European Journal of Operational Research 307 (2023) 860–886

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Dynamic demand management and online tour planning for same-day delivery

Vienna Klein[1,*], Claudius Steinhardt[1]

Chair of Business Analytics & Management Science, University of the Bundeswehr Munich, Germany

ARTICLE INFO

Article history:
Received 19 November 2021
Accepted 9 September 2022
Available online 14 September 2022

Keywords:
Logistics
Same-day delivery
Last-Mile logistics
Demand management
Approximate dynamic programming

ABSTRACT

For providers to stay competitive in a context of continued growth in e-retail sales and increasing customer expectations, same-day delivery options have become very important. Typically, with same-day delivery, customers purchase online and expect to receive their ordered goods within a narrow delivery time span. Providers thus experience substantial operational challenges to run profitable tours and generate sufficiently high contribution margins to cover overhead costs. We address these challenges by combining a demand-management approach with an online tour-planning approach for same-day delivery. More precisely, in order to reserve capacity for high-value customer orders and to guide customer choices toward efficient delivery operations, we propose a demand-management approach that explicitly optimizes the combination of delivery spans and prices which are presented to each incoming customer request. The approach includes an anticipatory sample-scenario based value approximation, which incorporates a direct online tour-planning heuristic. It does not require extensive offline learning and is scalable to realistically sized instances with multiple vehicles. In a comprehensive computational study, we show that our anticipatory approach can improve the contribution margin by up to 50% compared to a myopic benchmark approach. We also show that solving an explicit pricing optimization problem is a beneficial component of our approach. More precisely, it outperforms both a pure availability control and a simple pricing rule based on opportunity costs. The latter idea is one used in other approaches for related dynamic pricing problems dealt with in the literature.

1. Introduction

From 2014 to 2019, global retail e-commerce sales nearly tripled, and they are forecast to nearly double over the next few years to reach an expected USD 7.4 trillion in sales volume by 2025 (eMarketer, 2022). Many e-retail providers started out by offering same-day delivery (SDD), meaning that customers could shop online and receive the ordered goods on the very same day, typically within the next few hours, depending on the provider. For customers, such fast delivery brings instant gratification similar to shopping in brick-and-mortar stores, with the added convenience of online shopping. For this reason, the majority of customers are willing to pay higher fees for faster delivery (McKinsey & Company, 2016; PwC, 2018). However, despite high demand and the customers' willingness to pay more for faster delivery, many SDD-providers went out of service or shifted their service portfolio toward different business segments. This is because with SDD, they were not able to maintain profitable delivery operations and were not able to achieve sufficiently high overall contribution margins to cover their overhead costs. Thus, e-retail providers need to improve the profitability of their delivery operations in order to live up to increasing customer expectations regarding delivery speed.

Known from related fields of research regarding home delivery, there is two measures to increase the respective profitability. Those are (1) optimizing tour planning and (2) optimizing demand management. With the latter, it is possible to avoid unfavorable requests without loosing customer goodwill by making informed decisions on which delivery options and prices to offer each customer. This additionally holds potential to further improve tour planning with regard to profitability. Both measures have successfully been applied to related fields of research. The most related of those fields is attended home delivery (AHD), where customers have to be present when their goods are delivered. However, from a theoretical point of view, both measures, i.e., demand

* Corresponding author.
E-mail address: vienna.klein@unibw.de (V. Klein).

[1] This work was supported by the University of the Bundeswehr Munich.

https://doi.org/10.1016/j.ejor.2022.09.011

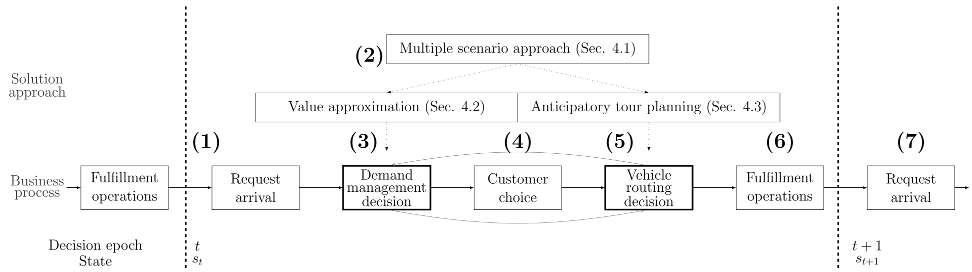


Fig. 1. Overview of the interaction of the considered business process, the respective MDP model, and our proposed solution approach.

management as well as tour planning, are substantially more difficult to optimize for SDD than for the broadly investigated AHD problems. This is due to the overlap of booking and service periods in SDD, which is typically not assumed for AHD. This overlap causes the necessity of closely integrating both previously named optimization measures with each other, with the additional requirement for the tour-planning optimization to be conducted online. More precisely, contrary to AHD, with SDD, not only the decision on which delivery options to offer at which prices must be made online, but also the decision on which orders to allocate to which tours and when to start each. Additionally, with SDD, both decisions have to be made under anticipation of potential future decisions. However, in the related literature, there only exist works that tackle the SDD problem in an anticipatory manner either with regard to optimizing the demand management, or with regard to optimizing the online tour planning. To the best of our knowledge, there is no approach that explicitly incorporates anticipation holistically for both components.

In this paper, we consider an SDD problem setting and approach increasing its profitability by holistically optimizing demand management and tour planning in an integrated manner. We thus refer to the problem under consideration as the SDD demand-management and tour-planning problem (SDD-DMTP). In particular, we approach such an SDD-DMTP from the perspective of some typical middle sized e-retailer offering the delivery of goods at the same day to a registered pool of customers in an urban area with a small number of delivery vehicles. We aim to make the concept of SDD profitable and to improve provider services and thus customer satisfaction. We do so by exploiting two demand-management levers, namely reserving more capacity for higher valued customers and guiding the stochastic customer choice toward efficient delivery options. Simultaneously, we improve online tour planning. Methodologically, we model the problem holistically as Markov decision process (MDP) and present a forward approximate dynamic programming (ADP) optimization approach (Powell, Simao, & Bouzaiene-Ayari, 2012) for its solution. Within the ADP approach, we combine ideas of multiple scenario approaches for online tour planning with the ideas of value approximation via sampled trajectories, such as those known from rollout algorithms/Monte Carlo methods (see for example Sutton & Barto, 2018). Fig. 1 shows the interaction of our proposed optimization approach with the SDD booking and service process of an e-retailer (in the following referred to as *provider*). The lower stream shows the actual business process of the provider, and the upper stream shows the main components of our solution approach and their temporal correspondence. In the following, we describe the interaction of our solution approach with the business process. Thereby, we do not assume that the individual components are already known and, thus, only provide a high-level overview, following the numbering within the figure: (1) A customer logs in to the website with information about their location and delivery preferences stored in the profile and chooses a shopping basket online while expecting a selection of narrow delivery time spans to be offered at affordable prices. This initiates a delivery request in response to which the provider has to make a demand-management decision. Therefore, simultaneously to the customer's login, (2) the provider samples different customer request trajectories and conducts tentative tour-planning optimization, called multiple scenario approach. From the solution of the multiple scenario approach, a value approximation can be conducted, which is then the input for an anticipatory demand-management decision (3). More precisely, based on the approximated value, the provider derives anticipatory decisions on which delivery time spans to offer the current customer and at what prices. (4) As a result, the customer chooses one of the options offered or leaves the website without purchasing, following their own individual preferences. If the customer chooses to purchase, the delivery request becomes a confirmed customer order with a delivery deadline and (5) the tour planning is updated on the basis of the previously sampled trajectories. (6) To enable prompt delivery, the execution of deliveries might start/continue immediately, even though further customer requests could arrive. (7) When a new customer request arrives, the whole process starts over again. Note, as typical in the related literature, we assume that customers log-in to the website one-by-one.

The contribution of this paper is threefold and regards *MDP modeling*, *solution approach*, and *practical application* as described in the following:

- *MDP modeling* – We contribute to the literature on modeling MDPs for SDD applications by being the first to explicitly formalize the interaction between two co-dependent types of decisions, i.e., the demand-management and tour-planning decisions, in a specifically adapted Bellman function.
- *Solution approach* – As the main contribution of our work we propose a holistic anticipatory solution approach to the integrated demand-management and online tour-planning problem, which does not require extensive offline learning, and guarantees applicability and scalability to realistically sized problem instances.
- *Practical application* – In a comprehensive computational study, we derive substantial contributions regarding the practical application of the SDD-DMTP. First, we derive a potential of increasing the contribution margin by anticipation in decision making within our approach of up to 50% and elaborate cases in which anticipation is particularly valuable. Therewith, we give a differentiated insight into the problem. Second, we benchmark our approach in relation to other demand-management and pricing approaches, adopting ideas from the existing literature. We show how explicit price optimization increases the contribution margin compared to the benchmarks, and discuss how the different approaches affect the solution structure. Generally, the results give a strong indication that our new approach can deliver decision support that helps to finally make SDD applications profitable in practice.

Table 1
Surveys that feature related problems.

Authors	Application	Perspective	Concepts	Models	Approaches
Agatz, Campbell, Fleischmann, Van Nunen, & Savelsbergh (2013)	AHD	DM	√	✗	√
Archetti & Bertazzi (2021)	G	VRP	√	✗	√
Boysen, Fedtke, & Schwerdfeger (2021)	G	VRP	√	✗	✗
Fleckenstein, Klein, & Steinhardt (2022)	G	i-DMVRPs	√	√	√
Klein, Koch, Steinhardt, & Strauss (2020)	G	DM	√	√	√
Snoeck, Merchan, & Winkenbach (2020)	AHD	VRP	√	✗	✗
Soeffker, Ulmer, & Mattfeld (2021)	G	VRP	√	√	√
Waßmuth, Köhler, Agatz, & Fleischmann (2022)	G	i-DMVRPs	√	✗	√

The remainder of the paper is organized as follows. In Section 2, we review the literature of related research streams. In Section 3, we model the problem as MDP with two integrated decisions. Then, in Section 4, we present our solution approach, which is subsequently evaluated in a numerical study. We discuss the results in Section 5, and in Section 6 summarize them, also giving the insights that emerged, as well as future research directions.

2. Literature review

In this section, we examine the existing literature related to SDD-DMTPs. In Section 2.1, we start with a short overview of existing surveys, that generally address related demand-management and tour-planning problems. Then, in Section 2.2, existing *solution approaches* for operational decision making in SDD-DMTP problem settings are analyzed. Note that we omit a detailed discussion of existing solution approaches for related AHD problem settings, because, as stated earlier, the overlap of the booking and service horizons in SDD yields substantially different challenges since it incorporates an online tour-planning component. Nevertheless, in Appendix B, we provide a list of recent solution approaches of related AHD literature. We conclude the discussion of the related literature by summarizing the identified research gaps in Section 2.3.

2.1. Surveys on integrated demand management and tour planning

In this section, we provide a brief overview of the existing surveys featuring integrated demand management and tour planning in home-delivery applications in general. We list the respective works in Table 1 and show whether AHD is considered in particular, or if general concepts are addressed (G). There is no survey that specifically addresses SDD literature. Further, Table 1 summarizes whether the authors focus on a specific component, i.e., the demand-management component (DM) or the tour-planning component (VRP), or consider the integrated demand-management and tour-planning problem (i-DMVRP) holistically. Additionally, Table 1 gives an overview of whether the respective survey addresses business concepts, mathematical models, and/or solution approaches. The reader is referred to Appendix A for a summary of the key insights of each of the listed surveys.

2.2. Solution approaches for the SDD-DMTP

Here, we describe existing solution approaches for the SDD-DMTP with a focus to anticipatory approaches. We structure our elaboration by first discussing research that involves learning-based anticipation (see Section 2.2.1). Next, we discuss research that involves non-learning-based anticipation (see Section 2.2.2). In each of these sections, we start by briefly discussing the most related research on pure online tour-planning approaches to show how the discussed integrated approaches evolve from them. Integrated approaches exceed basic feasibility control by (at least) allowing feasible customer requests to be rejected if the expected contribution to the objective is negative (Fleckenstein, Klein, & Steinhardt, 2022). Further, our discussion is structured along groups of publications that follow comparable ideas. For each group, we first describe the approaches and then highlight what distinguishes our approach from the presented works.

