of how well the supply, i.e., the fleet size K, can serve the demand, i.e., the number of requests n. Drawing from recent literature (VDV, 2023), we then define three supply–demand settings for each a-instance, depending on the maximum GAR for the particular *K_n* combination:

- *Low supply–demand (LSD)*: The supply–demand ratio is low with a GAR $\in [0.50, 0.70[$.
- *Balanced supply–demand (BSD)*: The supply–demand ratio is balanced with a GAR $\in [0.70, 0.95[$.

Table 3
Modified artificial data (size of vehicle fleet K).

Instance	n	T	LSD*	BSD*	HSD*	Instance	n	T	LSD*	BSD*	HSD*
aK_16	16	480	–	1	2	*aK_40*	40	480	1	2	3
aK_20	20	600	–	1	2	*aK_50*	50	600	–	2	3
aK_24	24	720	–	1	2	*aK_60*	60	720	2	3	4
aK_18	18	360	1	2	3	*aK_48*	48	480	–	2	3
aK_24	20	480	1	2	3	*aK_60*	60	600	2	3	4
aK_30	30	600	1	2	3	*aK_72*	72	720	–	2	4
aK_36	36	720	1	2	3	*aK_56*	56	480	–	3	4
aK_16	16	240	1	2	3	*aK_70*	70	600	2	3	4
aK_24	24	360	1	2	3	*aK_84*	84	720	2	3	4
aK_32	32	480	1	2	4	*aK_64*	64	480	2	3	4
aK_40	40	600	1	2	3	*aK_80*^	80	600	2	3	4
aK_48	48	720	1	2	3	*aK_96*	96	720	2	3	4

- *High supply–demand (HSD)*: The supply–demand ratio is high with a GAR ≥ 0.95.

Looking at Table 3, we select at most one *K_n* combination for each setting and *a*-instance, resulting in our modified set of instances. Therefore, we only consider the combination with the smallest *K* that yields a GAR in the respective range of the setting. Comparability among instances within the same supply–demand setting is ensured by the common GAR range. If no *K_n* combination exists for an instance and setting, the cell is marked with a "-" in Table 3.

In our computational study on artificial data, we solve the SDARP-REG to optimality for increasing levels of regulation and for each instrument separately. If no optimal solution is found within one hour for any level, the instance is marked with * (MPR) or ˆ (MSAR) and excluded from further analysis. In general, almost all instances are solved in less than 10 min across all supply–demand settings and regulatory instruments, with less than 1% of outliers, especially in the LSD setting with a high number of requests. We chose to focus on a-instances with uniformly one passenger per request and excluded experiments with *b*-instances, whose most relevant distinction is the increased number of passengers. Instead, we conduct experiments with multi-passenger requests using real-world data in Section 4.3, as descriptive analysis reveals that the assumptions about the number of passengers in b-instances are not entirely realistic.

4.2.2. Numerical results for minimum pooling rate

This section presents the numerical results for implementing an MPR based on artificial data. As stated in Section 4.1, we start with the feasibility analysis and then conduct the impact analysis.

Feasibility analysis: We illustrate the proportion of feasible instances at a specific level of regulation $\nearrow$ using feasibility graphs for each supply–demand setting in Fig. 1. The graphs are structured as follows: the abscissa shows the regulation level, starting from $\nearrow = 0$ and increasing in 0.05 steps, while the ordinate shows the proportion of feasible instances at each level. The graphs are color-coded for each supply–demand setting: green (LSD), black (BSD), red (HSD), and dashed vertical lines indicate the MARL for each setting.

Fig. 1 shows that the feasibility graphs follow a similar trajectory across all settings, indicating that feasibility for an MPR depends largely on the instances' demand structure and its pooling compatibility, rather than on the supply–demand setting. Thus, the provider can only improve MPR to a limited extent. Please note that the BSD and HSD settings yield identical results in the feasibility analysis; therefore, the black curve is covered by the green curve. The SMOD system achieves an MARL of $\nearrow = 0.95$ for all settings, followed by a steep drop at $\nearrow = 1.0$. This suggests that more than 10% of instances in each setting cannot comply with $\nearrow = 1.0$, as the provider is not able to generate more PKT than VKT. In Appendix C, we illustrate the limit for the achievable pooling rate with a didactic example. Since the MPR is distance-based (PKT divided by VKT), some pooling is possible even at $\nearrow < 1.0$. However, due to deadheading, i.e., empty vehicle rides generating high VKT and no PKT, only pooling rates $\nearrow < 1.0$ are achievable for some instances. A small number of instances (around 5% for BSD, HSD) remain feasible up to $\nearrow = 1.30$, but no instances comply with $\nearrow \geq 1.35$. In the following impact analysis, we focus on regulation levels up to the MARL of $\nearrow = 0.95$.

Impact analysis: We now evaluate the relative impact of $\nearrow$ on the metrics introduced in Section 4.1 compared to the benchmark $\nearrow = 0$. Specifically, we compute the mean values for GAR, DSP, and SDSA across the set of instances for each setting for $\nearrow \in \{0, 0.05, \ldots, 0.9, 0.95\}$ and calculate the relative deviation from the benchmark. The results are visualized in Fig. 2, with regulation levels $\nearrow$ on the abscissa and relative deviations from the benchmark (in percent) on the ordinate. Dashed vertical lines indicate the MARL for each supply–demand setting.

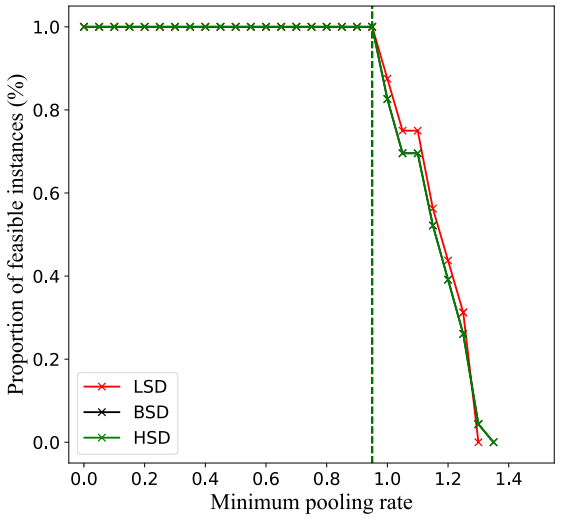

Fig. 1. Feasibility analysis on artificial data – MPR.

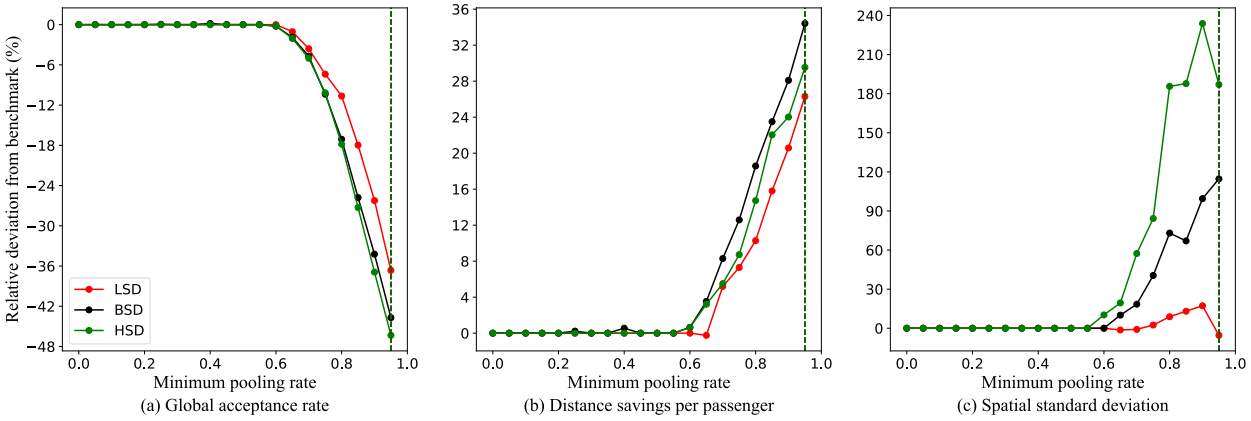

Fig. 2. Impact analysis on artificial data – MPR.

- *Impact – GAR*: Fig. 2a shows the mean deviation of GAR, which is consistently non-positive as $\nearrow$ increases, with similar magnitudes across all supply–demand settings: up to −36.64% (LSD), −43.70% (BSD), and −46.32% (HSD) for $\nearrow = 0.95$. This trend is expected, as the GAR (primary objective) decreases monotonically with increasing restrictions.
- *Impact – DSP*: Fig. 2b shows that DSP consistently becomes larger with increasing $\nearrow$, with a positive deviation from the benchmark until $\nearrow = 0.95$, reaching +26.30% (LSD), +35.41% (BSD), and +29.53% (HSD). The graphs are similar across all settings, indicating no noticeable dependence of DSP on the supply–demand ratio. Notably, Fig. 2a and Fig. 2b indicate a conflicting interaction between GAR and DSP: Around $\nearrow = 0.6$, the GAR decreases while DSP improves, suggesting that the provider may actively reject requests to adhere to MPR (Appendix F), thereby indirectly enhancing DSP. We discuss this interaction in more detail in Section 5.2.1.
- *Impact – SDSA*: Lastly, we examine the SDSA deviations in Fig. 2c, where a higher standard deviation of spatial acceptance rates is not desirable for spatial accessibility. As $\nearrow$ increases, SDSA almost consistently worsens, particularly for BSD (+ 114.65%) and HSD (+ 187.00%) at $\nearrow = 0.95$, indicating an increasingly negative impact on providers with a larger vehicle fleet. This is not surprising, as BSD and HSD show high GAR and thus have low SDSA without regulation, since most requests are served. As $\nearrow$ increases, rejected requests might lead to an imbalance in spatial acceptance rates.

4.2.3. *Numerical results for minimum spatial acceptance rate*

Analogous to Section 4.2.2, this section presents the numerical results of the feasibility and impact analyses for MSAR using artificial data.

Feasibility analysis: In Fig. 3, we display the regulation levels of β, increasing incrementally by 0.05 steps on the abscissa, and the proportion of still feasible instances on the ordinate with different supply–demand settings and their corresponding MARL.

Unlike MPR, Fig. 3 shows differences in the feasibility graphs for MSAR depending on the supply–demand setting. The MARL is

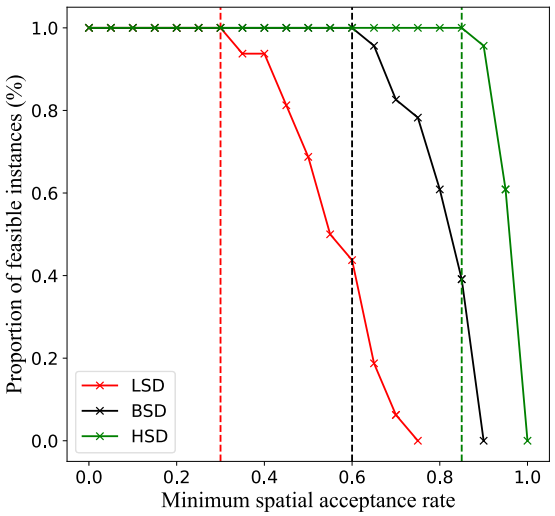

Fig. 3. Feasibility analysis on artificial data – MSAR.

$\beta_{LSD} = 0.30$, $\beta_{BSD} = 0.60$, and $\beta_{HSD} = 0.85$, highlighted by the vertical dashed lines. This is expected, as providers achieve higher GAR in BSD and HSD, leading to more feasible instances at larger β values. As β increases, the acceptance rates of the spatial subsets converge, but eventually, the MSAR becomes infeasible for at least one instance, leading to the MARL at the previous regulation level.

Impact analysis: As in Section 4.2.2, we limit the discussion of the numerical results up to the MARL of the respective supply–demand setting (Fig. 4).

- *Impact – GAR*: Fig. 4a shows only marginal negative deviations in GAR, with a decrease of -0.78% for $\beta_{LSD} = 0.30$ and no impact on BSD or HSD. Thus, increasing MSAR levels lead to "request swaps" rather than eliminations, with requests from demand-dense areas being replaced by those from low-demand areas. This indicates that the SMOD provider can balance spatial acceptance rates through selective demand management without substantially reducing GAR.
- *Impact – DSP*: The DSP graph for LSD (Fig. 4b) reveals a consistent negative deviation (-2.45% for $\beta_{LSD} = 0.30$), suggesting that DSP worsens as β increases, particularly when supply is limited. However, for BSD and HSD, the DSP graphs remain nearly constant.
- *Impact – SDSA*: Fig. 4c demonstrates a consistently positive effect on SDSA across all settings. The trend for SDSA is almost the inverse of that for DSP. Specifically, in the LSD setting, we observe notable improvements in SDSA (-14.80% for $\beta_{LSD} = 0.30$), highlighting the potential of MSAR to improve socially equitable mobility, especially in SMOD systems with low supply. For BSD and HSD, the impact on all performance metrics is less pronounced due to higher overall GAR and limited potential for improvement.

Note that the proposed quadrant-based partitioning approach is beneficial for MSAR as requests are uniformly distributed. Furthermore, despite promising results, there may still be spatial discrimination of certain parts of the service area, particularly for requests located at the edge of the service area. Therefore, we also apply a different approach that introduces asymmetry in the distribution of locations and addresses this issue. To achieve this, we conduct another experiment, dividing the requests into two subsets: the center area ($\sigma = 1$) covers the inner part of the Cartesian coordinate system, forming a square around the depot at $(0, 0)$ with edge lengths of 15, occupying a larger portion of the service area, while the peripheral area ($\sigma = 2$) covers the outer area. A request is assigned to the center area if the pick-up location lies within $(|x| \leq 7.5) \wedge (|y| \leq 7.5)$, and to the peripheral area if it falls outside this boundary $(|x| > 7.5) \vee (|y| > 7.5)$. This partitioning introduces asymmetry into the operational planning. An illustration of this partititoning approach and numerical results are provided in Appendix D. We will address issues related to the practical application of MSAR, including the selection of a suitable partitioning approach, in Section 5.2.2.

4.3. Study on Real-World data

In this section, we evaluate the SDARP-REG and M−SDARP−REG variants (Sections 3.1 and 3.2) using a real-world dataset provided by our industry partner, FLEXIBUS (Flexibus, 2024). Out of the different service areas in Bavaria, we investigate the city of Krumbach with its surrounding regions, which is a mature service area that was already established in 2009. Section 4.3.1 describes the data and parameters of the SMOD system, while Sections 4.3.2 and 4.3.3 present the numerical results for MPR and MSAR.

4.3.1. Problem instances

We use historical data starting in February 2022, and, for comparability, we consider the first 100 working days (Monday to Thursday) with an uniform service horizon (5:00 a.m. to 9:00p.m.). Fridays, weekends, and holidays were excluded due to differing service horizons and statistical dissimilarities. Each of the 100 service days represents one instance, with an average of 81.53 requests per day.

The fleet consists of two vehicles $K = 2$ that are continuously in service during each service day. In preliminary experiments, we solved all instances for the unregulated SDARP. The vehicle fleet is homogeneous with a passenger capacity of $\mathscr{C} = 8$. FLEXIBUS

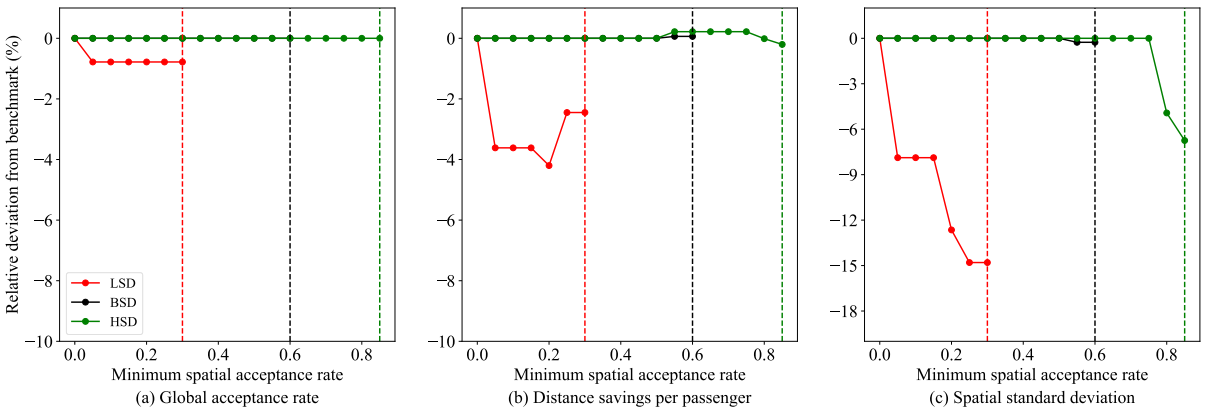

Fig. 4. Impact analysis on artificial data − MSAR.

operates with 563 virtual stops in the service area, including a depot. When placing a request, a customer specifies their pick-up and drop-off locations which are assigned to the closest virtual stops within the service area. The travel distance matrix and travel time matrix are computed using Open Source Routing Machine (OSRM, 2024) and we assume a constant uniform service time of one minute ($s_i = 1$). Each customer defines a desired time for an inbound or outbound request within the service horizon, represented by the depot's time window of $[0, 960]$. To derive the time windows for pick-up and drop-off, we use a uniform waiting time of $w = 30$ minutes and a maximum added travel time factor $\epsilon = 0.5$ to the direct travel time for detours. See Appendix A for a more detailed description of the resulting time windows.

We illustrate the spatial distribution of all requests within the observation period in a flow map in Fig. 5. The 563 stops are depicted as gray dots on the map. For clarity, we grouped the flow arrows into 24 service zones with each stop assigned to one zone. The flow arrows start and end at zone centers (red dots in Fig. 5), representing requests for rides between zones. The size of the arrow indicates the frequency of requests. Again to keep clarity, only the flow arrows between zones that occurred at least two times during the observation period are shown, and rides inside Krumbach (pick-up and drop-off location inside of Krumbach) are represented by the gray arrow in the center.

Looking at the flow arrows in Fig. 5, we identify an hub-and-spoke-type spatial distribution of demand, with the city of Krumbach being the most selected origin (54.9%) and the most selected destination (57.3%), and the peripheral area, except for two larger towns, showing low demand.

Similar to the experiments on artificial data, we spatially partition the set of requests into subsets as a basis for our experiments on MSAR. Motivated by the underlying demand structure, we define four different spatial subsets as follows:

- *Inside Krumbach*: Both pick-up and drop-off are within the city of Krumbach.
- *To Krumbach*: The pick-up is in peripheral areas, and the drop-off is in Krumbach.
- *From Krumbach*: The drop-off is in peripheral areas, and the pick-up is in Krumbach.
- *Peripheral areas*: Both the pick-up and the drop-off are in peripheral areas.

Each request $i \in R$ is assigned to one of the four subsets based on its OD-pair. By defining different types of subsets, we are going to analyze how spatial distribution affects reliability (acceptance rates) and how it can potentially be addressed by MSAR.

Unlike the experiments on artificial data, the real-world study includes results for the single-period variant (1P) and two multi-period variants with five (5P) monitoring periods $|\mathscr{H}| = 5$ and ten (10P) monitoring periods $|\mathscr{H}| = 10$, respectively. The three variants (1P, 5P, 10P) are solved for all 100 historical days, for both regulatory instruments, and for different levels of regulation. For the M−SDARP−REG variants, we consider five (or 10) consecutive working days as the monitoring period and enforce the regulatory minimum rates on the average of the multiple days. Thus, the 100 historical days result in 20 (10) instances. Each variant is solved to optimality for both MPR and MSAR, starting from $\gamma = 0$ and $\beta = 0$, with regulation levels increasing by a stepsize of 0.05. Despite the increasing size of the M−SDARP−REG, all instances are solved in under 10 min for every variant. Notably, the average runtime for all instances and regulation levels up to the 1P-MARL of MPR (MSAR) only increases by $+0.45\%$ ($+0.63\%$) for the 5P and $+2.66\%$ ($+1.58\%$) for 10P compared to the 1P variant.

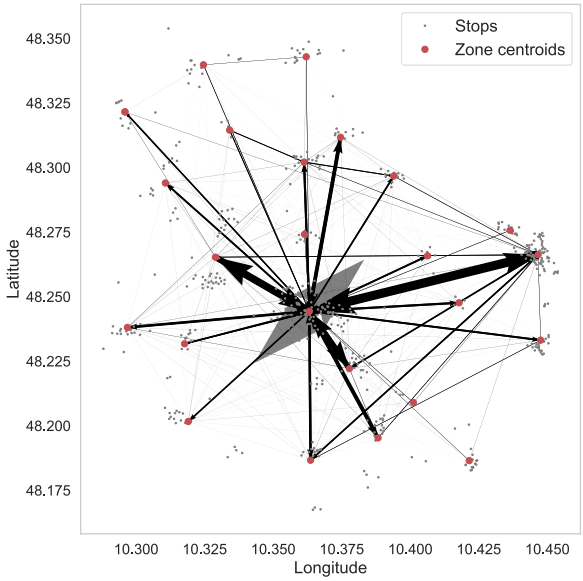

Fig. 5. Descriptive analysis on desired OD-pairs in Krumbach.

4.3.2. Numerical results for minimum pooling rate

The feasibility and impact analyses for MPR on real-world data follow the same structure as in Section 4.2.2.

Feasibility analysis: Fig. 6 shows the feasibility graphs for the variants 1P (red), 5P (black) and 10P (green) for MPR. At first glance, the instances tolerate higher levels of $\nearrow$, with MARLs of $\nearrow_{5P} = 1.30$ and $\nearrow_{10P} = 1.40$ compared to $\nearrow_{1P} = 1.10$. Almost generally, the multi-period feasibility graphs lie above the single-period variant, as multi-period variants offer more flexibility by offsetting "bad" days with "good" ones. The graphs converge at $\nearrow \geq 2.35$, where a few instances remain capable of achieving a high $\nearrow$.

Impact analysis: In Fig. 7, the levels of $\nearrow$ are displayed on the abscissa, while the relative deviations from the unregulated benchmark ($\nearrow = 0$) are plotted on the ordinate. Vertical dashed lines indicate the MARL for each variant (1P, 5P, 10P).

- *Impact – GAR*: Fig. 7a shows the relative deviations of GAR, which decrease monotonically as the MPR increases. Up to the single-period $\text{MARL}_{\nearrow_{1P}} = 1.10$, the multi-period variants (5P, 10P) consistently outperform the single-period variant, with fewer losses in GAR (-33.82% (1P), -30.04% (5P), and -29.57% (10P) for $\nearrow = 1.10$). The overall negative trend is expected, as GAR decreases with increasing restrictions.
- *Impact – DSP*: In Fig. 7b, there are only slight differences between the variants in their impact on DSP. After a plateau, DSP improves, reaching $+80.26\%$ (1P), $+83.38\%$ (5P), and $+84.39\%$ (10P) up to $\nearrow = 1.10$. In both our artificial data (Fig. 2a and Fig. 2b) and real-world data results, we observe a conflicting interaction between GAR and DSP, where rejecting requests indirectly enhances DSP. The DSP graphs flatten slightly for all model variants despite the rejection rates increasing linearly (Appendix F). This suggests that the indirect improvement in DSP diminishes beyond a certain point ($\nearrow > 1.05$), either because rejections become less effective at higher MPR levels or because the provider is enforced to artificially increase PKT to comply with high MPR levels. For example, providers may inflate PKT by introducing unnecessary detours with multiple passengers onboard, which could reduce the positive ecological impact of the MPR. We will address these issues and further implications critically in Section 5.2.1.
- *Impact – SDSA*: Fig. 7c depicts the impact on SDSA, which worsens across all variants. The multi-period variants, especially 10P, maintain lower SDSA deviation than the single-period variant up to $\nearrow = 1.00$. After that, the trend reverses, with the single-period variant yielding SDSA with $+113.93\%$ for 1P, $+123.71\%$ for 5P, and $+119.74\%$ for 10P at $\nearrow = 1.10$. Interestingly, for $\nearrow > 1.2$, the deterioration is less severe. This must be interpreted with caution, as the number of passengers served decreases (Fig. 7a) and acceptance rates for spatial subsets approach zero, leading to more balanced spatial acceptance rates.

4.3.3. Numerical results for minimum spatial acceptance rate

The feasibility and impact analyses follow the same structure as the MSAR results in Section 4.2.3.

Feasibility analysis: In Fig. 8, we observe steeper curves compared to Fig. 7 (MPR), indicating that the MSAR is more abrasive. As β increases, acceptance rates of the spatial subsets converge until MSAR becomes infeasible, i.e., at minimum one subset cannot achieve β. For the single-period variant, the MARL is $\beta_{1P} = 0.55$, while the multi-period variants (5P, 10P) yield an MARL of $\beta_{5P} = \beta_{10P} = 0.65$, reflecting the larger solution space for MSAR application. A closer look into the average acceptance rates per spatial subset shows that peripheral trips are the limiting factor, preventing $\beta_{1P} > 0.55$ or $\beta_{5P}, \beta_{10P} > 0.65$ on average for all instances. As with artificial data, MSAR feasibility strongly depends on the spatial partitioning approach.

Impact analysis: With regard to the impact analysis for MSAR, we expect that there will most likely be some spatial discrimination in the unregulated case, as the locations are not uniformly distributed. As with experiments on MPR, both single-period and multi-period variants are considered in Fig. 9.