2.2.1. Learning-based approaches

Pure online tour-planning: Learning-based approaches aim to learn accurate value function approximations (VFAs) either offline, by simulation in upstream learning phases, or online. To solve stochastic dynamic VRPs with unknown requests, VFAs are typically applied to derive tour-planning/routing decisions. Recent publications on solving stochastic dynamic VRPs with stochastic customer requests as considered in our problem are Ulmer (2017) and Ulmer (2019). These works present a variety of VFA approaches to make tour-planning decisions that will match as many customer requests as possible. In all approaches, the VFA is learned offline by a large number of simulation runs. The learned VFA can be applied to assess post-decision values in an online decision period in order to make good tour-planning decisions. Hildebrandt, Thomas, & Ulmer (2021) summarize solution frameworks for solving stochastic dynamic VRPs that originate from different research streams, namely *computer science* and *operations research*. They propose a high-level concept on how to combine those frameworks to build a reinforcement learning-based solution framework.

Integrated approaches: Ulmer, Goodson, Mattfeld, & Hennig (2019) combine an offline VFA with a simulation-based online rollout algorithm to solve a dynamic VRP with stochastic service requests for a single vehicle. They present an offline learned, dynamic look-up table which is generated by approximate value iteration using temporal information. When this look-up table is used for online decision making, it is combined with a simulation-based online rollout algorithm considering spatial information of potential post-decision states. In contrast to our approach, their demand-management decision results from the optimized tour-planning, but is not actively steered nor anticipated. Additionally, we consider multiple vehicles with multiple tours, and offer multiple delivery options to incoming customer requests. This results in very large state and action spaces, even if a state space aggregation is applied. Therefore, generally, look-up table based approaches cannot be implemented efficiently for the setting we consider.

In a different set of publications, researchers consider a pricing component within their integrated approaches. Ulmer (2020a) solves a dynamic routing and pricing problem for SDD by developing an anticipatory pricing and routing policy that is based on a sophisticated VFA approach and upstream policy learning. He is the first to present a VFA approach for a fleet of vehicles, which he does by separating the value function with regard to different vehicles. He includes the tour-plans of the vehicles in the state definition. To solve the pricing problem, the author relies on an opportunity cost estimate for different delivery options from comparing approximated state values. If the opportunity costs are low, the corresponding delivery options are offered for *budget prices* derived from the upstream policy learning. Those prices represent the typical base prices of the delivery options. Only in cases where the opportunity cost estimate exceeds the budget price, are the corresponding delivery options priced differently, setting prices to equal the opportunity costs. Therefore, this procedure ensures that only requests with a non-negative contribution to the overall objective are accepted. The routing aspect of the problem is solved by a simple, non-anticipatory insertion heuristic.

In the same set of publications, Prokhorchuk, Dauwels, & Jaillet (2019) introduce a stochastic dynamic pricing and routing problem for SDD with stochastic travel times. They also base decision making on the approximation of opportunity costs and amend the approach of Ulmer (2020a) by stochastic travel times, a different routing heuristic that accounts for stochastic travel times, and by using standard VFA procedures.

Our work differs decisively from the above-mentioned two decision-making approaches. In both, the authors construct a VFA around the post-decision state that is derived from myopic tour planning, i.e., from cheapest insertion algorithms. They approximate the corresponding value, and thus the opportunity cost, by anticipating customer orders that can be accepted. Note that the proposed learning based approaches rely on the use of base prices in their anticipation, which is a pre-requisite for this VFA (cf. Ulmer, 2020a). However, in contrast, we aim to integrate explicit anticipation in both tour planning and price optimization. More precisely, our tour-planning approach substantially differs from theirs, as we apply anticipatory replanning for every new customer order. Regarding the price optimization, we solve a choice-based pricing optimization problem with discrete, predefined price points and we aim to apply this optimization for anticipation in learning a state value as well, instead of relying on base prices. We propose a non-learning-based value approximation approach that incorporates a number of novel, problem-specific ideas.

2.2.2. Non-learning-based approaches

Here, we discuss non-learning-based solution approaches for stochastic dynamic VRPs with stochastic requests.

Pure online tour planning: Bent & Van Hentenryck (2004) introduce a multiple scenario approach to take tour-planning decisions in dynamic VRPs with stochastic customer requests. They aim to maximize the number of accepted customer requests by constantly generating multiple tour plans based on sampled customer requests. From those tour plans, a distinguished tour plan is chosen, repaired for feasibility, and frequently updated. It serves as input for taking decisions on which customers will be served next and by which vehicle. In Bent & Van Hentenryck (2007), the authors enhance the previous approach by including waiting and relocating strategies. With this approach, not only routing decisions, e.g., a vehicle's next destination, but also dispatching decisions, i.e., which orders to allocate to one tour, are taken.

Integrated approaches: Among the considered integrated, non-learning-based approaches, the most relevant group of papers is based on the idea of the multiple scenario approach by Bent & Van Hentenryck (2004), as described above. Azi, Gendreau, & Potvin (2012) introduce an initial demand-management approach to a dynamic VRP with stochastic requests and non-disjoint booking and service horizons. They consider a profit maximization problem in determining which requests to accept and which to reject. To solve the routing problem, they apply an adaptive large neighborhood search to scenarios that, like the ones in Bent & Van Hentenryck (2004), include already accepted customer orders and sampled customer requests. What is new about their approach is that they then compare scenario solutions with and without the current customer request and define the difference in solution quality as a scenario-specific *opportunity value*. If the sum of all scenario-specific opportunity values is positive, they accept the request. This approach delivers an estimate of whether or not the acceptance of a customer request yields a positive contribution to the overall objective, taking potential future developments into account.

Voccia, Campbell, & Thomas (2019) also adapt the ideas from Bent & Van Hentenryck (2004) and Bent & Van Hentenryck (2007). They aim to maximize the number of feasibly inserted customer requests for a stochastic dynamic VRP with time windows and stochastic requests. The customer requests that are not inserted in a feasible solution are outsourced to a third party logistics provider, which comes with a penalty cost per order. Their approach yields comprehensive tour-planning decisions including the set of orders allocated, vehicle assignment, as well as a schedule for each tour. Like Bent & Van Hentenryck (2007), they consider future, not yet realized customer requests by applying a sample-scenario approach. Thereby, they solve a multi-trip team orienteering problem with a standard implementation of a variable neighborhood search. Afterwards, the scenario solutions are used to construct anticipatory tour plans. Compared to Bent & Van Hentenryck (2007), they apply an enhanced consensus function that chooses partial plans according to their appearance frequency in the scenario solutions. Also, they include waiting strategies to improve the anticipatory quality of their solutions. Côté, de Queiroz, Gallesi, & Iori (2021) build on the approach by Voccia et al. (2019) and amend it by a regret heuristic, a different consensus function, and a specifically tailored branch-and-regret method. Further, they also consider settings in which pre-emptive depot returns are allowed.

Regarding the setting and the solution approach this set of publications is related to our work; however, the decisive difference is that they do not consider explicit demand management, i.e., which customers are served is a result of a pure tour-planning optimization. Compared to tour-planning problems without explicit demand management, for solving our problem it is critical to have a very accurate value approximation. This is needed to determine profitable prices across a relatively small set of close price points. At the same time, we need an online tour-planning approach. Therefore, we combine the ideas of multiple-scenario approaches for online tour-planning with a basic idea known from rollout algorithms, namely approximating the value of a decision by averaging the values of heuristic solutions of sampled trajectories (Soeffker et al. (2021)). More precisely, we extend the approach of Voccia et al. (2019) by incorporating a sophisticated demand-management approach that anticipates demand management in scenario solutions for value approximations. Finally we need to point out that there is another research stream dealing with SDD problems considering integrated approaches for SDD, but in this context neither the setting nor the approach is as closely related to our problem as those previously mentioned. Respective works are Klapp, Erera, & Toriello (2018), Klapp, Erera, & Toriello (2020), Chen, Ulmer, & Thomas (2019), and Soeffker, Ulmer, & Mattfeld (2017). Further, please note that there are other research streams in the SDD community that consider a variety of research questions which we disregarded here. For a review of respective fields of research, the interested reader is referred to Boysen et al. (2021).

2.3. Research gaps concerning the literature on MDP modeling and solution approaches regarding the SDD-DMTP

From a comprehensive analysis of the existing literature on related problems as summarized in Appendix B, the following research gaps can be conducted with regard to *MDP modeling* and *solution approach*:

- *MDP modeling* – There is no MDP model neither for SDD problem settings nor for AHD problem settings which ex-

plicitly accounts for the simultaneous integration of demand-management and tour-planning decisions.
- *Solution approach* – While there is a wide range of solution approaches that tackle AHD problem settings, the existing solution approaches for SDD are not sufficiently holistic and evolved to integrate anticipation in demand management and tour planning at the same time. Further, literature proposing approaches that aim at profit optimization and thereby consider revenue and cost at the same time, is scarce for SDD.

Further, the following insights can be derived: Among the approaches that address anticipatory tour planning of which demand-management decisions are an implicit result, there exist learning-based and non-learning-based approaches. Among the approaches that apply explicit anticipatory demand management but base that on myopic tour planning, there are only learning-based approaches. In the following, we close the identified research gaps by presenting a holistic MDP model, and by proposing a solution approach, which involves anticipation for the tour-planning optimization and the demand-management optimization simultaneously. The proposed approach is a non-learning based approach. In the following, we start by introducing the holistic MDP model for the SDD-DMTP.

3. Problem statement

In Section 3.1, we introduce the SDD-DMTP in detail and state our assumptions. In Section 3.2, we formalize it in a holistic MDP model formulation.

3.1. Problem description

The SDD-DMTP comprises two types of decisions, namely *demand-management decisions* and *tour-planning decisions*. Demand-management decisions have to be made for every *customer arrival* and comprise the decisions on which *delivery options* to offer each particular customer at which prices. The combination of a subset of delivery options with fixed prices is termed an *offer-set*. Every offer-set gives different customer *choice probabilities* according to which customers choose a delivery option, thus-either turning the request into a confirmed customer order or choosing to leave the system without purchasing anything. All confirmed customer orders have to be served by the provider's *delivery operations*. Therefore, the provider continuously takes tour-planning decisions and executes them, while the booking period is still running. Below, we describe the relevant components of the SDD-DMTP in detail:

Customer arrivals: Customer requests c can arrive at random times t within a pre-defined booking period with arrival rate λ. The arriving customers log in to the provider's website with registered profiles and fill their shopping basket. For every incoming customer request, the provider then knows the corresponding location $(x, y)_c$, as well as the shopping basket's potential value r_c.

Delivery options: *Delivery options* are predefined nested time spans in which the provider commits to deliver. The set of delivery options could, for example, comprise delivery within the next 90 minutes or within the next 300 minutes. Delivery options are referred to with indices in ascending order, so that the length of the delivery option with index i, denoted as $l(i)$, is shorter than the length of delivery option i', denoted as $l(i')$, if $i < i'$. The corresponding index set is denoted by $\mathcal{I}$.

Offer sets: Considering all delivery options, the provider decides on a subset to offer in response to an incoming customer request. In doing so, the provider also selects a price for each delivery option, either from predefined price points or from a continuous (potentially limited) price range. In defining offer sets, the provider has to take the following restrictions into consideration: (1) Within an offer set, each delivery option can appear only once. (2) Since a customer can always decide not to make a purchase, a fictive delivery option that represents a no-purchase option has to be included in every offer set. It is priced at zero and is referred to by index $i = 0$. (3) To ensure pricing consistency, the prices of deliveries with longer time spans can never exceed the prices of those with shorter spans. An offer set is denoted by g and the set of all offer sets by $\mathcal{G}$. It is either finite, if potential prices originate from a finite set, or infinite, if potential prices originate from a continuous range.

Table 2
Potential offer-sets.

	prices of delivery options		choice probabilities			
g	$i = 1$	$i = 2$	P^0	P^1	P^2	Σ
1	r^{12}	r^{22}	0.2	0.4	0.4	1
2	r^{12}	r^{21}	0.3	0.4	0.3	1
3	r^{12}	not offered	0.4	0.6	0.0	1
4	r^{11}	r^{22}	0.3	0.2	0.5	1
5	r^{11}	r^{21}	0.4	0.2	0.4	1
6	r^{11}	not offered	0.6	0.4	0.0	1
7	not offered	r^{22}	0.3	0.0	0.7	1
8	not offered	r^{21}	0.5	0.0	0.5	1
9	not offered	not offered	1	0.0	0.0	1

Customer choice probabilities: The utility u^i that an arriving customer experiences when choosing a certain delivery option i at a certain price r^i consists of an observable and an unobservable component. The deterministic (i.e. observable) component can be calculated and is influenced by the length of the delivery option and its price. The unobservable component can be drawn from a certain probability distribution and is represented here by ϵ^i. The *choice probability* with which a customer chooses a delivery option i from a set of offered options g is $P^i(g) = P(u^i \geq \max\{u^i : i \in g\})$. It is calculated differently for different random utility models which usually differ regarding the assumptions underpinning the distribution of the random component ϵ^i (see Talluri & Van Ryzin (2006) for more on random utility models). Table 2 illustratively shows a set of offer-sets $\mathcal{G}$ with $|\mathcal{G}| = 9$. Every $g \in \mathcal{G}$ is depicted in a row with artificial purchase probabilities. In the example, there are two different delivery options $\{1, 2\}$ with $l(1) < l(2)$, and two potential prices r^{i1}, r^{i2} each, with $r^{11} > r^{12} > r^{21} > r^{22}$. The no-purchase probability is denoted as P^0.

Delivery operations: If a customer chooses an option other than the no-purchase option, their request c turns into a *customer order* c. A customer order is assigned a *delivery deadline* t_c^{due} that is calculated from its request time t_c^{req} and the length $l(i)$ of the chosen delivery option i, i.e., $t_c^{due} = t_c^{req} + l(i)$. Since the delivery deadlines are typically narrow in SDD, the service period in which customer orders are being served starts with, or shortly after, the first realized customer order and ends when the last customer order of a day has been served. Hence, a particularity of SDD is that the booking and service periods overlap (Fleckenstein et al., 2022).

During the service period, a fleet of homogeneous vehicles $\mathcal{V}$ serves the customer orders from a centrally located depot. Once a customer order has been realized, it has to be loaded onto a vehicle in the depot. Thus, the order can only be served either by a vehicle that visits the depot after the request's arrival or by one that is idle in the depot when the request arrives. The provider continuously takes tour-planning decisions, i.e., decisions about whether and when a vehicle should leave the depot, and when it leaves which orders will be assigned to it. If a vehicle leaves the depot to serve customers, a *tour* is planned. A tour is denoted as θ^v for a vehicle $v \in \mathcal{V}$ and is defined by a start time t^{start} and a set of loaded customer orders $L = \{c_1, c_2, c_3, \ldots\}$. Further, to store the order in which a given tour will reach customer locations, we introduce a set of tuples that assign positions χ_{c_i} to customer orders $c_i, i \in L$. The set is denoted by X, i.e., $X = \{(c_1, \chi_{c_1}), (c_2, \chi_{c_2}), (c_3, \chi_{c_3}), \ldots\}$. Hence, $\theta^v = (t^{start}, L, X)$. Accordingly, we refer to the fields of the tuple of a given tour by $t^{start}(\theta^v), L(\theta^v)$ and $X(\theta^v)$.

After leaving the depot, a tour is always fully executed as planned, without pre-emptive depot returns. All tours have to be planned in such a way that no customer orders will be served later than their delivery deadline. We assume deterministic travel times $\tau_{cc'}$ from the location of the customer order c to the location of customer order c'. Without loss of generality, we assume that the service time of serving customer order c' is included in $\tau_{cc'}$. Vehicles can have several sequential tours during the day. In line with most SDD literature, we do not consider physical vehicles' capacities, because the narrow delivery deadlines of the customer orders are far more restrictive for the delivery operations than the available space in a vehicle (see for example Angelelli, Archetti, Filippi, & Vindigni, 2021, Lang, Cleophas, & Ehmke, 2021a, Berbeglia, Cordeau, & Laporte, 2010, Ulmer, 2020a, Voccia et al., 2019).

3.2. Markov decision process formulation

In the following, we model the SDD-DMTP as a Markov decision process (MDP). In every *decision epoch*, the provider has to take an *action* denoted as A_t, taking the current *state* of the system S_t into consideration. More precisely, the provider evaluates the current state of the delivery operations, as well as the customer orders already confirmed though not yet being delivered, to determine the feasibility of potential actions. The action then taken, yields a *transition* as well as a *reward* in that a customer chooses a delivery option from the set of offered delivery options (including the no-purchase option), which triggers the execution of the corresponding delivery decision. Accordingly, rewards follow: if a customer chooses to buy, they pay a delivery fee and the shopping basket value realizes. Further, tour costs realize for every vehicle that leaves the depot. After such realizations, the system transitions to the next state S_{t+1}, which differs from the previous one, potentially by the newly accepted customer order and/or new tours, and by the delivery execution's progress up to that time. The objective that the provider seeks is to maximize the total profit accrued over all decision epochs. In the following, we describe the MDP elements in detail:

Decision epochs: To model the SDD-DMTP as MDP, we assume a discretized booking period $\mathcal{T}^{book}$ in such a way that the stages of the MDP correspond to time steps $t \in \mathcal{T}^{book} = 1 \ldots T$. The time steps represent micro-periods in each of which no more than one customer request with arrival probability λ arrives.

State: The state S_t of a system at time t consists of all information relevant to making decisions and already revealed by time t. In the SDD-DMTP, two state components are required. The first component is the set of confirmed and not yet being delivered customer orders, denoted as $\mathcal{C}_t$. For all of those orders $c \in \mathcal{C}_t$, it contains information about their location $(x, y)_c$ and their due time t_c^{due}, stored in a tuple: $((x, y)_c, t_c^{due})$. The second component is the overall tour plan at time t, denoted by ϕ_t (see modelling of route-based MDPs in Ulmer, Goodson, Mattfeld, & Thomas, 2020). It contains the currently running tours θ_t^v for every vehicle $v \in \mathcal{V}$. If the vehicle v is idle in the depot, $\theta_t^v = ()$. So, the state is defined as $S_t = (\mathcal{C}_t, \phi_t)$.

All possible combinations of customer requests from the registered customer pool, with all possible arrival and due times and with all possible tour plans, define the state space $\mathcal{S}$, with $S_t \in \mathcal{S}$.

Action: Fig. 2 is a schematic representation of the stochastic decision process of the SDD-DMTP. We differentiate between actions in decision epochs in which, with probability λ, a customer arrival occurs and decision epochs in which, with probability $(1-\lambda)$, no customer request arrives. In the former case, two types of decisions have to be made integratively, namely demand-management and tour-planning decisions. In the latter case, only tour-planning decisions have to be made. We depict both cases in Fig. 2.

Customer request – In period t of Fig. 2, a customer request arrives. Therefore, a demand-management decision has to be made by selecting which offer set $g \in \mathcal{G}$ to offer the requesting customer. The offer set presented at time t is denoted as g_t. Further, for every delivery option $i \in g_t$, potential tour-planning decisions that are executed after observing the customer's actual choice, are made. Thus, a tour-planning decision consists of the subsequent state's possible overall tour plan which depends on the yet unknown customer choice for a delivery option i. Therefore, we introduce ϕ_t^i for $i \in g_t$ as the tour plans that will be executed if the customer were to choose delivery option i, and include it in the action definition.

No customer request – In $t+1$ of Fig. 2, no customer request arrives. Therefore, the corresponding action only comprises tour-planning decisions ϕ_{t+1}^0 without a new customer request.

Accordingly, the action A_t of micro-period t has two distinct cases:

$$A_t = \begin{cases} (g_t, (\phi_{t+1}^i)_{i \in g_t}) & \text{if there is a customer request at } t \\ \phi_{t+1}^0 & \text{else} \end{cases} \quad (1)$$

Correspondingly, the action space at decision epoch t, denoted as $\mathcal{A}_t$, is also defined for the above-mentioned two distinct cases:

Customer request – For the first case, if there is a customer request in t, the action space comprises two components. One component, denoted as $\mathcal{G}(S_t, c_t)$, defines all offer sets that only contain delivery options for which there is at least one feasible tour plan, given state S_t and customer c_t. The other component, denoted as $(\Phi_{t+1}^i(S_t, c_t))_{i \in \mathcal{G}(S_t, c_t)}$, defines all potential tour plans that are feasible given $\mathcal{C}_t$, and assuming that the current customer request c_t turns into a customer order with a deadline according to delivery option $i \in \mathcal{G}(S_t, c_t)$. This could also comprise the decision that no new tour will start, i.e., that the tour plan does not change.

No customer request – If there is no customer request in t, the action space accordingly comprises all feasible tour plans for the set of confirmed customers $\mathcal{C}_t$. In this case, the set of all potential tour plans is denoted as $\Phi_{t+1}^0(S_t)$. Consequently,

$$A_t \in \mathcal{A}_t = \begin{cases} (\mathcal{G}(S_t, c_t), & \text{if there is} \\ (\Phi_{t+1}^i(S_t, c_t))_{i \in \mathcal{G}(S_t)}) & \text{a customer request at } t \\ \Phi_{t+1}^0(S_t) & \text{else.} \end{cases} \quad (2)$$

Note that for both cases, the tour-planning component of the action space comprises all tours currently running at t and potentially new tours for vehicles that are standing idle at the depot ($\theta_t^v = ()$) or returning to the depot during the decision epoch.

Transition model: The transition model of the SDD-DMTP comprises demand-management related and tour-planning related transitions. While the former are stochastic, the latter are deterministic transitions from one state S_t to a successor state S_{t+1}. Fig. 3 is a schematic representation of the transitions involved in the SDD-DMTP, which shows the temporal relation between two consecutive states and transitions. As Fig. 3 shows, the stochastic event of whether there is a new customer request c_t arriving or not can be observed at the beginning of a decision epoch t, after observing state S_t. The resulting transitions differ accordingly.

If there is a request, integrated demand-management and tour-planning decisions are made and a transition, namely the customer choice i' follows. This is depicted in the upper stream of Fig. 3. This transition is stochastic, and potential outcomes i' can be observed with known probability $P^{i'}(g_t)$. It defines whether the first state component, namely the set of confirmed, yet still to be delivered customer orders $\mathcal{C}_t$, alters from one state S_t to a successor state S_{t+1} by adding a new customer order. Following this, another transition, namely the execution of deliveries, brings the system to the

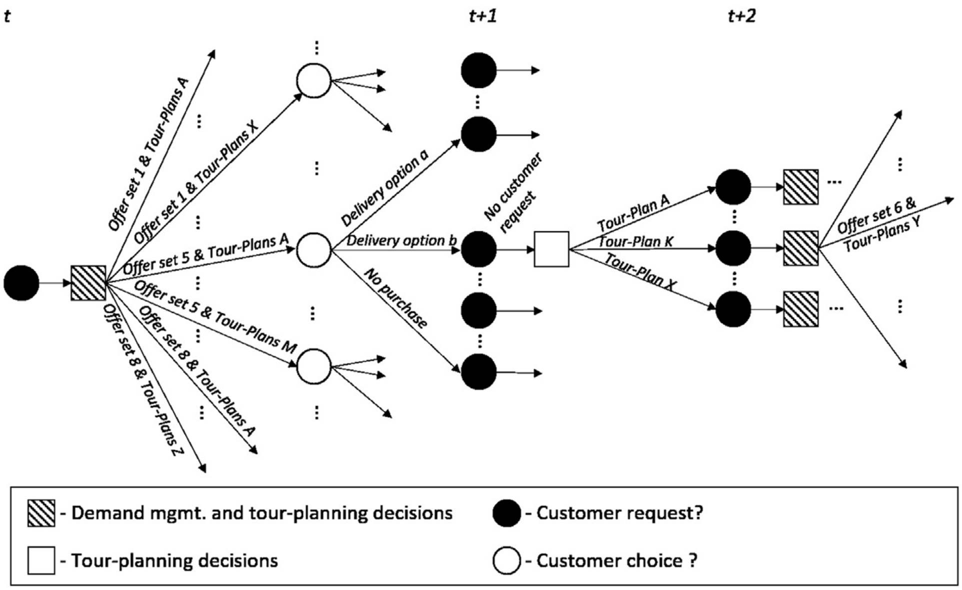


Fig. 2. Schematic representation of the stochastic decision process.

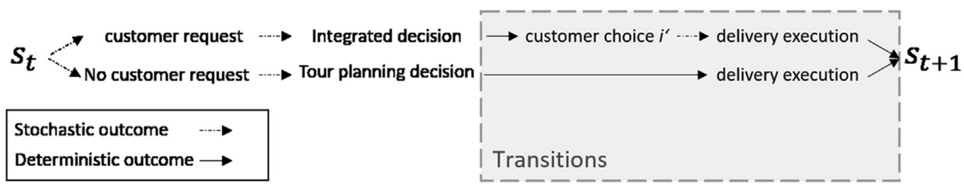


Fig. 3. Schematic representation of the transitions.

next state S_{t+1}. The latter strictly follows the tour-planning decision $\phi^{i'}_{t+1}$ in A_t, with $i' = 0$ representing the case that the current requesting customer in t has rejected all offered delivery options. As we assume deterministic travel times, this transition is purely deterministic, therefore in state S_{t+1}, ϕ_{t+1} is set to $\phi^{i'}_{t+1}$ from A_t. This also influences the first state component, because all customer orders from set $\mathcal{C}_t$ that are newly loaded onto a vehicle according to the new tour-plan ϕ_{t+1}, are removed from $\mathcal{C}_t$. We introduce the set $\Psi(\phi_{t+1} \mid \phi_t)$ that contains all those customers.