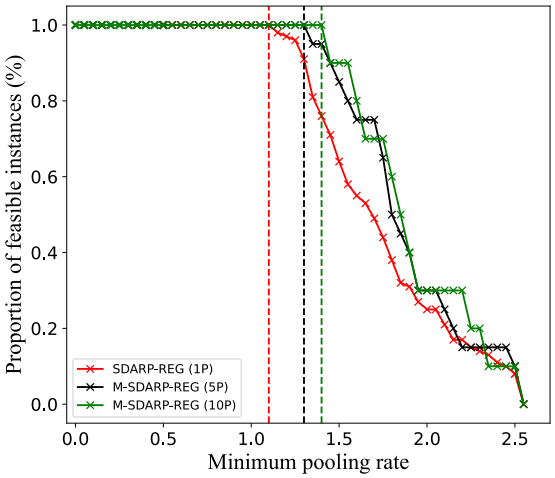

Fig. 6. Feasibility analysis on real-world data – MPR.

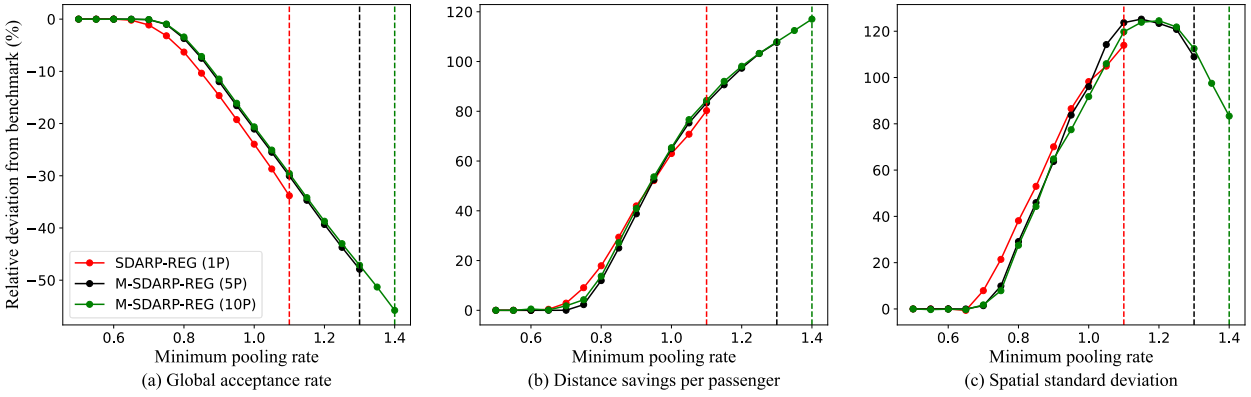

Fig. 7. Impact analysis on real-world data – MPR.

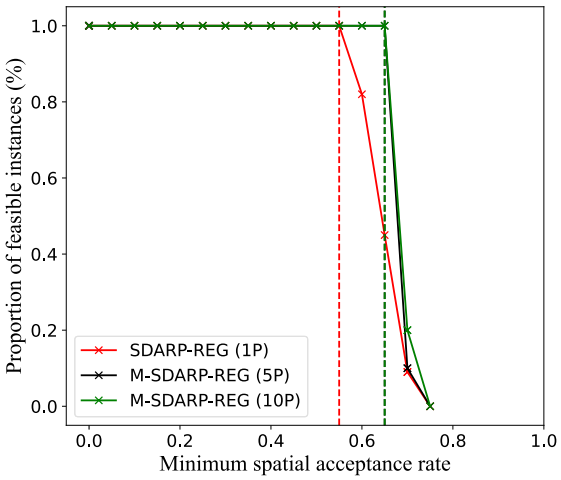

Fig. 8. Feasibility analysis on real-world data – MSAR.

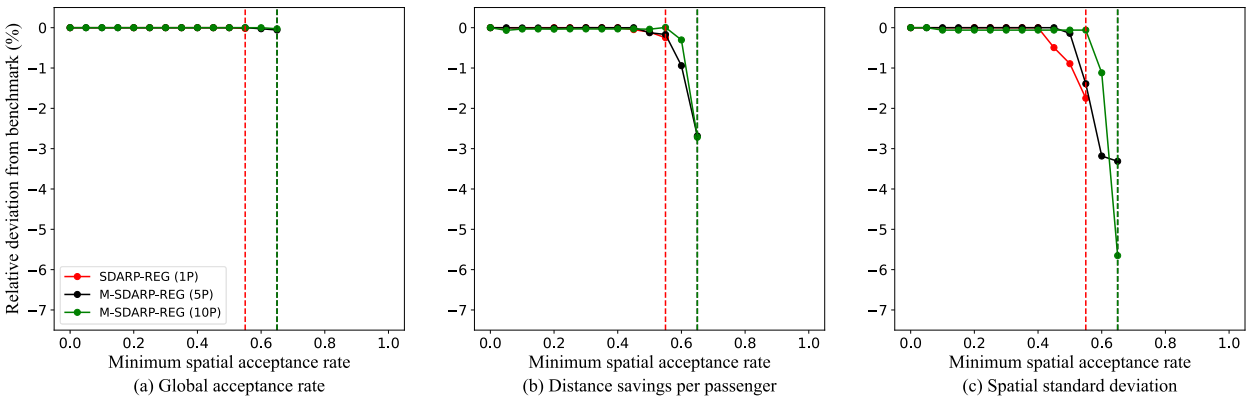

Fig. 9. Impact analysis on real-world data – MSAR.

- *Impact – GAR*: In Fig. 9a, GAR remains constant up to $\beta_{1P} = 0.55$ and $\beta_{5P} = \beta_{10P} = 0.65$. This implies that there is almost no loss of service reliability, with GAR reductions below 1%. Once again, the multi-period variants perform slightly better than the single-period variant.
- *Impact – DSP*: As β increases, we observe a decline in DSP, with multi-period variants dominating the single-period variant. For $\beta_{1P} = 0.55$, the deviations are -0.24% (1P), -0.17% (5P), and -0.01% (10P), while remaining within a moderate range (Fig. 9b).

- *Impact – SDSA*: SDSA improves across all variants (Fig. 9c), meaning that, with minimal GAR loss, acceptance rates become more balanced across spatial subsets. However, unlike GAR and DSP, the single-period variant shows greater impact on SDSA than the multi-period variants for the same β level. For example, at $\beta_{1P} = 0.55$, the SDSA improvements are -1.75% (1P), -1.39% (5P) and -0.06% (10P). This suggests that the single-period variant performs better in SDSA at lower MSAR, likely because it enforces strict minimum levels for each monitoring period. However, for $\beta_{5P} = \beta_{10P} = 0.65$, the multi-period variants achieve deviations of -3.31% (5P) and -5.65% (10P).

Note that the numerical results depend on the spatial partitioning approach. As the chosen approach is OD-based, alternative approaches could focus solely on pick-up locations, as in Section 4.2.1. For instance, the set of requests could be divided into the 24 service zones (Fig. 5) or aggregated into specific subsets of these service zones to improve accessibility and socially equitable mobility. For further discussion and results in this regard, we refer to Appendix E.

5. Criticial discussion of practical application

Based on the results of our computational study, in this section, we provide a critical discussion of the practical application of governmental regulation. To this end, we analyze and interpret the individual results of our computational study (Section 4) in more context and put them in relation to each other to gain overarching managerial insights with regard to several dimensions: First, in Section 5.1, we address two potential key factors for the practical application of the instruments (MPR and MSAR), namely, the available fleet size and the customer demand structure. Section 5.2 then elaborates on specific issues related to the application of the instruments, followed by recommendations in Section 5.3.

5.1. Key factors

First, we discuss key factors that might influence the practical application of MPR and MSAR.

5.1.1. Minimum pooling rate

The implementation of an MPR is challenging as the German Passenger Transportation Act (Deutscher Bundestag, 2021), for example, lacks explicit guidelines regarding specific MPR values. As a result, public authorities must find an adequate MPR level while considering the limitations of SMOD systems in achieving high MPRs. To better understand the factors that may constrain the MARL of an SMOD system, we analyze the fleet size on the provider side and demand structure on the customer side as two key factors:

- *Available vehicle fleet size:* Our numerical results suggest that implementing an MPR is largely independent of the size of the available vehicle fleet, as shown by the experiments with artificial data (Section 4.2.2). Across all supply–demand settings, the feasibility and impact graphs exhibit nearly identical patterns. This is logical, as the MPR primarily reflects how efficiently passengers can be consolidated by a single vehicle without generating excessive VKT.
- *Customer demand structure*: In contrast to fleet size, the demand structure impacts the MARL. Our numerical results confirm this: For real-world data, the MARL is $+15.79\%$ higher (1P) compared to that of artificial data (LSD, BSD, HSD). Two main factors contribute to this: First, requests in real-world data average of 1.18 passengers per request, unlike the uniform single-passenger requests in artificial data. Prioritized serving of multi-passenger requests increases PKT, supporting MPR compliance. This is evident in our real-world results: Comparing the MARL ($\nearrow = 1.10$) to the benchmark ($\nearrow = 0$), we observe a substantial drop in service reliability (GAR) by -33.82%. However, the acceptance rates for multi-passenger requests remain relatively stable compared to the benchmark (-1.10% for $\nearrow = 1.10$). This suggests that serving multi-passenger requests is more advantageous for MPR, as demand is naturally consolidated, provided there is enough vehicle capacity. Second, the spatial distribution of demand in real-world data follows a more structured hub-and-spoke pattern (Fig. 5), making it more conducive to pooling than the uniform distribution in artificial data. This can enable providers to better consolidate demand, resulting in higher MARLs.

5.1.2. Minimum spatial acceptance rate

Building on the discussion of MPR, we now turn our focus to MSAR. The application of an MSAR also presents several challenges. First, public authorities must establish a spatial partitioning approach that divides the service area into subsets to ensure equitable accessibility. Second, similar to MPR, the public authority must determine an appropriate MSAR level. Again, we focus on the available fleet size and the demand structure as key factors, potentially limiting the MARL of the MSAR:

- *Available vehicle fleet size:* Unlike MPR, MSAR clearly depends on the supply–demand ratio. In our numerical results for artificial data, the maximum applicable MSAR increases substantially ($\beta_{LSD} = 0.30$ up to $\beta_{HSD} = 0.85$). This is expected, as a larger fleet can serve almost all requests, resulting in high acceptance rates within the subsets, regardless of the spatial partitioning approach.
- *Customer demand structure*: The spatial distribution of demand plays a key factor in the practical application of MSAR, as it directly affects its effectiveness. Depending on this distribution, inequalities in the accessibility of requests across different OD-pairs may arise for a given fleet. For example, our artificial data (Section 4.2.1, Appendix D) reveals poor GAR in peripheral areas $\sigma = 2$ (49.68% at $\nearrow = 0$), with substantially better GAR in the center area $\sigma = 1$ (70.45% at $\nearrow = 0$), illustrating how accessibility

discrimination can occur. This emphasizes the need to identify such areas and develop spatial partitioning approaches that account for such discrimination.

5.2. Specific issues

In this section, we outline issues regarding the practical application of both regulatory instruments.

5.2.1. Minimum pooling rate

We first discuss feasibility issues, which serve as the foundation for addressing additional concerns, particularly the goal conflicts that arise in the multidimensional SMOD environment. The primary aim of implementing an MPR is to enhance the ecological impact (DSP), for which we observe improvements compared to the unregulated scenario. However, these gains come with notable trade-offs in service reliability (GAR) and socially equitable mobility (SDSA). These trends are evident in both artificial and real-world data (Fig. 2, Fig. 7). In the following, we will outline and discuss application issues associated with MPR application:

- *Feasibility issues*: As discussed in Section 5.1.1, the demand structure of an SMOD system influences the feasibility of an MPR. Therefore, determining feasible levels of regulation is crucial and must be tailored by public authorities. A dedicated feasibility analysis can assess, given the demand structure, what regulation level is possible to achieve and reasonable (Fig. 1, Fig. 6). Another closely related issue is the appropriate time horizon over which the MPR should be met on average (e.g., daily, weekly), as, the German Passenger Transport Act lacks clear guidelines on this matter.
- *Conflict with service reliability*: As the MPR increases, SMOD providers are enforced to reject requests that are difficult to consolidate due to "unfavorable" OD-pairs, such as those involving long distances to the pick-up or from the drop-off location. Even when supply is sufficient to serve these requests, providers may reject them because they do not generate enough PKT, but heavily increase VKT. This trend is evident in Appendix F, which shows a steady increase in the number of rejected requests for increasing MPR levels. Note that excessive rejections could also have a serious negative long-term impact on the public acceptability of an SMOD system.
- *Conflict with socially equitable mobility*: As noted earlier, MPRs encourage demand management that prioritizes requests easier to consolidate, such as those with similar OD-pairs. However, this also negatively impacts socially equitable mobility (SDSA). As the MPR increases, the growing number of rejected requests means that certain areas are served, leaving others underserved.
- *Negative ecological impact of rejections*: While rejecting requests to improve ecological impact (DSP) seems beneficial in our technical evaluation in Section 4, it can backfire if those customers drive their own cars, potentially increasing emissions. This is particularly concerning since SMOD services, like FLEXIBUS, typically use eco-friendly vehicles. However, certain requests, such as early morning rides with short OD-pairs in low-demand areas, are less ecologically desirable for SMOD systems. Without regulation, these requests would be served to maximize service reliability (GAR), but from an ecological perspective, alternative transport options might be more suitable. Serving these requests with SMOD would lead to excessive deadheading, undermining the positive ecological impact.
- *Unnecessary detours*: Beyond these goal conflicts with service reliability and socially equitable mobility, there is another major ecological concern related to MPR: Its current definition in the German Passenger Transport Act (Deutscher Bundestag, 2021) may incentivize unnecessary detours. Once more than one passenger is on board, SMOD providers generate detours instead of taking the shortest route between locations to inflate the MPR through unnecessary PKT. This also contributes to reduced positive ecological impact, as it can be problematic especially at high MPR levels, as indicated by the flattening of the DSP graphs in Fig. 7.

5.2.2. Minimum spatial acceptance rate

Unlike the MPR, the MSAR improves socially equitable mobility without substantial goal conflicts regarding service reliability and ecological impact (Fig. 4, Fig. 9). However, feasibility and impact remain major concerns, as the ability to adhere to MSAR and its impact on socially equitable mobility depend on the spatial partitioning approach and the available fleet size. In the following, we outline and discuss specific issues associated with MSAR application:

- *Feasibility issues*: Similar to the MPR, implementing an MSAR presents feasibility issues for public authorities: Feasible regulation levels are influenced by the customer demand structure and the available vehicle fleet (Section 5.1.2) as well as the spatial partitioning approach. A dedicated feasibility analysis is essential to assess what regulation level is practical for SMOD systems (Fig. 3, Fig. 8). Also, selecting the appropriate time horizon for meeting the MSAR (e.g., daily or weekly) is another critical decision.
- *Impact dependency on choice of spatial partitioning*: Beyond feasibility, the impact of the MSAR is also shaped by the spatial partitioning approach. An ineffective partitioning scheme can still result in inequitable accessibility, with low-demand areas being underserved. For example, the quadrant-based partitioning approach in artificial data (Section 4.2.3) still results in a structural discrimination of peripheral areas (low demand), since only acceptance rates across the quadrants are balanced. However, a tailored alternative partitioning approach could address this imbalance, ensuring better equitable accessibility across both the center and peripheral areas. We discuss the feasibility and impact analysis of the alternative partitioning approach in detail in Appendix D and visualize the direct effects on spatial acceptance rates with heatmaps.
- *Impact dependency of available fleet size:* The MSAR's impact is also dependent on the available fleet size. A larger fleet can easily meet low spatial acceptance rate targets. In contrast, a smaller fleet may struggle to achieve these targets, particularly in low-demand areas. This highlights the importance of considering fleet size when setting MSAR targets, as it directly impacts the

regulatory instrument's impact. For SMOD providers in LSD settings, increasing the number of vehicles could enhance overall service reliability and thereby socially equitable mobility. However, this comes with additional monetary and environmental costs and is a strategic decision. Please note that implementing an MSAR might lead to more expensive operations and ultimately to increased prices for customers, potentially making the service less attractive and socially sustainable. Therefore, the public authority should be aware of the direct consequences of this regulatory instrument, such as the potential need to provide subsidies for serving specific customers or to support the expansion of the vehicle fleet.

5.3. Recommendations

Based on the analysis and identified issues, we propose a number of recommendations for public authorities regarding the practical application of MPR (Section 5.3.1) and MSAR (Section 5.3.2).

5.3.1. Minimum pooling rate

The recommendations for MPR focus on three main areas: strategies for mitigating feasibility issues, addressing issues with the current definition in the German Passenger Transport Act, and proposing a general approach to determine an adequate level of regulation.

Mitigation of feasibility issues: Our numerical results highlight the benefits of multi-period variants for MPR. Extending the number of monitoring periods (e.g., 5P, 10P) expands the solution space, resulting in higher MARLs compared to the single-period variant (1P), with improvements of +18.18% for 5P and +27.27% for 10P (Fig. 6). Thus, while accounting for the demand structure as a key factor, we recommend adopting multi-period variants to mitigate the feasibility issues.

Refining MPR – Eliminating unnecessary detours with bMPR: Throughout this work, we evaluated the precise MPR definition as outlined in the legislation. However, there are alternative definitions worth mentioning. A leg-based pooling rate measures the proportion of shared legs, meaning the percentage of rides where customers share a trip between two locations. A customer-based pooling rate measures the proportion of customers who are pooled with others in the same vehicle. However, like the MPR, all these definitions suffer from the problem of unnecessary detours, as SMOD providers could artificially inflate the pooling rate by non-meaningful rides. As demonstrated in Appendix G, both leg-based and customer-based pooling rates exhibit similar issues.

To overcome this shortcoming, we recommend refining the current MPR definition to address unnecessary detours. We propose the booked MPR (bMPR), where constraint (8), based on the ratio of PKT to VKT, is replaced by constraint (25).

$$\sum_{i\in R} p_i q_i c_{i^+,i^-} \geq \nu^b \sum_{a\in A} x_a c_a \tag{25}$$

Thus, we substitute the PKT $\sum_{a\in A} x_a c_a \omega_a$ with the booked direct OD-pair length $q_i c_{i^+,i^-}$ multiplied by the number of passengers. This ensures that the bMPR ν^b cannot be achieved through unnecessary detours, as only the constant value of $q_i c_{i^+,i^-}$ is added if the request is served ($p_i = 1$). We conducted the computational study of the bMPR using both artificial and real-world data and observed improvements in ecological impact. While a detailed analysis is beyond the scope of this section, further information can be found in Appendix H. Instead, we focus on a comparative study of MPR and bMPR regarding service reliability (GAR) and ecological impact (DSP), as illustrated in Fig. 10.

Service reliability (GAR) is plotted on the ordinate and ecological impact (DSP) on the abscissa, with "x" representing bMPR and "o" representing MPR data points. These points correspond to pooling rates, ranging from the unregulated case to the respective MARL ($\nu_{1P} = 1.10, \nu^b_{1p} = 0.85, \nu_{5P} = 1.30, \nu^b_{5p} = 1.25, \nu_{10P} = 1.40, \nu^b_{10p} = 1.25$). As regulation levels increase, both service reliability (GAR) and ecological impact (DSP) tend to deteriorate. However, the bMPR curve shifts to the right, indicating that the SMOD system achieves better results in both dimensions.

The pareto-optimal regulation levels for both MPR and bMPR, considering service reliability and ecological impact, are highlighted

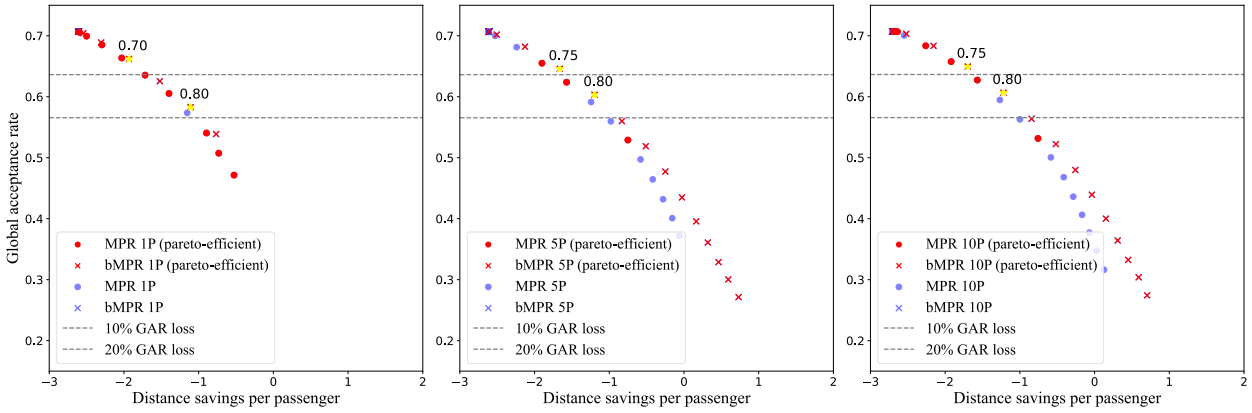

Fig. 10. Pareto analysis on real world data – MPR vs. bMPR.

in red. As the level of regulation increases, bMPR increasingly dominates, improving ecological impact with fewer rejections. This trend is evident across all model variants. This suggests that bMPR actually reduces unnecessary detours, as it benefits from the fixed distance of the booked OD-pair rather than increased PKT. We observe the same trends for bMPR in artificial data across all supply–demand settings. For further details, we refer to Appendix I.

Determining the adequate level of regulation: Recommendations for determining an adequate level of regulation involve balancing service reliability, ecological impact, and socially equitable mobility, depending on the public authority's prioritization of goals for SMOD systems. As discussed in Section 5.2.1, the application of an MPR comes with serious goal conflicts. However, a moderate level of regulation might offer a balanced trade-off. There is no universal "sweet spot", as the optimal level depends largely on the public authority's priorities, which can vary based on the intended purpose of the SMOD system. For example, a public authority might decide that a 10% or 20% reduction in service reliability (GAR) is acceptable for SMOD providers in exchange for substantial improvements in ecological impact (DSP). In Fig. 10, we illustrate the regulation levels that yield the best DSP while maintaining an acceptable GAR deviation range for all model variants, highlighted by yellow stars (-10% GAR : $\rho^b_{1p} = 0.70, \rho^b_{5p} = \rho^b_{10p} = 0.75$; -20% GAR: $\rho^b_{1p} = \rho^b_{5p} = \rho^b_{10P} = 0.80$).

Please note that these regulation levels are relatively moderate and well below the MARL. This is expected, given the severe negative impact on service reliability (Fig. 2a, Fig. 7a) as the regulation level increases. Public authorities should also recognize that improvements in ecological impact (DSP) must be interpreted with caution. Enforced rejections may implicitly reduce ecological benefits, as some rejected customers might switch to less environmentally friendly private vehicles. However, serving requests with certain OD-pairs, particularly in low-demand areas, often leads to excessive deadheading, making them inherently inefficient for an SMOD system. In such cases, rejecting these requests is sensible but might be reasonable.

Furthermore, the pareto analyses (Fig. 10, Appendix H) do not account for the negative impact on socially equitable mobility. This suggests that the recommended regulation level must be even lower, as SDSA tends to deteriorate as ρ increases (Fig. 2c, Fig. 7c). To incorporate socially equitable mobility into the decision-making process, we recommend a goal-weighting approach that reflects the public authority's prioritization.

5.3.2. Minimum spatial acceptance rate

Recommendations for MSAR focus on addressing the mitigation of feasibility issues and determining the adequate level of regulation.

Mitigation of feasibility issues: Analogous to MPR, we observe multi-period MSAR variants outperforming the single-period variant by a $+18.18\%$ increase in MARL (Fig. 8). Thus, we also recommend adopting multi-period variants to mitigate the feasibility issues. Notably, available fleet size and customer demand are key factors influencing the practical application (Section 5.1.2). MSAR, in particular, is highly dependent on the fleet size, allowing SMOD providers to mitigate feasibility issues by expanding their fleets. However, this decision entails a strategic trade-off between better adherence to MSAR, improved service reliability, and enhanced socially equitable mobility, weighed against additional financial and environmental costs.

Determining the adequate level of regulation: The MSAR generally has a positive impact on socially equitable mobility, as it consistently improves SDSA across all experiments using both artificial (Fig. 4c) and real-world data (Fig. 9c) for different partitioning approaches (Appendix D, Appendix E). However, there is a mild negative impact on environmental sustainability (DSP) across all experiments (Fig. 4b, Fig. 9b), which is expected since enforced acceptance of unfavorable requests hinders efficient consolidation. Interestingly, the application of MSAR does not substantially affect service reliability (GAR) (Fig. 4a, Fig. 9a), supporting the idea that MSAR can enhance social sustainability without compromising the SMOD provider's primary goal of high service reliability.

We recommend that public authorities focus on selecting the appropriate spatial partitioning approach, as this directly affects the success of MSAR. Based on the findings from our computational study on artificial and real-world data, the highest possible regulation level, considering feasibility limitations, can be recommended, as it improves socially equitable mobility with minimal sacrifices in service reliability and ecological impact.

6. Conclusion

This section is divided into three parts: First, we provide a summary of our work. Second, we offer managerial implications on the application of regulatory instruments, separately for MPR and MSAR. Finally, we address the limitations of our study and derive possible future research directions.

6.1. Summary

In this paper, we proposed two new MILP formulations for the SDARP that incorporate regulatory instruments inspired by the German Passenger Transport Act (Deutscher Bundestag, 2021) as constraints. In the single-period variant, SDARP-REG, these constraints must be adhered to for each period, whereas in the multi-period variant, M-SDARP-REG, the constraints must be met on an average over a multi-period horizon. We introduced two regulatory instruments which aim at improving environmental and social sustainability in SMOD systems through a minimum pooling rate (MPR) and a minimum spatial acceptance rate (MSAR).