If no customer request is observed in state S_t, only tour-planning decisions are made. The corresponding deterministic transition of the delivery execution alters the system from state S_t to the successor state S_{t+1}. This is depicted in the lower stream of Fig. 3.

The transitions of the state components can be formalized as follows:

$$\phi_{t+1} = \phi^{i'}_{t+1} \tag{3}$$

$$\mathcal{C}_{t+1} = \begin{cases} \mathcal{C}_t \setminus \Psi(\phi_{t+1} \mid \phi_t), & \text{if there is no customer request in } t \text{ with probability} (1 - \sum_{c_t \in C} \lambda_{c_t}(t)), \text{ or if the incoming request } c_t \text{ does not turn into a customer order with probability } \lambda_{c_t}(t) \cdot P^0_{g_t} \\ (\mathcal{C}_t \cup \{c_t\}) \setminus \Psi(\phi_{t+1} \mid \phi_t), & \text{if there is a customer request } c_t, \text{ that turns into a customer order with probability} \lambda_{c_t}(t) \cdot \sum_{i \neq 0} P^i_{g_t} \end{cases} \tag{4}$$

Rewards: The SDD-DMTP rewards can also be attributed to demand-management and tour-planning decisions. The rewards accrued through the demand-management related transitions are positive. They are the sum of the contribution margin of the customer order c, denoted as r^i_c, and the delivery fee of the chosen delivery option $r^i(g_t)$, determined by the offer set g_t. Here $r^i_c = r_c$ applies, for all $i \neq 0$, and $r^i_c = r^i(g_t) = 0$, for all $g_t \in \mathcal{G}$, if $i = 0$. The rewards that are induced by the deterministic tour-planning related transitions are negative. Such rewards are called logistics-related rewards of a transition from S_t to S_{t+1} given a decision ϕ^i_{t+1}, formally denoted as $r^l_{\phi^i_{t+1}}$. The logistics-related rewards equal the sum of delivery costs of all tours that start in $t+1$ according to action A_t.

Objective: The objective of solving the SDD-DMTP is to maximize the overall profit, i.e., to maximize the difference between the sum of positive rewards and the sum of negative rewards. Positive rewards are accrued across all decision periods by selling shopping-baskets and delivery options. Negative rewards equal the delivery costs accrued across all decision periods. The objective can be represented by the well-known Bellman equation that captures the value of being in a given state (Powell et al., 2012). We specify the Bellman equation for the SDD-DMTP and explicitly integrate the two interdependent decisions that have to be made:

$$\begin{aligned} V(S_t) = {} & \sum_{c_t \in C} \lambda_{c_t}(t) \cdot \arg\max_{g \in \mathcal{G}} \Bigg(\sum_{i \in g} P^i(g) \cdot \Big[r^i(g) \\ & + r^i_{c_t} + \arg\max_{\phi^i_{t+1} \in \Phi^i_{t+1}} (r^l_{\phi^i_{t+1}} + V(S_{t+1} \mid S_t, \phi^i_{t+1})) \Big] \Bigg) \\ & + (1 - \sum_{c_t \in C} \lambda_{c_t}(t)) \cdot \arg\max_{\phi^0_{t+1} \in \Phi^0_{t+1}} \\ & \cdot \Big(r^l_{\phi^0_{t+1}} + V\big(S_{t+1} \mid S_t, \phi^0_{t+1}\big) \Big), \end{aligned} \tag{5}$$

with boundary condition:

$$V(s_{T+1}) = 0. \tag{6}$$

The first two lines of Eq. (5) reflects the value and decision making in cases where a customer request arrives. The last two lines reflect the corresponding value and decision making if no customer request arrives. If a customer request arrives, the provider derives the demand-management decision by solving arg $\max_{g\in\mathcal{G}}(\cdot)$. To do so, the provider needs to consider the value of all delivery options (including the no-purchase option) i that the current customer might choose. This is obtained by solving arg $\max_{\phi^i_{t+1}\in\Phi^i_{t+1}}(\cdot)$. If no customer request arrives, the tour-planning decisions equal those of $i = 0$.

4. Solution approach

The presented SDD-DMTP is a dynamic stochastic optimization problem with large state and action spaces. Since solutions have to be determined in near real-time, it is not possible to solve it to optimality. Therefore, we develop a heuristic solution approach which takes the SDD-DMTP's two types of decision into account, namely dynamic demand-management decisions and online tour-planning decisions. The approach is based on the following consideration: If the second part of the second line of Eq. (5), i.e.,

$$V'(S^i_{t+1}) := \arg\max_{\phi^i_{t+1}\in\Phi^i_{t+1}}\left(r^l_{\phi^i_{t+1}} + V\left(S_{t+1}\mid S_t,\phi^i_{t+1}\right)\right) \tag{7}$$

was known for all potential customer choice outcomes $i \in g$, solving the Bellman Eq. (5) would be simplified tremendously. $V'(S^i_{t+1})$ captures the optimal tour-planning decision, i.e., arg $\max_{\phi^i_{t+1}\in\Phi^i_{t+1}}(\cdot)$, the resulting logistics reward, i.e., $r^l_{\phi^i_{t+1}}$, as well as the value of the associated successor state, i.e., $V(S_{t+1} \mid S_t, \phi^i_{t+1})$. Thus, with known $V'(S^i_{t+1})$, and if $|\mathcal{G}|$ is not large, Eq. (5) could be solved to optimality by total enumeration across all $g \in \mathcal{G}$ (Yang, Strauss, Currie, & Eglese, 2016). In the SDD-DMTP, we indeed assume $|\mathcal{G}|$ to be of tractable size. However, $V'(S^i_{t+1})$ cannot be determined exactly. Thus, we propose a problem-specific approximation of $V'(S^i_{t+1})$, i.e., an approximation of the optimal tour-planning decision, the related reward, and the resulting value, which is carried out every time a customer request arrives and decisions have to be made. As already outlined in the introduction and schematically depicted in the upper part of Fig. 1, the underlying procedure consists of three main components: a *multiple scenario approach* (see Section 4.1), from which a *value approximation* can be derived (see Section 4.2), and that, at the same time, returns *anticipatory tour-planning* decisions (see Section 4.3).

In the following, we describe each component of the approach separately, starting with a description of how to generate and solve scenarios.

4.1. Multiple scenario approach

The sample-scenario value approximation and tour-planning approach adapts the online tour-planning ideas of Bent & Van Hentenryck (2004) as well as Voccia et al. (2019) and substantially extends them in order to include demand-management decisions. The basic idea is to sample scenarios, and then, to solve a deterministic version of the SDD-DMTP (d-SDD-DMTP) for every scenario. The resulting solutions are then used to derive state values as input for decision making. In this section, we present the newly developed multiple scenario approach (MSA) in more detail. We first describe how scenarios are generated. Then, we describe how to solve these scenarios.

Generation of scenarios: Every time a customer request arrives, different customer request realizations are sampled into the future to generate scenarios. In order to reach the previously described goals, i.e., to derive a value approximation for demand-management decisions and to derive tour-planning decisions, scenarios are needed that are state, time, and customer-choice-specific. Consequently, a scenario $\omega \in \Omega^i_t$ at time t and for a certain delivery option i consists of three types of customers: first, the confirmed and not yet being delivered customer orders $\mathcal{C}_t$; second, the current customer request c_t with assigned deadline according to i, i.e., $t^{due}_{c_t} = t + l(i)$, if $i \neq 0$; third, a sampled realization of customer requests N^ω, sampled from t on until the end of a predefined sampling horizon length. For those sampled customer requests, for now we assume a preliminary delivery deadline according to the longest available delivery span, $t^{due}_c = t^{req}_c + \max\{l(i) \mid i \in \mathcal{I}\}$. Further, for the sampled customers, we simplify the demand-management decisions, such that the respective requests can only be accepted or rejected. This allows us to formulate the d-SDD-DMTP for each scenario, as a deterministic, profitable multi-trip vehicle routing problem, as formalized in Appendix C. In the following, we present a specifically tailored heuristic to solve the d-SDD-DMTP.

Solving scenarios: The d-SDD-DMTP, as presented in Appendix C, is a profitable multi-trip vehicle routing problem (PVRPMT), and thus, belongs to the class of NP-hard problems (Chbichib, Mellouli, & Chabchoub, 2012). Even more, the d-SDD-DMTP has to be solved for every scenario. Consequently, we cannot solve the presented MIP for all instances in reasonable time. Instead, we propose a heuristic approach, which consists of the following three steps:

Relaxation – First, for the moment, we relax explicitly considering depot returns in the d-SDD-DMTP. The resulting problem is a profitable single trip vehicle routing problem with time windows (P-VRPTW) (Toth & Vigo, 2014). A customer request's arrival time now forms the start of their delivery time window, while the delivery deadline remains unchanged. The trick is that all vehicles can now start only one tour, but can, theoretically, serve customer orders that have not yet realized at the time the tour starts. Therewith, we enable to apply standard tour-planning software in the next step and, thus, ensure practical applicability of our heuristic.

Solving the relaxed problem – Next, we solve the resulting P-VRPTW heuristically by means of a standard tour-planning software (e.g. Google OR Tools). The result is a tour plan with one tour per vehicle, including confirmed and sampled customer orders.

Feasibility repair – When the tours start, not all sampled customer orders have already realized, which is why we have to add depot returns to the planned tours. Thus, for feasibility, we repair the respective tours as follows: We interrupt a vehicle's tour for a depot return each time a sampled customer order has to be served of which the request had not yet arrived when the tour started in the depot. For the same vehicle, a new tour is planned to serve the original tour's remaining customers in the same order, until it has to be interrupted for another depot return. If a depot return causes a late delivery for a sampled customer, the customer is removed from the tour; yet, if the depot return causes a late delivery for a confirmed customer, the latest sampled customer is removed from the tour and, according to vehicle availability, the departure time is updated to an earlier time. This procedure is repeated until all late deliveries have been removed. If the algorithm does not find a feasible solution without late deliveries, at the end of the algorithm, an empty scenario tour plan ϕ^ω and a scenario value $\tilde{V}^\omega(S^i_{t+1}) = -\infty$ is returned. For the original decision problem (5), this results in not offering the corresponding delivery option i to the current customer c_t. Note that since a tour plan can now comprise more than just one tour per vehicle, we add an index k to the tour notation, i.e., θ^{vk} denotes the k^{th} tour of vehicle v. This procedure is more formally presented in Algorithm 1 in Appendix D.

4.2. Value approximation and demand-management decision

The heuristic presented in Section 4.1 is used to solve the d-SDD-DRMP for scenarios $\omega \in \Omega^i_t$. In this way, we generate scenario-specific tour plans ϕ^ω. Those are anticipatory in the sense that they anticipate future customer requests. However, the demand management that was considered for those requests, only comprised accept/reject decisions. In particular, neither did it involve prices for delivery options and the respective rewards $r^i_{c_t}$, nor were choice and no-choice probabilities of the sampled customers and their resulting delivery revenues considered. Thus, in order to determine accurate scenario values, denoted as $\tilde{V}^\omega(S^i_{t+1})$, these aspects have to be captured retrospectively and integrated into the scenarios' solutions. In the following, we first give a verbal description of the idea underlying our re-integration of demand management. Afterwards, we formalize the respective procedure and show how to approximate a state value:

Idea underlying the re-integration of demand management: The main target is to reconstruct demand-management decisions for all a scenario's sampled customers in such a way that the same scenario tour plan would result as in the scenario's d-SDD-DMTP solution. Consequently, for every sampled customer request of such a solution, it has to be determined which offer sets provoke purchase choices with which the corresponding scenario solution is feasible. Across those offer sets, the expected contribution for every sampled customer is maximized and a close estimation of the scenario value $\tilde{V}^\omega$ can be determined. More precisely, in order to derive the best possible estimate, we want to imitate, as closely as possible, the original demand management of the SDD-DMTP as the Bellman equation (Eq. (5)) solved it. Therefore, imagine solving Eq. (5) by hand: in a first step one would intuitively define the feasible action space by excluding all infeasible decisions from being considered. For the demand-management decision this means determining which delivery options can be feasibly offered to the current customer, i.e., defining the set of feasible offer sets. In a next step, the offer set with the highest expected sum of immediate reward and successor state value is offered to the requesting customer. This last step includes making tour-planning decisions.