In our computational study, we defined three goals: service reliability, measured by the global acceptance rate (GAR), which captures the proportion of served passengers; ecological impact, quantified by distance savings per passenger (DSP), comparing SMOD to individual motorized transport; and socially equitable mobility, expressed through the standard deviation of spatial acceptance rates

(SDSA), reflecting balanced spatial accessibility to the SMOD system. We conducted an extensive computational study using both artificial and real-world data to critically evaluate the proposed regulatory instruments, focusing on key factors, specific issues, and recommendations in regards with their practical application for public authorities on how to apply them to a specific SMOD system. For artificial data, we differentiated three individual settings for low (LSD), balanced (BSD), and high (HSD) supply–demand ratios to test the robustness of our results. In real-world data, we applied the regulatory instruments in our single- and multi-period variants.

6.2. Managerial implications

Minimum Pooling Rate: Superficially, implementing an MPR might seem promising, but several major issues must be addressed. First, we observe feasibility issues, as achieving high MPRs is limited ($\gamma = 0.95$ for artificial data, $\gamma_{1P} = 1.10$ for real-world data). Thus, we recommend public authorities employing a multi-period enforcement, which may help to alleviate these concerns (Section 5.3.1), with up to +27.27% higher maximum applicable regulation levels (Fig. 6).

Second, another issue involves the conflict to service reliability and socially equitable mobility: As MPR levels increase, requests that are harder to consolidate, particularly those from/to low-demand areas, tend to be rejected, resulting in reduced service reliability and socially equitable mobility.

Third, the deterioration in service reliability can also implicitly worsen the ecological impact, as rejected passengers may switch to less environmentally friendly private vehicles.

Fourth, the current MPR definition in the German Passenger Transport Act (Deutscher Bundestag, 2021) can induce unnecessary detours, further diminishing the positive ecological impact (Fig. 10, Appendix I). To address this, we recommend to implement the booked MPR (bMPR, Section 5.3.1), which mitigates the issue of unnecessary detours. However, even with this adjustment, trade-offs between goals persist.

Therefore, if a pooling rate is to be applied, we recommend a moderate regulation level of bMPR enforced over a multi-period horizon, e.g., for real-world data $\gamma^{b}_{10P} = 0.75$, well below the maximum applicable threshold ($\gamma^{b}_{10P} = 1.25$). This approach may help contain the negative trade-offs within a reasonable range (GAR: −8.33%, DSP: + 35.04%, SDSA: +46.33% for $\gamma^{b}_{10P} = 0.75$), though it does not resolve the inherent conflicts that cannot be ignored.

Minimum Spatial Acceptance Rate: Also implementing the MSAR can offer benefits, but several application issues must be considered. First, public authorities must select a suitable spatial partitioning approach, which heavily depends on the customer demand structure and the available vehicle fleet. As discussed in Section 5.2.2, an ineffective partitioning scheme might still fail to address some unwanted spatial discrimination, e.g., between high-demand (center area) and low-demand (peripheral area) regions (see Appendix D).

Second, feasibility is another critical issue, as regulation levels must be tailored to the specific SMOD system ($\beta_{LSD} = 0.30$, $\beta_{BSD} = 0.60$, and $\beta_{HSD} = 0.85$ for artificial data, $\beta = 0.55$ for real-world data). However, similar to MPR, we recommend applying multi-period enforcement to mitigate feasibility issues (Section 5.3.2), with up to +27.27% higher maximum applicable regulation levels (Fig. 8).

Unlike MPR, MSAR does not exhibit severe goal conflicts, as improving socially equitable mobility does not greatly compromise service reliability or ecological impact, as demonstrated by our results for both artificial data (e.g., GAR: −0.78%, DSP: −2.45%, SDSA −14.80% for $\beta_{LSD} = 0.30$), and real-world data (e.g., GAR: −0.03%, DSP: −2.72%, SDSA: −5.65% for $\beta_{10P} = 0.65$). Therefore, we recommend applying the MARL.

We conclude that MSAR is solely meaningful for providers operating in low or balanced supply–demand settings as the impact is limited for a larger vehicle fleet. Note that providers can improve their adherence to MSAR by expanding their vehicle fleet, though this comes with trade-offs in terms of monetary sacrifices in exchange for higher service reliability and socially equitable mobility.

6.3. Limitations and future research directions

Our findings are restricted by a few limitations, some of which suggest further research directions:

- *Methodological approach:* Our study adopts a static-deterministic modeling perspective, showing the potential extent to which an MPR and an MSAR can be implemented (feasibility analysis) and what their impact might be (impact analysis) if providers make optimal decisions with perfect information. It is important to note that our results can be seen as upper bounds for the stochastic-dynamic problems providers face in the real world, as the operational planning solutions cannot be expected to be as good as the static ones. Future research should focus on developing solution approaches capable of capturing regulatory instruments in a stochastic-dynamic environment.
- *Demand management:* In both single-period and multi-period variants, we relied on a simple accept-reject mechanism. There is likely significant potential to improve MPR and MSAR outcomes by implementing more sophisticated demand management strategies, such as an availability control with time shift mechanism (Anzenhofer et al., 2024), alongside selectivity mechanisms.
- *Enhanced artificial data*: Our computational study on artificial data shows that existing artificial instances lack real-world applicability. We find mismatches in both spatial and temporal dimensions, as desired times and OD pairs are randomly chosen. Future research should focus on developing artificial datasets that reflect real world conditions more closely.
- *Enforcement of regulatory instruments*: While our study analyzes the potential of regulatory instruments to enhance environmental and social sustainability, a key question remains about the practical enforcement of MPR and MSAR in real-world settings. In the

context of the German Passenger Transport Act (Deutscher Bundestag, 2021), it is stipulated that both the responsible authority and the SMOD provider must monitor the impact of such instruments on public transportation and sustainability goals. However, regulatory monitoring requires SMOD providers to share customer and operational data, which could cause unease among providers, as this data contains valuable insights about the service. Note that this also creates a significant burden for providers, as they must invest considerable effort into thoroughly anonymizing and carefully preparing their data before sharing. Further research should explore how regulatory instruments can be effectively enforced by public authorities.

- *Alternative enforcement via incentives or penalties*: Instead of strict regulatory requirements that must be met within a single or multiple periods, an alternative approach could involve using penalties for non-compliance or offering monetary incentives for compliance. This approach would differ from hard regulation levels by allowing providers more flexibility, as they would not be strictly bound to meet specific MPR or MSAR levels. Instead, MPR and MSAR would serve as targets within the provider's economic calculus, where adherence is rewarded, and non-adherence is penalized. This could create a more adaptable and economically motivated framework, encouraging providers to integrate these sustainability goals into their decision-making processes without the rigidity of fixed-period enforcement.

CRediT authorship contribution statement

Fabian Anzenhofer: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Simon Schmidbaur:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis. **Robert Klein:** Writing – review & editing, Supervision, Conceptualization. **Claudius Steinhardt:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We sincerely thank the FLEXIBUS KG, in particular Josef Brandner and Daniel Mayer, for providing access to real-world data used in this study. In addition, we highly appreciate their valuable support, the insightful discussions, and the practical experiences they shared with us. Also, we thank the two anonymous reviewers for their valuable evaluation and comments to further improve this work.

Appendix A

In the following, we briefly define how the time windows for pick-up ($[e_{i^+}, l_{i^+}]$) and drop-off ($[e_{i^-}, l_{i^-}]$) locations of a request $i \in R$ are computed based on the desired time t_i^d (Jaw et al., 1986). To guarantee a certain service level for all requests, we define a waiting time w and a maximum added ride time ϵ to the direct ride time t_{i^+,i^-} from the pick-up location i^+ to the drop-off location i^-.

The time windows for an inbound request are constructed as follows. For the pick-up node, the earliest time is the desired time $e_{i^+} = t_i^d$ and the latest pick-up time is the desired time plus the waiting time: $l_{i^+} = t_i^d + w$. The earliest drop-off time is the desired time plus the direct travel time from pick-up to drop-off, hence $e_{i^-} = t_i^d + t_{i+,i^-}$. The latest drop-off time consists of the latest pick-up time, the direct travel time and the maximum added ride time: $l_{i^-} = l_{i^+} + t_{i^+,i^-} + t_{i^+,i^-} \cdot \epsilon$.

The time windows for an outbound request are defined as follows. The latest drop-off time is the desired time $l_{i^-} = t_i^d$ and the earliest drop-off time is the desired time minus the waiting time: $e_{i^-} = t_i^d - w$. The latest pick-up time is the desired time minus the direct travel time, hence $l_{i^+} = t_i^d - t_{i^+,i^-}$. The earliest pick-up time consists of the earliest drop-off time, the direct travel time and the maximum added ride time: $e_{i^+} = e_{i^-} - t_{i^+,i^-} - t_{i^+,i^-} \cdot \epsilon$. For an overview of the time windows, we refer to Table A.1.

Table A1
Calculation of time windows.

	Outbound Request	Inbound Request
$i^+ \in P$	$e_{i^+} = e_{i^-} - t_{i^+,i^-} - \epsilon \cdot t_{i^+,i^-}$	$e_{i^+} = t_i^d$
	$l_{i^+} = l_{i^-} - t_{i^+,i^-}$	$l_{i^+} = t_i^d + w$
$i^- \in D$	$e_{i^-} = t_i^d - w$	$e_{i^-} = t_i^d + t_{i^+,i^-}$
	$l_{i^-} = t_i^d$	$l_{i^-} = l_{i^+} + t_{i^+,i^-} + \epsilon \cdot t_{i^+,i^-}$

Appendix B

Table B1
Notations for sets, parameters and decision variables.

Element	Description
$a \in A = \bigcup_{i=1}^{6} A_i$	Set of Arcs
$B_j \geq 0$	Decision Variable indicating the Start of Service at Location j
β	Minimum Spatial Acceptance Rate
$c_{i,j}$	Travel Distance between Locations i and j
$c_{v,w}$	Travel Distance between Events v and w
$D = \{i^-, \cdots, n^-\}$	Set of Drop-off Locations
$[e_j, l_j]$	Time Window of Location j
ϵ	Maximum Additional Ride Time Factor
$\delta^{in}(v)$	Set of Incoming Arcs of Event v
$\delta^{out}(v)$	Set of Outgoing Arcs of Event v
$G = (V, A)$	Event-based Graph
H	Number of Monitoring Periods
$h \in \mathcal{H} = \{1, \cdots, H\}$	Set of Monitoring Periods
K	Number of Vehicles
$k \in \mathcal{K} = \{1, \cdots, K\}$	Set of Vehicles
$j \in J = P \cup D \cup \{0^+, 0^-\}$	Set of Locations
L_i	Maximum Ride Time of Request i
$M_{i,j}$	Big-M Parameter for Locations i and j
n	Number of Requests
$\sigma \in \mathcal{C} = \{1, \cdots \Sigma\}$	Set of Subsets
ω_a	Number of Transported Passengers during Arc a
μ, μ^b	Minimum Pooling Rate, booked Minimum Pooling Rate
Q	Maximum Number of Request in a Vehicle
$\mathcal{Q}$	Maximum Passenger Capacity of a Vehicle
q_i	Number of Passengers for Request i
$P = \{1^+, \cdots, n^+\}$	Set of Pick-up Locations
$p_i \leq 1$	Decision Variable indicating whether Request i is served
$i \in R = \{1, \cdots, n\}$	Set of Requests
R_σ	Set of Requests of Subset σ
s_i	Service Time of Request i
Σ	Number of Subsets
T	Service Horizon
$t_{i,j}$	Travel Time between Locations i and j
t_i^d	Desired Time of Request i
$x_a \in \{0, 1\}$	Decision Variable indicating whether Arc a is used
$v \in V = V_0 \cup \bigcup_{i=1}^{n} V_{i^+} \cup \bigcup_{i=1}^{n} V_{i^-}$	Set of Events
w	Waiting Time

Appendix C

In the following, we illustrate with a didactic example how a SMOD provider's ability to achieve high minimum pooling rates largely depends on the demand structure. Let $i \in R = \{1, 2\}$ be two requests with $q_1 = q_2 = 1$ and identical pick-up and drop-off locations, served by a single vehicle (fleet size of $K = 1$). Figure C.1 depicts the pick-up and drop-off locations for both requests, a depot denoted as 0, and the uniform distances between them.

Given these distances, the maximum achievable regulation level for MPR is $\frac{PKT}{VKT} = \frac{2 \cdot 1}{1+1+1} = \frac{2}{3}$ even though both requests are efficiently pooled. Thus, MPR is ultimately constrained by unavoidable empty trips, such as those to or from the depot. This limits achievable pooling rates, which can approach zero depending on the distances between the depot and customer locations, even when customers share the entire ride.