For re-integrating demand management into a scenario's solution ϕ^ω, this previously described procedure is mimicked with two modifications:

(1) When identifying the offer sets for the accepted, sampled customer requests $c \in N^\omega \cap \{L(\theta^{vk}) : \theta^{vk} \in \phi^\omega\}$ that are feasible with respect to the scenario's d-SDD-DMTP solution, all resulting tour-planning decisions have already been determined. Thus, the specific delivery times for customer orders a^{vk}_c, are already defined. Consequently, delivery options are only feasible, if a^{vk}_c can be matched within the delivery option.
(2) When selecting which offer set to offer, only the expectation regarding the immediate rewards is considered. Displacement cost and marginal cost to serve can be neglected.

The second modification can be made without sacrificing accuracy because the scenario solution, i.e., the acceptance and delivery times of all requesting customers in the scenario under consideration, has already been decided. Thus, it does not matter whether the currently considered customer chooses one of the offered delivery options or the no-purchase option. The value that might be incurred with subsequent customer requests will not change for this scenario.

Formalization and value approximation: More formally, re-integrating demand management into a scenario's solution ϕ^ω can be described as follows: For every $c \in N^\omega \cap \{L(\theta^{vk}) : \theta^{vk} \in \phi^\omega\}$, the procedure determines which delivery options $i \in \mathcal{I}$ can feasibly be offered according to their planned delivery time a^{vk}_c when following ϕ^ω. Next, for each of those customers, a subset $\mathcal{G}'_c(\phi^\omega) \subset \mathcal{G}$ defines all offer sets that include only the valid delivery options i. To approximate the sampled customer's contribution $r_{c\ \phi^\omega}$ to a scenario's value $\tilde{V}^\omega(S^i_{t+1})$, the expected reward across all $g \in \mathcal{G}'_c(\phi^\omega)$ is maximized: $r_{c\ \phi^\omega} = \max_{g\in\mathcal{G}'_c(\phi^\omega)} \sum_{i\in g} P^i(g) \cdot (r^i_c + r^i)$, if a customer order c is being accepted in the scenario's solution, otherwise $r_{c\ \phi^\omega} = 0$.

A scenario's value is then defined as $\hat{V}^\omega(S^i_{t+1}) = \sum_{c\in N^\omega} r_{c\ \phi^\omega} - r^l_{\phi^\omega}$. Following this, $V'(S^i_{t+1})$ is approximated by

$$\hat{V}'(S^i_{t+1}) = \frac{\sum_{\omega\in\Omega^i_t} \tilde{V}^\omega(S^i_{t+1})}{\mid \Omega^i_t \mid}. \tag{8}$$

Finally, the SDD-DMTP's demand-management decision is taken by substituting (8) in the Bellman equation. That yields the following demand-management decision policy for when a customer request arrives:

$$g^* = \arg\max_{g\in\mathcal{G}} \left(\sum_{i\in g} P^i(g) \cdot [r^i(g) + r^i_{c_t} + \hat{V}'(S^i_{t+1})] \right). \tag{9}$$

Note that the value approximation described above relies on solving scenarios ex-post, under the assumption that all customer arrivals were known. This could lead to a systematic over-estimation of the actual value of a state. However, for deciding on which offer set to present to an incoming customer, this over-estimation is not a major issue for the reason that when solving Eq. (9), not the absolute level of the values $\hat{V}'(S^i_{t+1})$ for $i \in g$ is decision-relevant, but the differences between them. As the potential over-estimation is systematic, it applies similarly to all those values.

Other approaches that approximate values/costs via heuristically solving scenarios ex-post in order to derive tour-planning decisions are for example Azi et al. (2012), Campbell & Savelsbergh (2005), and Angelelli et al. (2021).

4.3. Anticipatory tour planning

Having described how we approximate values to make demand-management decisions based on tour plans resulting from a scenario's d-SDD-DMTP solution (in the following referred to as scenario tour plans), we now explain how tour-planning decisions are derived.

For every potential customer choice i and the corresponding successor state S^i_{t+1}, a set of scenario tour plans $\phi^\omega \in \Omega^i_t$ with values $\tilde{V}^\omega(S^i_{t+1})$ is available from the scenarios' solutions. These can be used to derive tour-planning decisions. Typically, in MSAs, at this point, a consensus function measures the robustness of partials of those tour-plans by evaluating, which partials appear most frequently among the solutions. Then, from the result, it constructs a robust overall tour plan, called a distinguished plan (see for example in Bent & Van Hentenryck, 2004, Voccia et al., 2019). Due to the large number of stochastic influences in our problem, i.e., customer location, request arrival time, and customer choice, the scenario solutions exhibit high variability. This is why typical consensus functions proved not to perform well in pre-tests. Therefore, we derive tour-planning decisions from the one sampled tour plan ϕ^ω, which has the highest value $\tilde{V}^\omega(S^{i'}_{t+1})$ of all tour plans in $\Omega^{i'}_t$. Note, we are fully aware of and accept that the derived tour plan's performance might naturally be lower in entirely different realizations. The highest value tour plan is selected as distinguished plan ϕ^* and comprises planned tours θ^{vk} for all $v \in \mathcal{V}$. The tours of one vehicle v start sequentially at given start times A^{vk} and they contain sampled and confirmed customer orders. Then, in line with the literature on MSAs, all sampled customer orders are removed from those tours and the delivery times a^{vk}_c of all remaining confirmed customer orders c, as well as the return times to the depot, are updated according to A^{vk} and relevant $\tau_{cc'}$. This procedure is more formally described in Algorithm 2 in Appendix E.

An executable tour at state S^i_{t+1}, derived from the tour-planning decision ϕ^* for vehicle $v \in \mathcal{V}$ is denoted as θ^{*vk}. All θ^{*vk} for $v \in \mathcal{V}, k \in \mathcal{K}$ of ϕ^* form the tour-planning decision ϕ_{t+1} in t and for all subsequent t' until a new customer request arrives. If a new customer request arrives, the full decision-making procedure as presented in Fig. 1 starts all over again. For the tour-planning decisions that means all tours in ϕ_{t+1} that have not already started by the time of the new customer request, can be revised.

Note that the tour-planning decisions are based on predictions into the future, which means they consider potential future customer requests, potential time-steps in between future customer arrivals, and, especially, also future vehicle departures. Thus, a tour-planning decision in t also includes potential future tour-planning decisions. Accordingly, it is not necessary to revise tour plans if no customer request arrives in a new decision epoch. Instead, in decision epochs in which no customer request arrives, the provider analyses the latest MSA solution to derive tour-planning decisions. More precisely, the provider checks whether, for the current decision epoch, a new tour was planned to start and if yes, which customer orders are assigned to it. Then, the tour-planning decision for the current decision epoch with no customer request is derived respectively. Consequently, unlike the MDP model of the SDD-DMTP, the solution approach is not defined across all decision epochs t in the booking period. Instead, it is event driven, i.e., customer request arrival driven.

Regarding the literature discussed in Section 2, our solution approach falls in the class of non-learning approaches. It uses an information model internally, i.e., for decision making, in a predictive matter (Soeffker et al., 2021) and is conducted fully online.

5. Computational study

In this section, we present a computational study on a variety of parameter settings for which we apply our solution approach in different variants, e.g., with different lengths of the sampling horizon. Additionally, we solve some benchmark approaches and compare the results. In particular, we assess the effectiveness of our approach and evaluate the value of anticipation, as well as that of an explicit price optimization. In Section 5.1, we describe the parameters of the settings under consideration and explain how instances are generated. In Sections 5.2 and 5.3, we discuss our extensive computational experiments' results on the two evaluations, i.e., of anticipation and of explicit price optimization.

5.1. Setup

The computational study is based on a number of *settings* that we examine in a stochastic simulation, applying and comparing different anticipation and pricing approaches. In Section 5.1.1, we specify the parameters that are commonly used throughout all considered settings. In Section 5.1.2, we discuss the parameters that may vary across settings. In Section 5.1.3, we describe how we generate instances for each setting within our stochastic simulation.

5.1.1. Setting-independent parameters

The following parameters are defined identically for all settings considered in our computational study.

Time horizon and delivery options: The considered time horizon corresponds to the booking and service course of one day. It is represented by 900 episodes, which could be thought of as representing 900 minutes from 7am to 10pm. The booking period consists of 600 minutes, i.e., it starts at 7am and ends at 5pm. The service period starts with the first accepted customer order and ends at 10pm, latest. In all settings, offer sets can be generated based on two possible delivery options, i.e., delivery within 90 minutes or within 300 minutes.

Customer segments: Customers are defined by a segment affiliation, their location, their arrival times, and arrival rates. A customer's segment affiliation defines the potential contribution margins of selected shopping baskets. More precisely, it indicates a probability distribution across the potential contribution margins in connection with a purchase decision. Further, it defines their utility for different delivery options with different prices. In our computational study, we assume there are two segments, distinguishing between *segment-one* customers and *segment-two* customers. The contribution margin of a segment-one customer is drawn from a uniform distribution over $[75, 85, 100]$ monetary units (MU). The contribution margin of a segment-two customer is drawn from a uniform distribution over $[20, 35, 40]$ MU. Additionally, segment-one customers have a higher observable utility for shorter delivery options than segment-two customers. The basic observable utilities before pricing u^i_{basic} of segment-one customers are 22 and 14, and those of segment-two customers are 13 and 10.5 for the short and the long delivery options. To calculate the observable utility for a delivery option with a certain price u^i, the corresponding basic utility u^i_{basic} is reduced by the offered price r^i, but it cannot be negative, i.e., $u^i = \max\{u^i_{basic} - r^i, 0\}$. Also, the no-purchase option has a utility for customers from both segments. For segment-one customers, this utility equals 2, while for segment-two customers it equals 3. This reflects that segment-two customers are more likely to purchase via a traditional, non-SDD delivery option or in a brick-and-mortar store.

We model the purchase probabilities for different delivery options within the offer sets according to a basic attraction model. Therefore, the purchase probabilities for delivery options i in an offer set $g \in \mathcal{G}$ can be calculated by solving $P^i(g) = \frac{u^i}{\sum_{i \in g} u^i}$ (Luce, 1959).

Service area and customer locations: We simulate the service area on a squared grid with a width of 120 distance units (DU), with a centrally located depot. On this grid, we generated 200 customer locations from a uniform distribution in advance, which we will use later on in instance generation. Travelling a DU equals one minute in the simulation run and costs 0.3 MU. Thus, all potential customer locations on this grid can be visited within 120 minutes, i.e., if vehicle capacity allows, every customer can at least be offered the longer delivery option.

Customer arrivals: In every decision epoch t, customers arrive according to their individual, time-dependent arrival rates $\lambda_{c_t}(t)$. Thereby, decision epochs are sufficiently small that at most one customer arrives. In practice, these arrival rates can be derived from historic data. For our computational study, we assume the following pattern regarding the sum of customer specific arrival rates (cumulative arrival rate) per decision epoch: We assume two peaks in the cumulative arrival rate in order to mimic common online shopping behavior, namely customers placing orders during their lunch break or after returning home from work. Further, we design arrival rates in a way that the cumulative arrival rate over the lower valued segment-two customers forms lower and wider peaks than for the higher valued segment-one customers. This reflects more flexible working conditions with lower income for segment-two customers and, vice versa, less flexible working conditions with higher income for segment-one customers. The distribution of arrival rates is illustratively depicted in Fig. 4.

Pricing approach: We assume two price points per delivery option: 8 or 10 MU for guaranteed delivery within 90 minutes and 5 or 7 MU for guaranteed delivery within 300 minutes. We also assume that no delivery option other than the no-purchase option needs to be offered. These pricing parameters result in nine potential price lists from which the provider can select one to offer to an incoming request.

5.1.2. Setting-dependent parameters

The settings we consider differ in terms of the expected number of incoming customer requests and the number of delivery vehicles. More specifically, we consider settings resulting from each

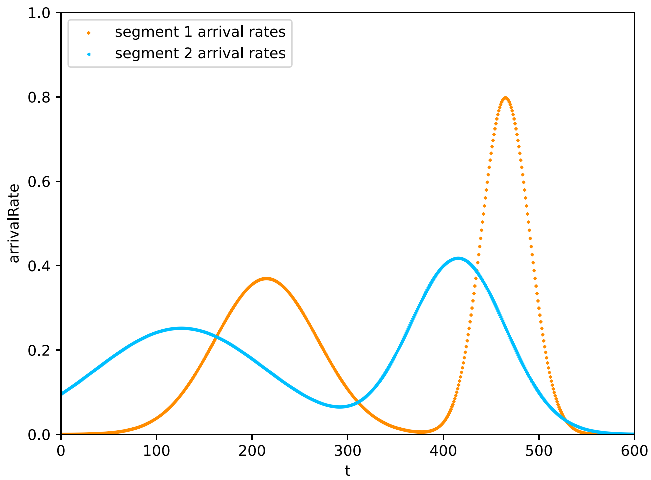


Fig. 4. Customer arrivals.