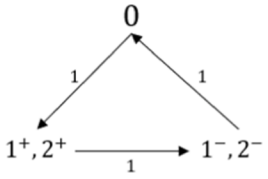

Fig. C1. Example of limited pooling capability

Appendix D

As the numerical results for MSAR for artificial data have shown, there are clear trends related to spatial discrimination and how MSAR can mitigate it. However, as discussed in Section 4.2.3, we also explore MSAR using an alternative partitioning approach with a center and a peripheral area that is not symmetric with uniformly distributed requests. As a remainder we divide the requests into two subsets: the center area ($\sigma = 1$) forms a square around the depot at $(0, 0)$ with edge lengths of 15, while the peripheral area ($\sigma = 2$) covers the outer area. A request is assigned to the center area if the pick-up location lies within $(|x| \leq 7.5) \wedge (|y| \leq 7.5)$, and to the peripheral area if it falls outside this boundary $(|x| > 7.5) \vee (|y| > 7.5)$. An illustration of the partititon is provided in Figure D.1. We discuss the results for feasilibity and impact in the following analyses.

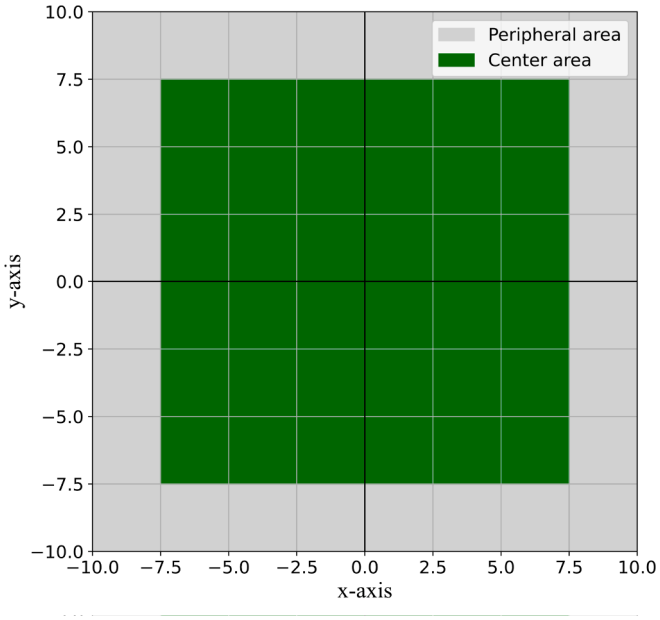

Fig. D1. Alternative partitioning for MSAR (artificial data)

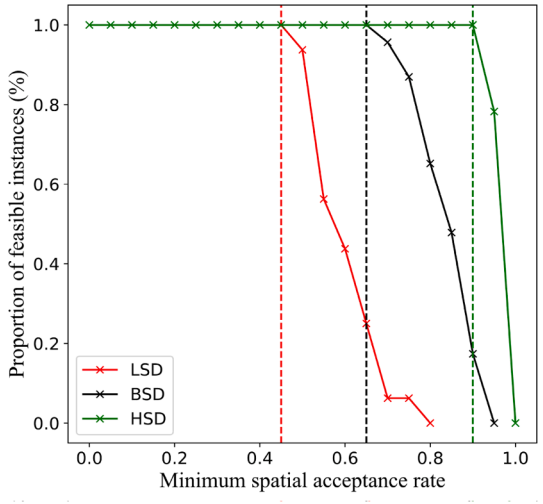

Fig. D2. Feasibility analysis on artificial data – Alternative partitioning approach for MSAR

Feasibility analysis: In Figure D.2, we observe similar patterns to those in Fig. 3, with the feasibility graphs highly dependent on the three supply–demand settings. The MARLs are $\beta_{LSD} = 0.45$, $\beta_{BSD} = 0.65$, and $\beta_{HSD} = 0.90$, represented by the vertical dashed lines. Notably, this partitioning approach allows the provider to achieve higher MARLs for each setting compared to the experiments in Section 4.2.3. This could be because MSAR only needs to be satisfied for two subsets, rather than four as in the original partitioning.

Impact analysis: We limit the discussion of the numerical results up to the MARLs in Figure D.3.

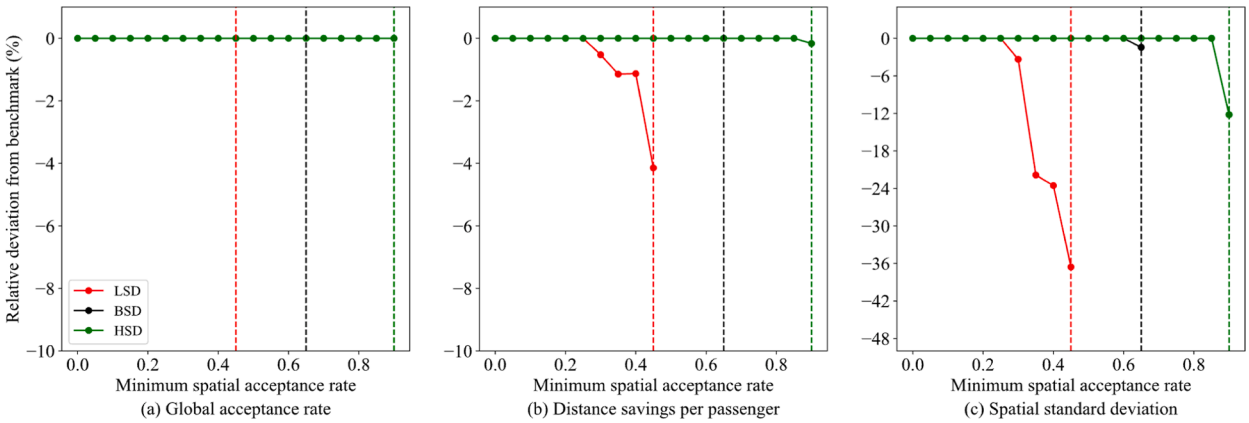

Fig. D3. Impact analysis on artificial data – Alternative partitioning approach for MSAR

- *Impact – GAR*: Figure D.3a illustrates that increasing β has no impact on GAR, which remains constant across all supply–demand settings. This suggests that any impact on the other two metrics (DSP, SDSA) does not result in a loss of GAR. The provider swaps specific requests to balance acceptance rates between both areas, while maintaining the same number of passengers served.
- *Impact – DSP*: The DSP graph for LSD (Figure D.3b) shows a consistent negative deviation (-4.14% for $\beta_{LSD} = 0.45$), providing further evidence that DSP worsens, particularly with a limited fleet size. In contrast, the DSP for BSD and HSD remain nearly constant.
- *Impact – SDSA*: For SDSA, improvements are observed across all settings (Figure D.3c). LSD, in particular, shows significant potential for MSAR, with notable improvements in SDSA (-36.57% for $\beta_{LSD} = 0.45$). BSD (-1.45% for $\beta_{BSD} = 0.65$) and HSD (-12.21% for $\beta_{HSD} = 0.90$) also show improvements, but to a lesser extent.

To further illustrate the impact of implementing the MARL (LSD) with both partitioning approaches ($\beta_{LSD} = 0.3$ for quadrant-based partitioning, and $\beta_{LSD} = 0.45$ for alternative partitioning) compared to the unregulated case, we use heat maps in Figure D.4. The service area is divided into granular subareas (2.5×2.5), creating a chessboard-like pattern with 64 individual regions. The heat map shows the average spatial acceptance rate for each request, based on whether the request's origin lies in a specific subarea, across all artificial instances.

Examining the heat maps, we observe shifts in acceptance rates: in Figure D.4b, acceptance rates become more balanced across the four quadrants, particularly pronounced in the lower-left quadrant. Clear patterns also emerge in Figure D.4c, where the outskirts show higher average acceptance rates compared to the other two scenarios. Please note that these heat maps do not reflect the frequency of requests, as there may be imbalances in demand intensity across subareas. However, since request locations are drawn from a uniform distribution, we do not believe that this is a significant issue.

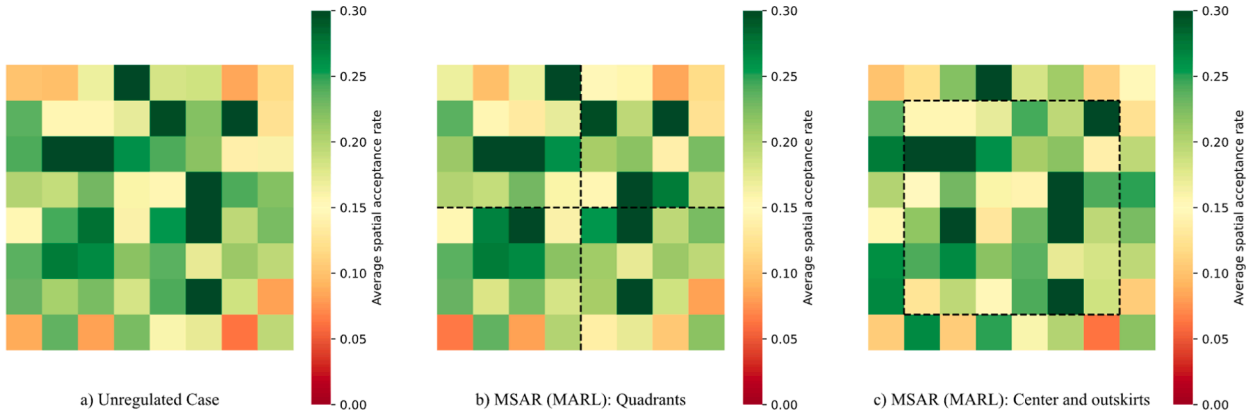

Fig. D4. Heat maps on spatial acceptance rates (MSAR, MARL) for different partitioning approaches – Artificial data

In our work, we examine only two spatial partitioning approaches for artificial data. However, we recognize that there exist other, potentially more effective partitioning approaches. For instance, further dividing the outskirts into two or even four distinct areas could achieve more balanced minimum spatial acceptance rates at a higher granularity. However, as the number of spatial subsets increases, so does the trade-off with a higher likelihood of feasibility issues in some instances, which may lead to smaller and thus less impactful MARLs. We suggest that future research could explore a range of spatial partitioning approaches to examine this inherent trade-off on a general level.

Appendix E

As an alternative partitioning approach for MSAR in real-world data, we exemplary define five subsets based on the descriptive analysis in Fig. 5: north-west ($\sigma = 1$), north-east ($\sigma = 2$), center ($\sigma = 3$), south-west ($\sigma = 4$), south-east ($\sigma = 5$). Each request is assigned to one of these subsets solely based on their pick-up location (origin-based partitioning). An illustration of the partititon is provided in Figure E.1. We present the graphs for the feasibility analysis and impact analysis in Figs. E.2 and E.3.

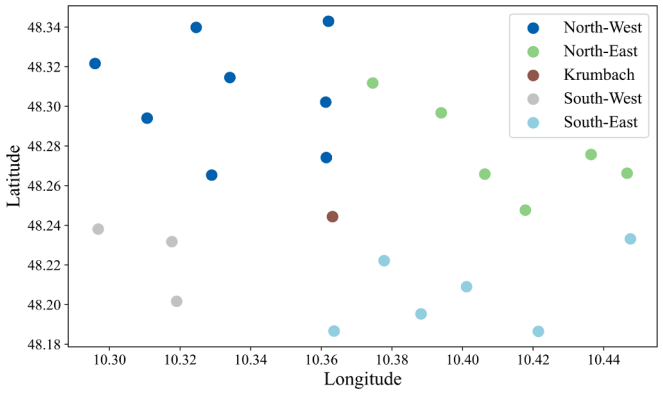

Fig. E1. Alternative partitioning for MSAR (real-world data)

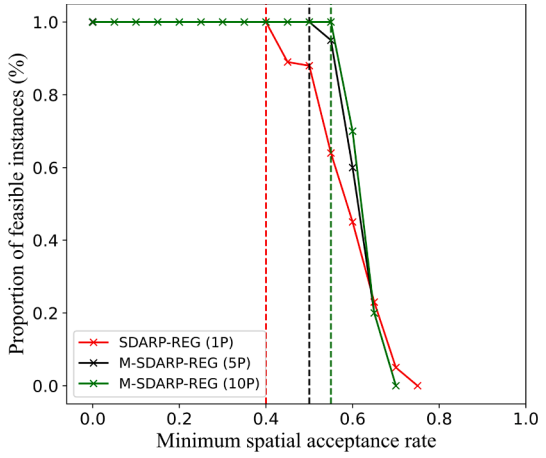

Fig. E2. Feasibility analysis on real-world data – Alternative partitioning approach for MSAR

Feasibility analysis: In Figure E.2, we observe similar steep curves for the feasibility graphs as in Fig. 8. The MARLs are $\beta_{1P} = 0.40, \beta_{5P} = 0.50, \beta_{10P} = 0.55$, which again emphasizes the potential of multi-period variants. All MARLs lie under the respective MARLs from Fig. 8 due to the refined subsets. This indicates a trade-off between higher MARLs and the respective feasibility.

Impact analysis: At first glance, it is immediately apparent that this spatial partitioning scheme has a very limited impact. All three target metrics (GAR, DSP, and SDSA) remain within a moderate range. From this, we can conclude that, at least up to the MARLs, no significant impact can be observed, and the spatial subsets appear relatively balanced up to this point.

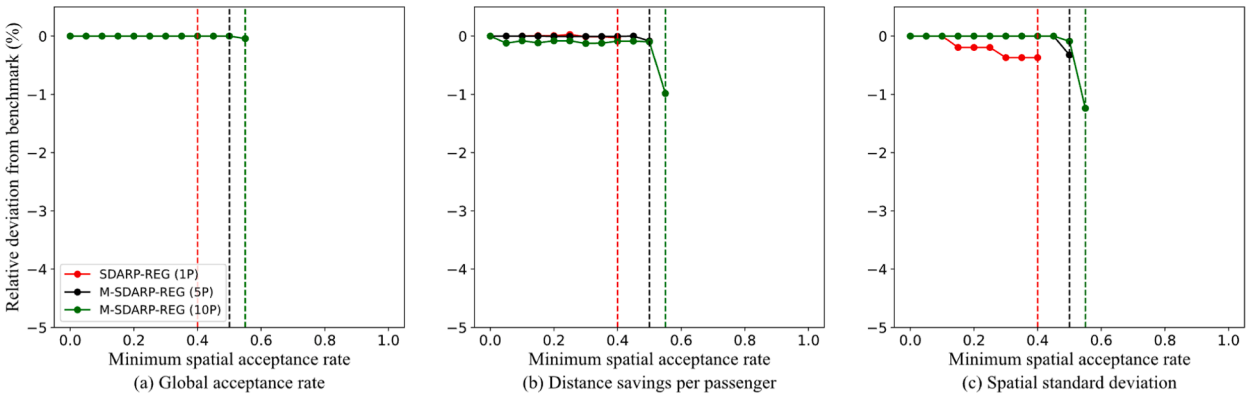

Fig. E3. Feasibility analysis on real-world data – Alternative partitioning approach for MSAR

Appendix F

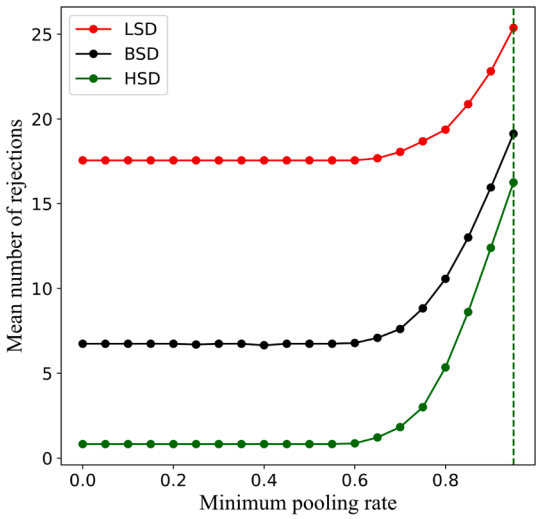

Fig. F1. Analysis of rejections (artifical data)

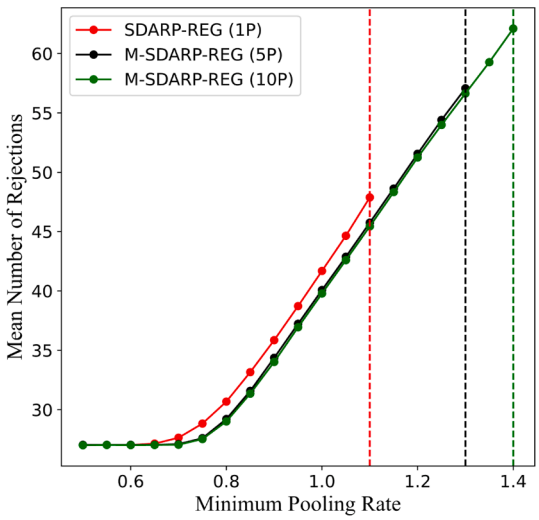

Fig. F2. Analysis of rejections (real-world data)

Appendix G

Let $i \in R = \{1, 2\}$ be two requests with $q_1 = q_2 = 1$ which are served by a fleet size of $K = 1$. We assume that, due to time constraints, $i = 1$ must be picked up first by the vehicle and both requests must be served. We provide an illustration of the example in Figure G.1 including distances between the pick-up and drop-off locations, respectively.

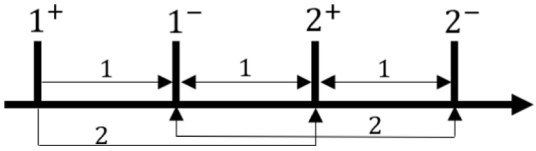

Fig. G1. Didactic example

There are three potential route plans:

a) $1^+ \rightarrow 1^- \rightarrow 2^+ \rightarrow 2^-$ (3 VKT)
b) $1^+ \rightarrow 2^+ \rightarrow 1^- \rightarrow 2^-$ (5 VKT)
c) $1^+ \rightarrow 2^+ \rightarrow 2^- \rightarrow 1^-$ (5 VKT)

Obviously, the ecological reasonable route plan is a), as it minimizes the VKT. Now, we evaluate the following four different approaches to define a pooling rate:

1) MPR (German Passenger Transport Act): $\frac{PKT}{VKT}$
2) Leg-based MPR: $\frac{SharedLegs}{DrivenLegs}$
3) Customer-based MPR: $\frac{Customerssharingaride}{ServedCustomers}$
4) bMPR: $\frac{bookedODdistance}{VKT}$

In Table G.1, we observe that definitions 1–3 might benefit from unnecessary rides to inflate the respective pooling rate, as requests 1 and 2 share part of their route, which negatively impacts ecological impact. In contrast, 4) (bMPR) does not benefit from these additional VKT, as serving both requests only adds a fixed value based on the booked direct OD distances.

Table G1
Comparison of MPR definitions.

Route	MPR	Leg-based MPR	Customer-based MPR	bMPR
$1^+ \rightarrow 1^- \rightarrow 2^+ \rightarrow 2^-$	$\frac{2}{3}$	$\frac{0}{3}$	$\frac{0}{2}$	$\frac{2}{3}$
$1^+ \rightarrow 2^+ \rightarrow 1^- \rightarrow 2^-$	$\frac{6}{5}$	$\frac{1}{3}$	$\frac{2}{2}$	$\frac{2}{5}$
$1^+ \rightarrow 2^+ \rightarrow 2^- \rightarrow 1^-$	$\frac{6}{5}$	$\frac{1}{3}$	$\frac{2}{2}$	$\frac{2}{5}$

Appendix H

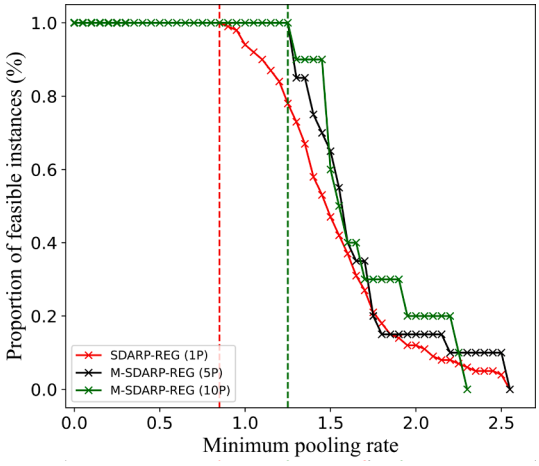

Fig. H1. Feasibility analysis on real-world data – bMPR

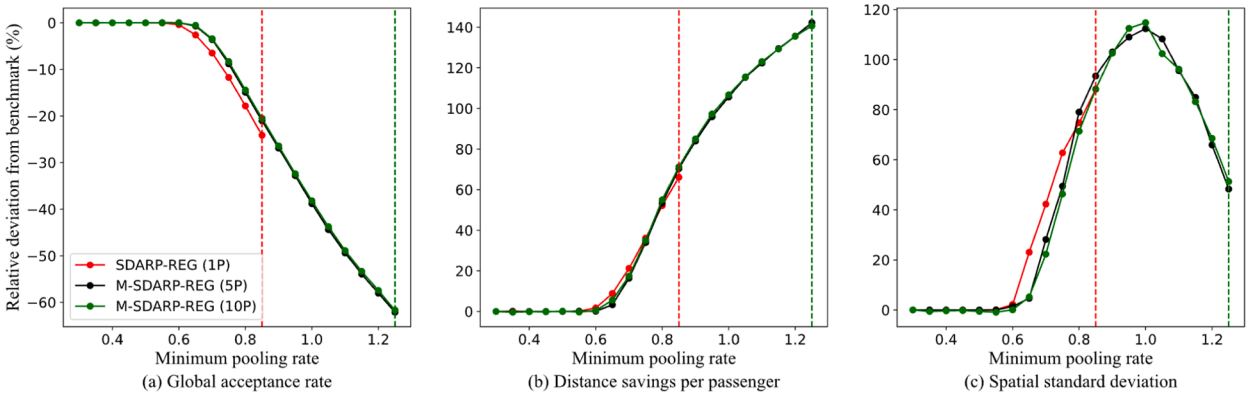

Fig. H2. Impact analysis on real-world data – bMPR

Appendix I

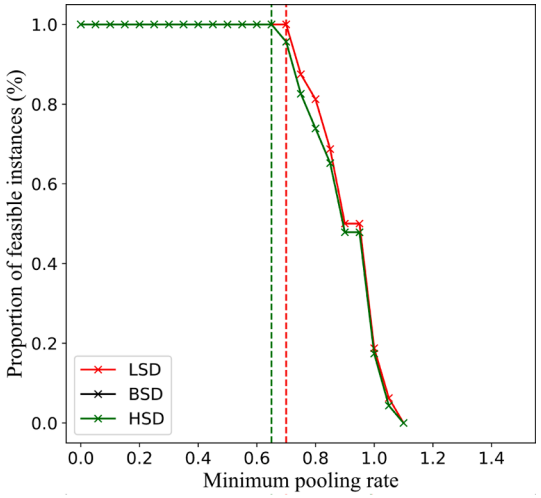

Fig. I1. Feasibility analysis on artificial data – bMPR

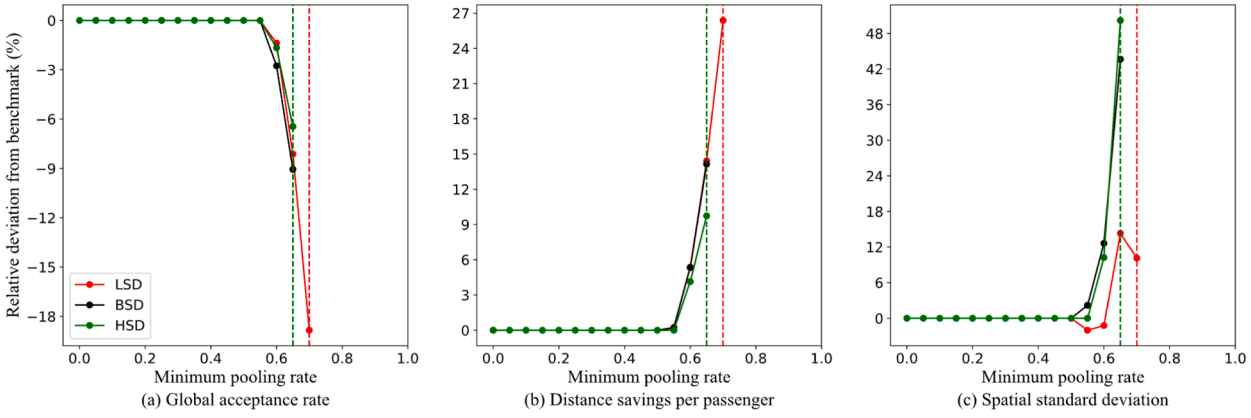

Fig. I2. Impact analysis on artificial data – bMPR

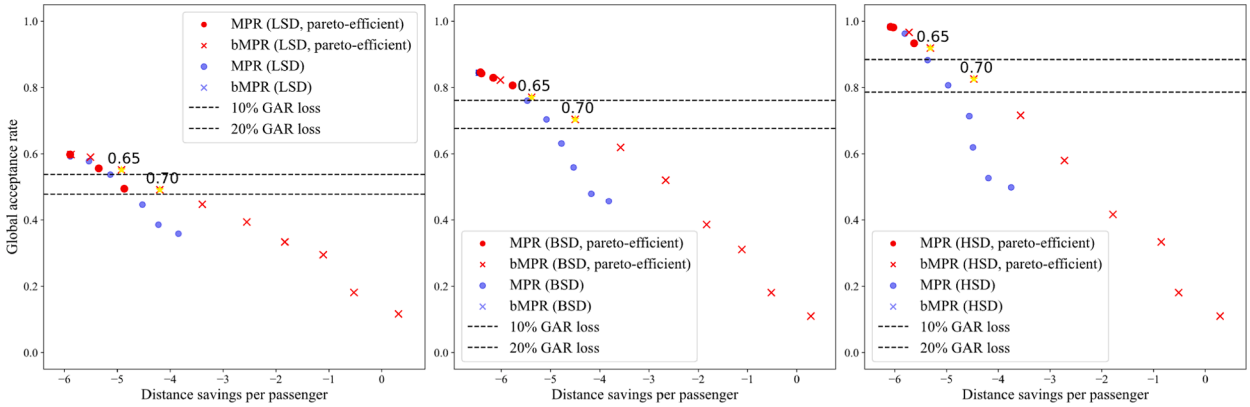

Fig. I3. Pareto analysis on artificial data – bMPR vs. MPR

Appendix J

Table J1
List of Abbreviations.