Table 3
Setting-dependent parameters.

		customer requests		
		100	150	200
vehicles	1	1V_100	1V_150	1V_200
	2	2V_100	2V_150	2V_200
	3	3V_100	3V_150	3V_200

possible combination of 100, 150, and 200 expected customer requests with one, two, and three delivery vehicles. The corresponding settings are shown in Table 3.

5.1.3. Instance generation

To ensure comparability, we test our approach and the benchmark approaches on the same set of registered customers, which we refer to as the *customer base*. More precisely, based on the customer segments' and customer locations' characteristics described in Section 5.1.2, we initially generate a customer base of 3000 different customers once. Thirty percent of the customers in this customer base are segment-one customers. Then, for each setting, instances basically represent particular demand streams that we obtain by event-based discrete simulation based on the arrival rates and according to the setting's expected number of customers. Requests' characteristics are obtained by sampling from the customer base. We generate 300 instances for each setting. Note that, again to ensure comparability, we use the same 300 instances for settings that differ only in the number of delivery vehicles.

5.2. Value of anticipation

In the following, we discuss the value of anticipation for the SDD-DMTP with respect to the developed approach as presented in Section 4.

5.2.1. Experimental design and performance metrics

In studying the impact of different levels of anticipation, we apply different variants of our approach. They differ as to the length of the sample horizon used for approximating the scenario values and tour planning (see Section 4). We consider sample horizon lengths of 30, 60, 90, and 120 minutes. Here, we base the decision making on the anticipation of a total of 15 scenarios, a number that led to good decisions in the pre-tests we performed. We sample the scenarios by drawing new customer requests from the customer base each time a decision has to be taken. Further, we benchmark our anticipatory approach against myopic decision making. Myopic decisions are taken in exactly the same way as in the anticipatory approach, except that all potential successor state values in Eq. (5) are set to 0. Further, the tour-planning decisions are taken without anticipated customer requests. Thus, in this approach the demand-management decision is based only on myopic marginal costs of serving a request.

Table 4
Averaged results 1V_100.

	look-ahead				
	myopic	30	60	90	120
# segment 1 customers	5.01	7.14	8.71	9.85	10.67
# segment 2 customers	14.16	12.36	10.53	8.77	7.27
# 90 minutes choice	1.25	1.60	1.44	1.21	1.67
# 300 minute choice	17.92	17.90	17.80	17.41	16.26
average price 90 minutes	8.00	8.49	8.55	8.69	8.78
average price 300 minutes	5.00	5.36	5.45	5.52	5.53
active vehicle minutes	809	760	734	704	686

To measure performance, for each setting and each length of the sample horizon, we evaluate the deviation from the myopic benchmark with respect to the following metrics:

Metric		Description
Revenue Shopping Baskets	(RSB)	sum of contribution margins of all shopping baskets sold in one instance
Revenue Deliveries	(RD)	sum of delivery fees accrued by selling delivery options throughout one instance
Delivery Costs	(DC)	overall cost of delivery operations, i.e., all executed delivery tours in one instance
Contribution Margin	(CM)	RSB + RD - DC
Number Of Deliveries	(NOD)	number of accepted customer requests that turned into orders and are being served in the course of one instance

The deviation of a given metric from the myopic benchmark for a given setting with a given sample horizon length is determined as follows: We average the results of the metric across the 300 test instances of the setting under consideration, and compare them to the corresponding averaged values resulting from solving the same 300 instances with the myopic benchmark approach. For example, the deviation of the CM with a sample horizon length of 30 minutes from the myopic results is calculated by $\frac{CM^{30}}{CM^{myopic}} - 1$.

5.2.2. Numerical results

The results we obtained are shown in Fig. 5. On the tested settings, it is possible to achieve an increase in CM of 15 to 50%. First, the increase grows degressively as the sample horizon length increases, until it reaches a peak at a sample horizon length of 90 or 120 minutes for most settings. It then slowly decreases for longer sample horizon lengths, which is displayed in more detail in Appendix F.1), where we depict the absolute values of the mean CM across all 300 instances, as well as the corresponding 95%-confidence intervals. For almost all settings, these intervals of the myopic approach and the anticipatory approaches do not overlap. For those settings, this implies with a confidence of 95%, that the increase in CM results from our anticipation approach. The only setting in which the increase in CM is smaller than 10% and where 95%-confidence intervals overlap, is the setting with low resource scarcity, in which the myopic approach also yields good results. Regarding the degressive course of the CM increase with increasing sample horizon length, pre-tests have shown that using fewer samples flattens the growth and shifts the peak to a shorter sample horizon length. This is shown illustratively for setting 1V_100 in Fig. 6. Increasing the sample size does not significantly shift the peak to a longer sample horizon length. Table 4 shows further numerical results for the 1V_100 setting, namely the average absolute values of customer choices, the segments of customer orders,

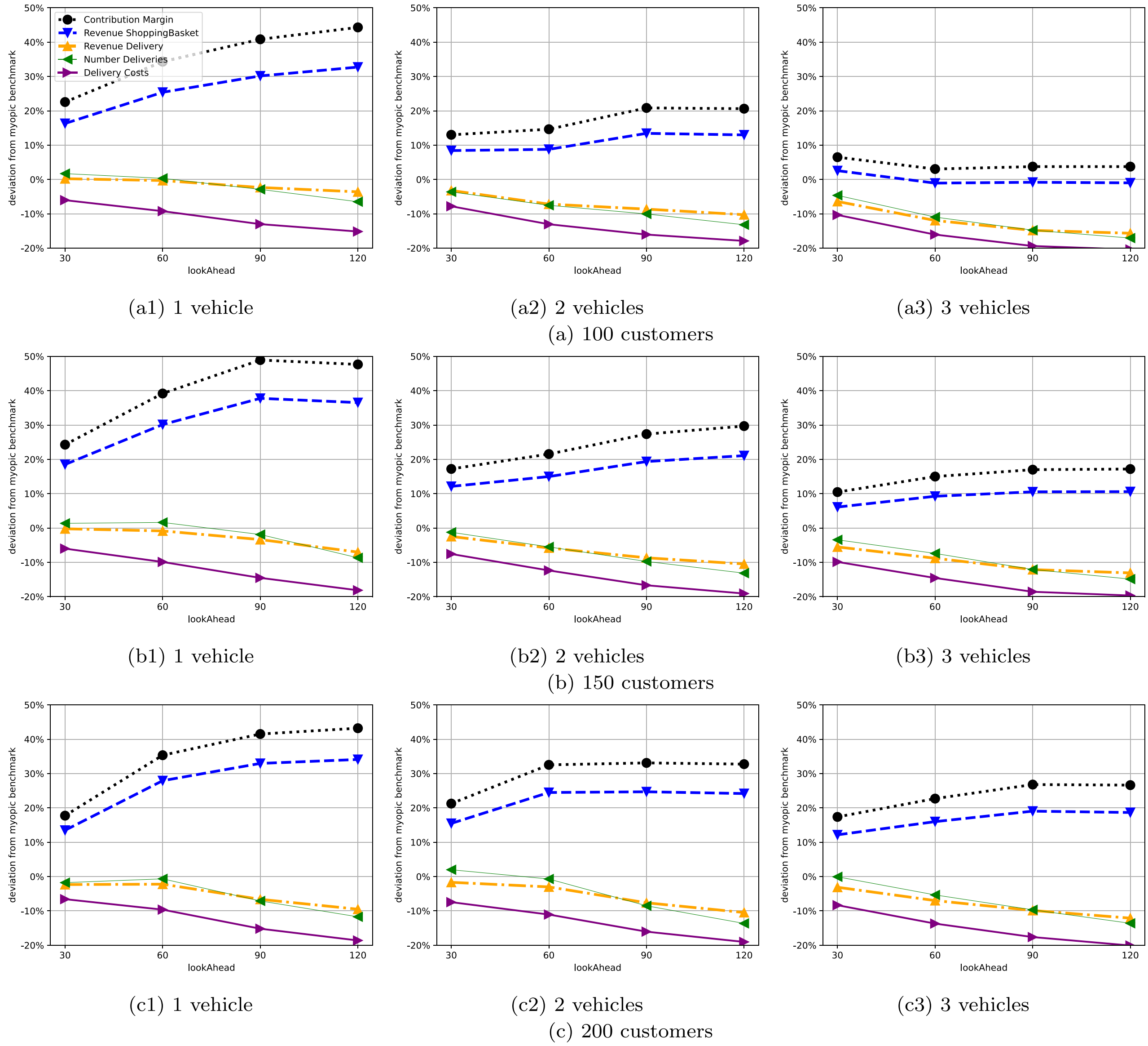


Fig. 5. Value of anticipation.

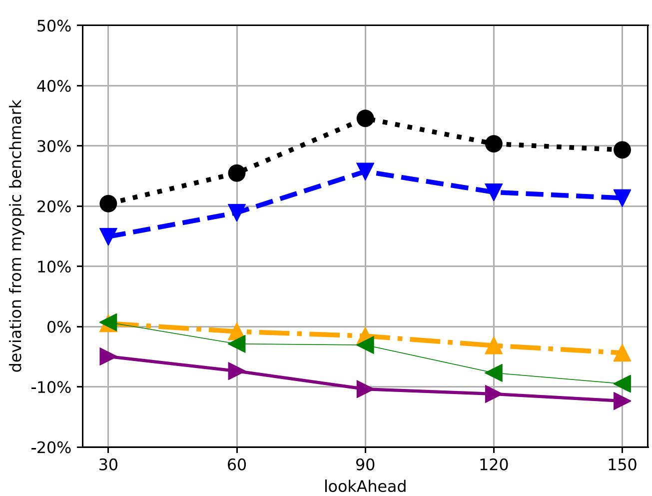
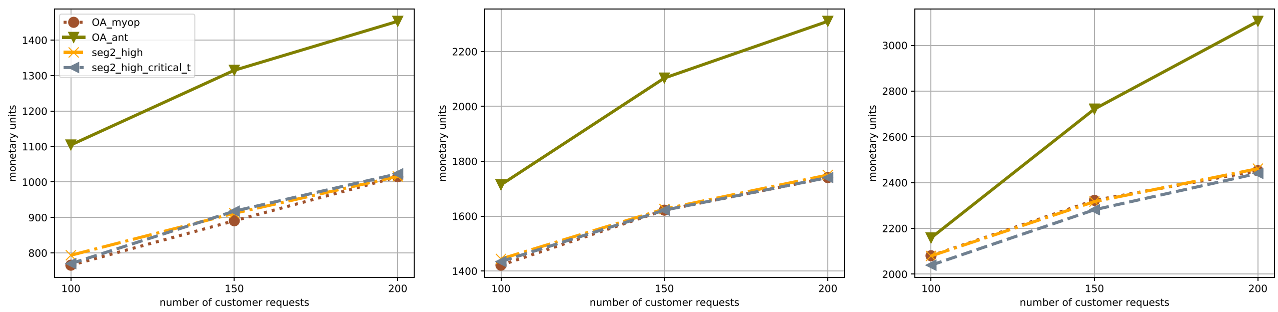


Fig. 6. 1V_100, 5 samples.

and the average prices paid for delivery options per instance. Here, we observe that as the length of the sampling horizon increases, the average number of highly valued customer orders accepted in an instance increases, and correspondingly, the average number of low-value customer orders accepted, decreases. Another trend observed is the increase in the average prices paid for the delivery spans, as the length of the sample horizon increases. The average number of customer choices for the different delivery spans shows no obvious pattern. All of these observations are representative of the results in the other instances, as can be seen in Appendix F.2.

To track down the demand-management that underlies the previously discussed trends, we further analyze the relationship of price lists offered and resulting customer choices, per customer segment. More precisely, for every customer segment, for the myopic and the best anticipatory approach (120 minutes look-ahead), we compare the partials of the different price lists offered, and the resulting ratios of customer choices. Again, we analyze the same 300 instances as before and summarized our results for time intervals of width 50 minutes. The results for setting 1V_100 are depicted in Fig. 8. The results for further settings are found in Appendix F.3. The price lists are represented by their prices with



(a) 1 vehicle (b) 2 vehicles (c) 3 vehicles

Fig. 7. Rule-based benchmark - Contribution margin.

the following pattern: the first element represents the price for the short delivery span, the second element is the price for the long delivery span, and the third element represents the no-purchase option with a price equal to 0*MU* in all price lists. If the price of a delivery span equals 100*MU*, the resulting choice probability equals 0 for all customer segments.

When comparing the anticipatory results with the myopic ones, it can be observed that the acceptance rate, i.e., the ratio of customer requests that were offered any delivery span for a price lower than 100*MU*, decreases over all time intervals for segment 2 customers. In turn, the respective numbers increase for all segment 1 customers. Further, it can be observed for both customer segments, that the anticipatory approach accepts less customer requests at the beginning of the booking horizon. Additionally, for segment 2 customers, i.e., customers with a low valued shopping basket, the ratio of customer requests being accepted substantially decreases for certain time intervals. Those time intervals correspond to the time intervals, in which the demand of segment 1 customer rises.

This last observation led to the idea, that our anticipatory approach could be imitated by a simple rule-based demand-management policy, if the demand pattern is known. To test this hypothesis, we derived two rule-based approaches, in which there is no explicit anticipation. Instead, the rule-based decision making is following these demand-management patterns observed in our anticipatory solutions, i.e., to lower the demand of segment two customers. Thus, we evaluated the following two simple rules for demand-management decisions: The first rule ('seg2-high') is to only offer the higher prices for each feasible delivery span to customers with a shopping basket value $\leq 50MU$. The second rule ('seg2-high-critical-t') is to only offer the higher prices for each feasible delivery span to customers with a shopping basket value $\leq 50MU$, if they request in certain decision epochs. Those decision epochs were derived from analyzing the results of our anticipatory approach and are the intervals [100;250] and [400; 500]. The average CM that can be observed with these rule based approaches is depicted in Fig. 7. The respective results for MOD, RSB, RD, and DC are depicted in Appendix F.4. It can be observed, that the two rule-based approaches yield comparable results as our original myopic approach ('OA-myop') and much worse results than our original anticipatory approach ('OA-ant').

5.2.3. Analysis and insights

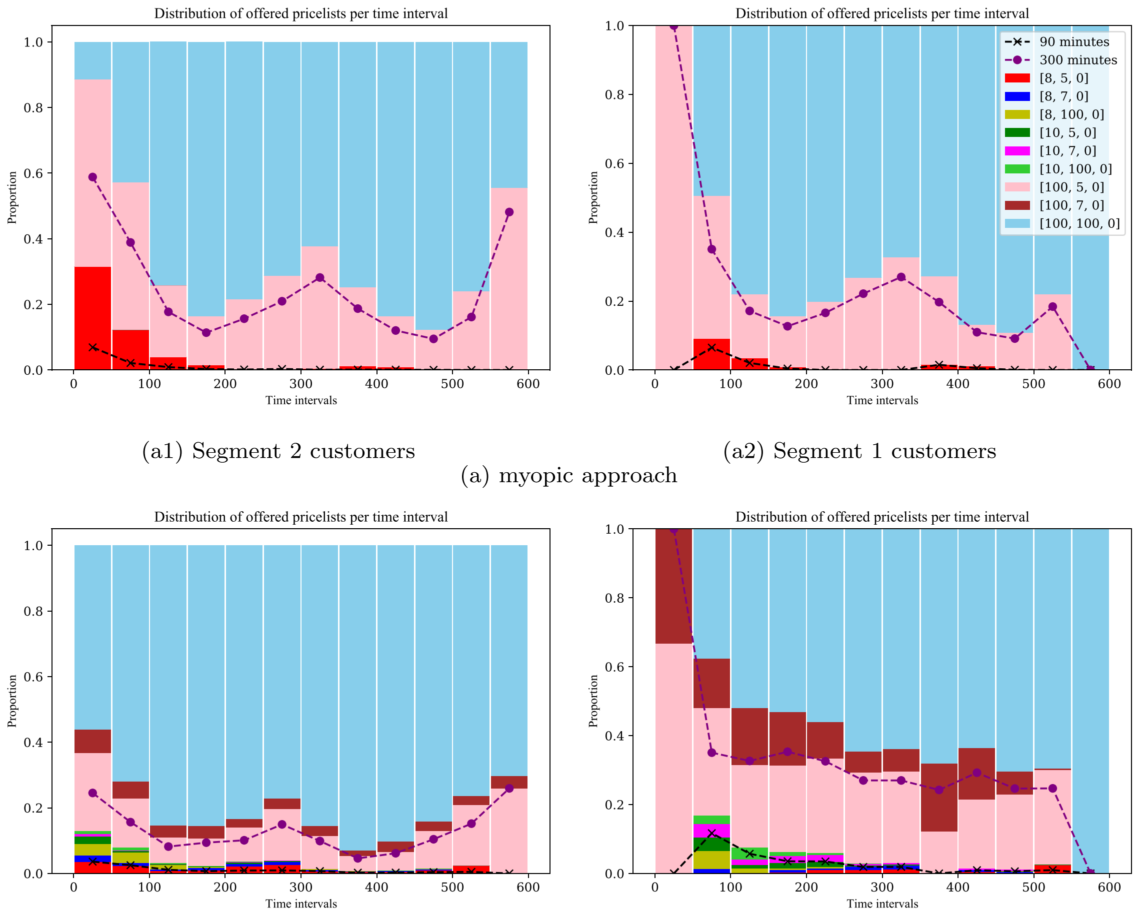
According to our observation, the contribution margin that can be achieved with anticipation is always higher than the contribution margin of any myopic benchmark. This is mainly due to the fact that the revenues generated by selling shopping baskets increase and the delivery costs decrease disproportionately to the decrease in delivery orders. Combined with Table 4, Fig. 8, and Appendices F.2 and F.3, this shows that anticipation indeed allows us to preserve capacity for high-value customer orders, and also to generally guide customer choice with respect to a favorable spacial structure. Thus, compared to myopic decision making, through anticipation delivery efficiency can be improved. Further, we observed a degression in the increase of contribution margin with an increasing sample horizon length. Such degression is explained by the lengths of the sample horizon becoming longer, and as this happens, the proportion of uncertainty in decision making increases. Thus, these results indicate that the solutions' quality decreases if the sample horizon is too long or if too few samples are used. This is because, for every decision, increasing the sampling horizon length also increases the number of sampled, and hence uncertain requests, while the number of certain orders does not increase. Additionally, due to the tight delivery spans that distinguish SDD from other last mile logistics services, all certain orders in the scenarios will be served shortly after the time when the sampling starts. Hence, sampling into the future too far leads to decision making based on tours that include only uncertain orders. This distorts the precision of the value approximation.

5.3. Value of explicit pricing optimization

Here, we elaborate the value of the explicit pricing approach as described in Section 5.1.1, and compare it with three benchmark pricing approaches.

5.3.1. Experimental design

To determine the value of (explicitly) using a pricing optimization model within our approach, we benchmark three variants. The first pricing benchmark reflects pure availability control, in which the provider can only decide whether to offer certain delivery options or not. Thereby, all prices are set to the corresponding lower prices from the explicit pricing approach described in Section 5.1.1. The second pricing approach equals the first, but prices are set corresponding to the higher prices from the explicit pricing approach. The third pricing benchmark replaces solving an explicit pricing optimization problem in our approach by a simple pricing rule based on opportunity cost estimation, which mimics an idea followed by Ulmer (2020a). If a delivery option's calculated opportunity costs are low, its base price (as before the lower price point used in the explicit pricing optimization) is set. If the opportunity costs of an option exceed this base price, the price is set to the opportunity costs. For calculating opportunity cost, Ulmer (2020a) follows a definition by Yang et al. (2016). They define opportunity cost as the difference between the values of the states that result from rejecting a customer and those from accepting the customer (for a certain delivery option). In our benchmark study, we also follow this definition and calculate opportunity cost accordingly, based on state values resulting from our approximation



(a1) Segment 2 customers (a2) Segment 1 customers

(a) myopic approach

(b1) Segment 2 customers (b2) Segment 1 customers

(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. 8. Offers and choices per customer segment - myopic and anticipatory - 1V_100.

approach (see Section 4.2). We refer to the first benchmark as 'AC-BP-low' (for 'availability control with low base prices'), to the second benchmak as 'AC-BP-high' (for 'availability control with high base prices'), and to the third as 'OCBP' (for 'opportunity costs based pricing'). Further, we refer to our explicit pricing approach as 'OP' (for 'original pricing approach').

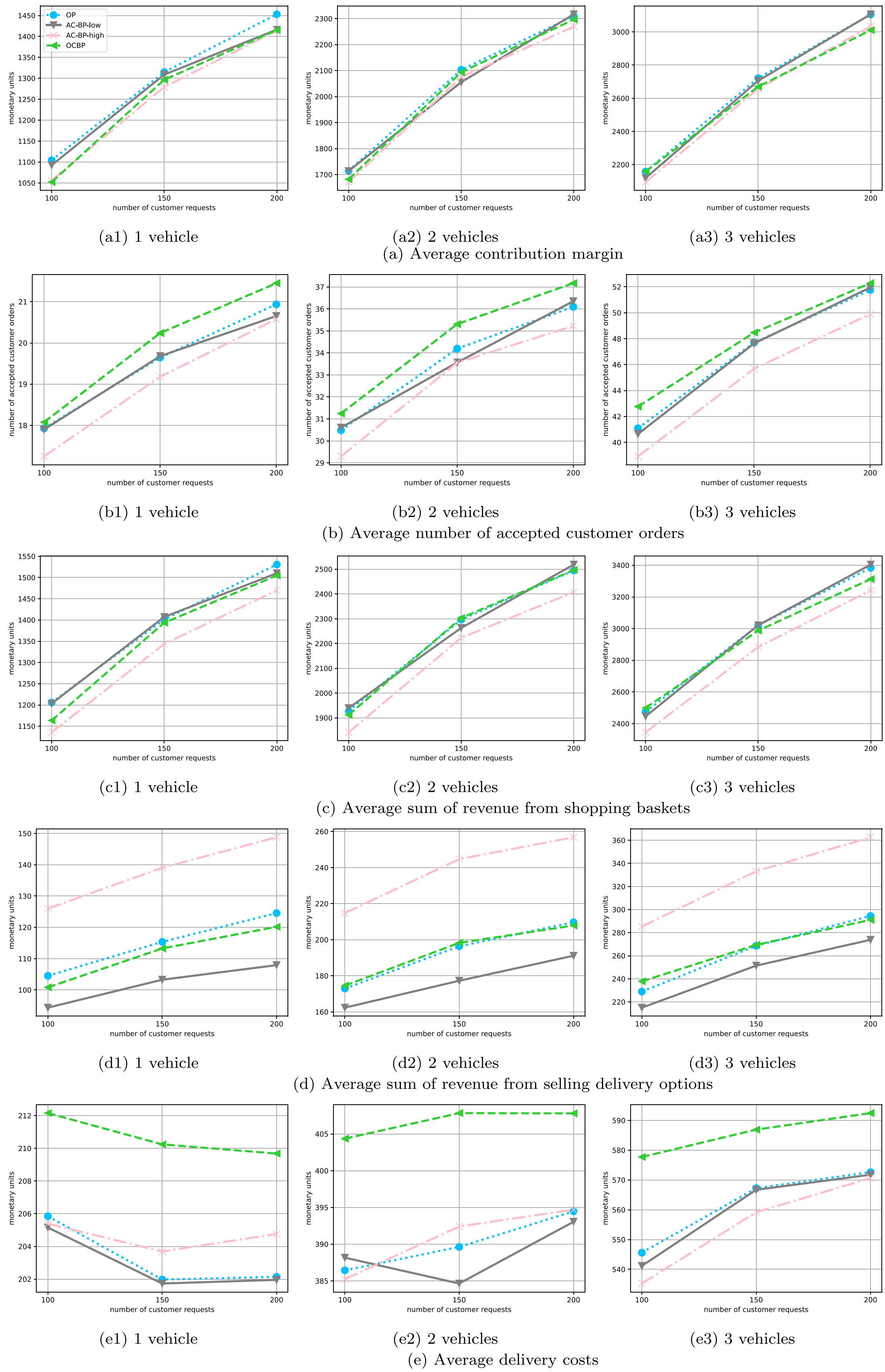
We conduct the study on the same 300 instances for each setting as in Section 5.2.1. We approximate state values, and thus also the opportunity costs, by averaging the values of 15 samples across a sample horizon of 120 minutes length. Based on the analysis in Section 5.2.2, this has proven to be the best combination for the considered settings. In this way, we minimize the effects of bad opportunity cost estimation by sub-optimal sampling horizon lengths/number of samples. Again, we measure performance by evaluating the average of the contribution margins, the number of accepted customer orders, the sum of revenues from shopping baskets and from selling delivery options, as well as of the delivery costs.