Abbreviation	Description
BSD	Balanced Supply-Demand
DARP	Dial-a-Ride Problem
DSP	Distance Savings per Passenger
GAR	Global Acceptance Rate
HSD	High Supply-Demand
LSD	Low Supply-Demand
MOD	Mobility-on-Demand
MARL	Maximum Applicable Regulation Level
MILP	Mixed-Integer Linear Program
(b)MPR	(booked) Minimum Pooling Rate
MSAR	Minimum Spatial Acceptance Rate
M−SDARP−REG	Multi-period Selective Dial-a-Ride Problem with Regulatory Instruments
OD	Origin-Destination
PKT	Passenger Kilometers Traveled
SDARP	Selective Dial-a-Ride Problem
SDARP-REG	Selective Dial-a-Ride Problem with Regulatory Instruments
SDSA	Standard Deviation of Spatial Acceptance Rates
SMOD	Shared Mobility-on-Demand
VKT	Vehicle Kilometers Traveled
VRP	Vehicle Routing Problem

Data availability

The data that has been used is confidential.

References

Anzenhofer, F., Fleckenstein, D., Klein, R., Steinhardt, C., 2024. Analyzing the Impact of Demand Management in Rural Shared Mobility-on-Demand Systems. https://doi.org/10.2139/ssrn.4682056.

Berbeglia, G., Cordeau, J.-F., Gribkovskaia, I., Laporte, G., 2007. Static pickup and delivery problems: a classification scheme and survey. TOP 15, 1–31. https://doi.org/10.1007/s11750-007-0009-0.

Berbeglia, G., Pesant, G., Rousseau, L.-M., 2011. Checking the feasibility of dial-a-ride instances using constraint programming. Transp. Sci. 45, 399–412. https://doi.org/10.1287/trsc.1100.0336.

Bösch, P.M., Becker, F., Becker, H., Axhausen, K.W., 2018. Cost-based analysis of autonomous mobility services. Transp. Policy 64, 76–91. https://doi.org/10.1016/j.tranpol.2017.09.005.

Braekers, K., Kovacs, A.A., 2016. A multi-period dial-a-ride problem with driver consistency. Transp. Res. B Methodol. 94, 355–377. https://doi.org/10.1016/j.trb.2016.09.010.

Cai, H., Wang, X., Adriaens, P., Xu, M., 2019. Environmental benefits of taxi ride sharing in Beijing. Energy 174, 503–508. https://doi.org/10.1016/j.energy.2019.02.166.

Camarero, L., Oliva, J., 2019. Thinking in rural gap: mobility and social inequalities. Palgrave Commun 5, 95. https://doi.org/10.1057/s41599-019-0306-x.

Cauchi, M., Scerri, K., 2020. An improved variable neighbourhood search algorithm for selective dial-a-ride problems. In: 20th Mediterranean Electrotechnical Conference (MELECON). IEEE, pp. 652–657. https://doi.org/10.1109/MELECON48756.2020.9140695.
Chen, M.H., Jauhri, A., Shen, J.P., 2017. Data driven analysis of the potentials of dynamic ride pooling. In: Proceedings of the 10th ACM SIGSPATIAL Workshop on Computational Transportation Science, pp. 7–12. https://doi.org/10.1145/3151547.3151549.
Cordeau, J.-F., 2006. A branch-and-cut algorithm for the dial-a-ride problem. Oper. Res. 54, 573–586. https://doi.org/10.1287/opre.1060.0283.
Cordeau, J.-F., Laporte, G., 2007. The dial-a-ride problem: models and algorithms. Annals of Operations Research 153, 29–46. https://doi.org/10.1007/s10479-007-0170-8.
Curtis, C., Stone, J., Legacy, C., Ashmore, D., 2019. Governance of future urban mobility: a research agenda. Urban Policy Res. 37, 393–404. https://doi.org/10.1080/08111146.2019.1626711.
Dandl, F., Engelhardt, R., Hyland, M., Tilg, G., Bogenberger, K., Mahmassani, H.S., 2021. Regulating mobility-on-demand services: tri-level model and bayesian optimization solution approach. Transp. Res. Part C Emerging Technol. 125, 103075. https://doi.org/10.1016/j.trc.2021.103075.
Deutscher Bundestag, 2021. Personenbeförderungsgesetz [WWW Document] accessed 1.11.23 https://www.gesetze-im-internet.de/pbefg/.
Engelhardt, R., Dandl, F., Bilali, A., Bogenberger, K., 2019. Quantifying the benefits of autonomous on-demand ride-pooling: a simulation study for Munich, Germany. In: Intelligent Transportation Systems Conference (ITSC). IEEE, pp. 2992–2997. https://doi.org/10.1109/ITSC.2019.8916955.
Ennen, D., Heilker, T., 2020. Ride-hailing services in Germany: potential impacts on public transport, motorized traffic, and social welfare. Institut für Verkehrswissenschaft.
Erhardt, G.D., Roy, S., Cooper, D., Sana, B., Chen, M., Castiglione, J., 2019. Do transportation network companies decrease or increase congestion? Sci. Adv. 5, eaau2670. https://doi.org/10.1126/sciadv.aau2670.
Flexibus, 2024. Der FLEXIBUS in Krumbach [WWW Document]. Flexibus. URL https://www.flexibus.net/krumbach (accessed 9.11.24).
Gaul, D., Klamroth, K., Stiglmayr, M., 2022. Event-based MILP models for ridepooling applications. Eur. J. Oper. Res. 301, 1048–1063. https://doi.org/10.1016/j.ejor.2021.11.053.
Gaul, D., Klamroth, K., Pfeiffer, C., Stiglmayr, M., Schulz, A., 2025. A tight formulation for the dial-a-ride problem. Eur. J. Oper. Res. 321, 363–382. https://doi.org/10.1016/j.ejor.2024.09.028.
Gurobi, 2024. Version 11.0.0 [WWW Document]. Gurobi Optimization. URL https://www.gurobi.com/ (accessed 5.14.24).
Ho, S.C., Szeto, W.Y., Kuo, Y.-H., Leung, J.M.Y., Petering, M., Tou, T.W.H., 2018. A survey of dial-a-ride problems: literature review and recent developments. Transp. Res. B Methodol. 111, 395–421. https://doi.org/10.1016/j.trb.2018.02.001.
Hosni, H., Naoum-Sawaya, J., Artail, H., 2014. The shared-taxi problem: formulation and solution methods. Transp. Res. B Methodol. 70, 303–318. https://doi.org/10.1016/j.trb.2014.09.011.
Hungerländer, P., Maier, K., Pachatz, V., Truden, C., 2021. Improving sharing rates of a dial-a-ride problem implemented for an Austrian mobility provider. Transp. Res. Procedia 52, 525–532. https://doi.org/10.1016/j.trpro.2021.01.062.
Jaw, J.-J., Odoni, A.R., Psaraftis, H.N., Wilson, N.H.M., 1986. A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows. Transp. Res. B Methodol. 20, 243–257. https://doi.org/10.1016/0191-2615(86)90020-2.
Kirchler, D., Wolfler Calvo, R., 2013. A granular tabu search algorithm for the dial-a-ride problem. Transp. Res. B Methodol. 56, 120–135. https://doi.org/10.1016/j.trb.2013.07.014.
Liobikienė, G., Miceikienė, A., 2022. Influence of informational, social, convenience and financial tools on sustainable transport behaviour: the case of Lithuania. J. Clean. Prod. 362, 132457. https://doi.org/10.1016/j.jclepro.2022.132457.
Ma, Z., Koutsopoulos, H.N., 2022. Near-on-demand mobility. The benefits of user flexibility for ride-pooling services. Transp. Res. Part C Emerging Technol. 135, 103530. https://doi.org/10.1016/j.trc.2021.103530.
Martinez, L.M., Viegas, J.M., 2017. Assessing the impacts of deploying a shared self-driving urban mobility system: an agent-based model applied to the city of Lisbon, Portugal. Int. J. Transp. Sci. Technol. 6, 13–27. https://doi.org/10.1016/j.ijtst.2017.05.005.
Molenbruch, Y., Braekers, K., Caris, A., 2017. Typology and literature review for dial-a-ride problems. Annals of Operations Research 259, 295–325. https://doi.org/10.1007/s10479-017-2525-0.
Parragh, S.N., Pinho De Sousa, J., Almada-Lobo, B., 2015. The dial-a-ride problem with split requests and profits. Transp. Sci. 49, 311–334. https://doi.org/10.1287/trsc.2014.0520.
Pillac, V., Gendreau, M., Guéret, C., Medaglia, A.L., 2013. A review of dynamic vehicle routing problems. Eur. J. Oper. Res. 225, 1–11. https://doi.org/10.1016/j.ejor.2012.08.015.
Powell, W.B., 2011. Approximate Dynamic Programming: Solving the Curses of Dimensionality, 1st ed, Wiley Series in Probability and Statistics. Wiley. https://doi.org/10.1002/9781118029176.
Python, 2024. Version 3.9 [WWW Document]. Python.org. URL https://www.python.org/ (accessed 5.14.24).
Qiu, H., Li, R., Zhao, J., 2018. Dynamic Pricing in Shared Mobility on Demand Service.
Reinhardt, L.B., Clausen, T., Pisinger, D., 2013. Synchronized dial-a-ride transportation of disabled passengers at airports. Eur. J. Oper. Res. 225, 106–117. https://doi.org/10.1016/j.ejor.2012.09.008.
Riedler, M., Raidl, G., 2018. Solving a selective dial-a-ride problem with logic-based benders decomposition. Comput. Oper. Res. 96, 30–54. https://doi.org/10.1016/j.cor.2018.03.008.
Rist, Y., Forbes, M.A., 2021. A New Formulation for the Dial-a-Ride Problem. Transp. Sci. 55, 1113–1135. https://doi.org/10.1287/trsc.2021.1044.
Ropke, S., Cordeau, J., Laporte, G., 2007. Models and branch-and-cut algorithms for pickup and delivery problems with time windows. Networks 49, 258–272. https://doi.org/10.1002/net.20177.
Shaheen, S., Cohen, A., 2020. Mobility on demand (MOD) and mobility as a service (MaaS): early understanding of shared mobility impacts and public transit partnerships, in: Demand for Emerging Transportation Systems. Elsevier, pp. 37–59. https://doi.org/10.1016/B978-0-12-815018-4.00003-6.
Shaheen, S., Martin, E., Cohen, A., Broader, J., Davis, R., 2022. Managing the curb: understanding the impacts of on-demand mobility on public transit, micromobility, and pedestrians. https://doi.org/10.31979/mti.2022.1904.
Sörensen, L., Bossert, A., Jokinen, J.-P., Schlüter, J., 2021. How much flexibility does rural public transport need? – Implications from a fully flexible DRT system. Transp. Policy 100, 5–20. https://doi.org/10.1016/j.tranpol.2020.09.005.
VDV, 2023. Linienbedarfsverkehr: zukunftsgerecht, integriert und nachfragegesteuert [WWW Document]. URL https://www.vdv.de/positionensuche.aspx?id=ecb3bf26-1e30-42ba-80b9-579f839075a9&mode=detail&coriander=V3_43e70abd-7d56-e24f-47b1-ba66f00c017d.
Wolfler Calvo, R., Colorni, A., 2007. An effective and fast heuristic for the dial-a-ride problem. 4OR 5, 61–73. https://doi.org/10.1007/s10288-006-0018-0.
Yu, B., Ma, Y., Xue, M., Tang, B., Wang, B., Yan, J., Wei, Y.-M., 2017. Environmental benefits from ridesharing: A case of Beijing. Appl. Energy 191, 141–152. https://doi.org/10.1016/j.apenergy.2017.01.052.
Zhang, Z., Liu, M., Lim, A., 2015. A memetic algorithm for the patient transportation problem. Omega 54, 60–71. https://doi.org/10.1016/j.omega.2015.01.011.
Zwick, F., Axhausen, K.W., 2022. Ride-pooling demand prediction: a spatiotemporal assessment in Germany. J. Transp. Geogr. 100, 103307. https://doi.org/10.1016/j.jtrangeo.2022.103307.