5.3.2. Numerical results

The obtained results are given in Fig. 9. Although the results of the average contribution margins are close, the OP yields better results than the benchmark approaches in nearly all settings. Only in the settings with 200 customers, with two as well as with three vehicles, does the AC-BP-low yield a higher averaged CM; however, the results of the OP are exceeded by less than 0.5% and 0.005%, respectively. In the setting with 100 customers and three vehicles, the OCBP yields a less than 0.05% higher CM than the OP (see Fig. 9a). The OCBP, on average, accepts the most customer requests of all settings (see Fig. 9b), but at most settings its average RSB falls below the other approaches' RSB. Also, it yields a substantially higher DC for all settings and yields the highest RD in only three settings, where it does not substantially exceed the RD of the OP. In most instances the AC-BP-low accepts the lowest number of customer requests, also with substantially lower RD than the other approaches, but it still accrues a comparably high RSB. It even exceeds the other approaches' RSB in four settings. Also, the AC-BP-low yields the lowest DC of all instances except one.

5.3.3. Analysis and insights

The results in Section 5.3.2 show that the different pricing approaches rely on three different levers to increase the CM, and that each of the various approaches exploits those levers to a different extent. The levers we observed are (1) increasing the overall revenue by setting higher prices where possible (mainly observed for the OCBP and the AC-BP-high), (2) increasing the overall revenue by preserving capacity for high-value customer orders (mainly observed for the BP and AC-BP-low), and (3) reduce overall delivery costs by steering customer choices toward the most efficient delivery options and rejecting those requests that negatively affect routing efficiency (mainly observed for the AC-BP-low and OP). The OCBP has the highest pricing flexibility, as prices originate from a



(a1) 1 vehicle (a2) 2 vehicles (a3) 3 vehicles

(a) Average contribution margin

(b1) 1 vehicle (b2) 2 vehicles (b3) 3 vehicles

(b) Average number of accepted customer orders

(c1) 1 vehicle (c2) 2 vehicles (c3) 3 vehicles

(c) Average sum of revenue from shopping baskets

(d1) 1 vehicle (d2) 2 vehicles (d3) 3 vehicles

(d) Average sum of revenue from selling delivery options

(e1) 1 vehicle (e2) 2 vehicles (e3) 3 vehicles

(e) Average delivery costs

Fig. 9. Pricing benchmark II.

continuous range instead of being chosen from a predefined, finite set of price points. Therefore, this approach can exploit lever (1) the most (see Fig. 9d) and hence can also accept the most customer requests. Still, regarding the CM, for most settings the OCBP performs worse than the other approaches due to exploiting levers (2) and (3) less effectively. This can be derived from the lower or under-proportionally higher RSB (see Fig. 9c), and from the over-proportionally higher DC (9e).

The AC-BP-low and the AC-BP-high, in turn, have the lowest pricing flexibility. Thus, the AC-BP-low cannot exploit lever (1) as the much lower RD (Fig. 9d) shows. On the contrary, the AC-BP-high yields a high RD, but cannot exploit lever (3) and, thus, also yields substantially higher delivery cost. Generally, the AC-BP-low is a performant approach regarding the exploitation of levers (2) and (3). We observed the same for the OP, as well as recognizing that the OP also exploits lever (1). In addition to exploiting lever (1), the OP enables us to enlarge the provider's service provision, as the OP can offer delivery of customer requests that the AC-BP-low would deny and the customers can themselves decide whether to accept or reject the corresponding offer.

6. Conclusion and outlook

In this paper, we investigated the SDD demand-management and tour-planning problem, with special attention to explicitly incorporating two types of required decisions, namely demand-management decisions and tour-planning decisions. The problem under consideration is characterized by overlapping booking and service periods. This adds an online tour-planning component to the demand-management problem, which itself is computationally intractable. Thus, it makes the overall problem substantially more difficult to optimize than related problems dealt with in the literature.

We have developed a non-learning based solution approach that provides integrated decision making for the two types of decisions and does not require extensive offline learning. In this approach, both decisions are anticipatory and based on the combination of two central ideas – multiple scenario approaches for online tour-planning and approximation of state values – which is done by averaging across sampled trajectories, such as those known from rollout algorithms.

In the first part of our extensive numerical study, we assessed how our approach performed at different levels of anticipation. The assessment showed that anticipation can increase the contribution margin with as much as 10–50% in our settings, especially if delivery resources are scarce (in a low ratio of vehicles to customers). When we incorporated anticipation through sampling, we found that appropriately limiting the length of the sample-horizon can improve decision making. The main reason for this is that as the length of the sample horizon increases, decisions are made with increasing uncertainty. This is especially relevant for practical settings where booking and service periods overlap, as in the SDD case. If the sampling horizon is too long, anticipatory decisions are based on tours that contain only sampled orders and no confirmed ones.

In the second part of our study, we compared three different pricing approaches: pure availability control, our proposed explicit pricing approach, and a simple pricing rule based on opportunity cost. Comparing the different approaches, we found that as price flexibility increases (from fixed prices to a limited number of possible price points to possible prices from an unbounded continuous space), the quality of the resulting tours decreases. This demonstrates that the integrated state value approximation and decision-making approach does indeed allow us to steer customer choice toward efficient delivery options, while at the same time preserving capacity for high-value customer orders. Compared to the other two approaches, this one has the best ratios of number of customer requests accepted to the corresponding sum of revenues from shopping baskets, and delivery efficiency. Further, we found the approach that accepts the most customer requests is not necessarily the best in terms of contribution margin, as it yields the highest delivery costs. In practice, when choosing a pricing approach, one has to examine closely which is more relevant for long-term success – losing a customer's goodwill due to being rejected or due to higher delivery costs.

We believe that our study's results provide starting points for future efforts in several directions. The first direction concerns anticipation in solving integrated demand-management and tour-planning problems with overlapping booking and service periods. In future studies, it could be useful to examine hybrid anticipation approaches that combine learning based and non-learning based decision making. Thus, a good starting point would be to explore whether adding a previously learned end-of horizon valuation to the presented approach would improve its performance. The second direction concerns the pricing component of our approach and the different variations we compared. Our results show that an increase in price flexibility leads to a decrease in cost efficiency, which is a very interesting direction for deeper analysis, especially when dealing with continuous explicit price optimization and more complex customer choice models. The third direction concerns an entirely different, more revenue management oriented view. It would be very interesting to further investigate the hierarchical demand-management decomposition approach we have developed. Particularly, we could study how this approach performs in different environments and for different problems, e.g., with more complex pricing and choice models, and whether it would then still be possible to apply it in online algorithms.

Appendix A. Literature reviews addressing integrated demand management and online tour planning

In this section, we shortly discuss the existing surveys featuring integrated demand management and tour planning in home-delivery applications in general. We first outline surveys with a focus to *demand management*. Then, we discuss the respective literature with a focus on the *tour-planning* problem. At last, we review literature that considers both perspectives in an integrative manner.

Demand-management perspective – Agatz, Fleischmann, & Van Nunen (2008) provide the first review on the distributional challenges in e-fulfillment, including initial ideas to connect demand management and tour planning. The authors name two features of e-fulfillment systems that enable demand management. Those are pricing flexibility and extensive availability of data concerning purchasing behavior. In those two features, the authors see the foundation for segment-specific pricing as well as promotion. In a later review, Agatz et al. (2013) compare the demand-management-related processes of a large e-grocer with those prevalent in airline revenue management and elaborate similarities as well as decisive features of both concepts. Therewith, they provide starting points for incorporating differentiated and/or dynamic slotting/pricing into home delivery business concepts with a focus on the demand-management side. The same holds for Klein et al. (2020) who review recent generalizations and advances of revenue management techniques in traditional applications and new industry applications. They show how to transfer availability control to AHD problem settings and present the corresponding DP formulation.

Tour-planning perspective – Archetti & Bertazzi (2021) consider the problems under consideration with a focus on tour-planning aspects. They review recent advancement and challenges of home delivery systems. They see pricing as a measure to balance demand

Table B5
Anticipatory solution approaches for i-DMVRPs in LMD.

Authors	Appli-	Anticipatory		Learning	Objective	OC
	cation	DM	TP	based		
Asdemir, Jacob, & Krishnan (2009)	AHD	√	✗	analytical	rev	DPC
Lebedev, Goulart, & Margellos (2021)	AHD	√	✗	analytical	profit	√
Dumouchelle, Frejinger, & Lodi (2021)	AHD	√	✗	√	profit	-
Koch & Klein (2020)	AHD	√	✗	√	profit	√
Lang et al. (2021a)	AHD	√	✗	√	rev	DPC
Lang, Cleophas, & Ehmke (2021b)	AHD	√	✗	√	rev	DPC
Lebedev, Margellos, & Goulart (2020)	AHD	√	✗	√	profit	-
Ulmer & Thomas (2020)	AHD	√	✗	√	rev	-
Yang & Strauss (2017)	AHD	√	✗	√	profit	√
Vinsensius, Wang, Chew, & Lee (2020)	AHD	√	✗	√	cost	MCTS
Angelelli et al. (2021)	AHD	√	✗	✗	profit	√
Campbell & Savelsbergh (2005)	AHD	√	✗	✗	profit	√
Giallombardo, Guerriero, & Miglionico (2020)	AHD	√	✗	✗	profit	-
Klein, Mackert, Neugebauer, & Steinhardt (2018)	AHD	√	✗	✗	profit	√
Mackert (2019)	AHD	√	✗	✗	profit	√
Strauss, Gülpınar, & Zheng (2021)	AHD	√	✗	✗	profit	√
Yang et al. (2016)	AHD	√	✗	✗	profit	MCTS
Chen et al. (2019)	SDD	✗	√	√	accept	-
Chen, Wang, Thomas, & Ulmer (2020)	SDD	✗	√	√	accept&fair	-
Ulmer (2020b)	SDD	✗	√	√	accept	-
Ulmer, Mattfeld, & Köster (2018)	SDD	✗	√	√	accept	-
Ulmer et al. (2019)	SDD	✗	√	√	accept	-
Azi et al. (2012)	SDD	✗	√	✗	profit	-
Côté et al. (2021)	SDD	✗	√	✗	accept&cost	-
Klapp et al. (2018)	SDD	✗	√	✗	cost&serv	-
Klapp et al. (2020)	SDD	✗	√	✗	profit	-
Voccia et al. (2019)	SDD	✗	√	✗	cost	-
Prokhorchuk et al. (2019)	SDD	√	✗	√	rev&penalty cost	DPC
Soeffker et al. (2017)	SDD	√	✗	√	accept&fair	-
Ulmer (2020a)	SDD	√	✗	√	rev	DPC

among favorable and unfavorable delivery time windows, but do not further elaborate demand-management measures in particular. The same is true for the survey by Snoeck et al. (2020). Although the authors specifically address *revenue management* in AHD problem settings, they focus on the influences of potential extensions and future developments on the tour-planning component only. Boysen et al. (2021) survey research on home delivery problems with a focus on newly emerged business concepts and Soeffker et al. (2021) discuss the related stochastic dynamic vehicle routing problems (VRPs) and embed them into a prescriptive analytics framework. Both consider pricing as an essential decision dimension in existing business concepts that influences stochastic dynamic vehicle routing respectively.

Integrated perspective – The most recent and comprehensive survey of related literature that integrates both of the previously discussed perspectives, is the survey by Fleckenstein et al. (2022). The authors provide a generalized problem definition and outline AHD and SDD applications, as well as literature on mobility-on-demand (MOD). They propose a high-level, generic MDP modeling formulation and outline typically involved customer choice models. Further, they provide a comprehensive survey of general solution concepts and describe solution approaches for all involved subproblems, i.e., demand-management-related subproblems and tour-planning-related subproblems. Another holistic survey is Waßmuth et al. (2022). The authors review recent literature dealing with demand management in home delivery on the strategic, tactical and operational level and specifically differentiate between the two demand-management levers *offering* and *pricing*.

Appendix B. Tabular overview of anticipatory solution approaches for integrated demand management and online tour planning

Table B.5 summarizes the literature that addresses anticipatory solution approaches for AHD and SDD in home-delivery applications, as well as literature that deals with demand management for such home delivery applications analytically. The second column shows for which applications, i.e., AHD or SDD, a solution approach is designed. In the next two columns, it is indicated whether an approach involves anticipatory demand management (DM) (√) and/or tour planning (TP) (√) or not (✗). The fourth column shows whether the addressed anticipation is analytical, learning-based (√), or non-learning-based (✗). The fifth column summarizes the objectives addressed by an approach. The observed objectives are the maximization of revenue (rev), profit (profit), customer request acceptances (accept), fairness (fair), coverage of the service area (serv), the minimization of cost (cost), or a (hierarchical) combination of those objectives. Approaches that aim at minimizing the number of rejected customer requests are counted as those maximizing customer request acceptances. The last column shows whether opportunity cost are considered explicitly and, if so, whether they are considered comprehensively, accounting for displaced acceptances and variable fulfillment cost (√), or whether the displacement of expected revenue (DPC) or variable fulfillment cost (MCTS) are considered only.

Appendix C. MIP Formulation of the d-SDD-DMTP with first-tier demand management

The d-SDD-DMTP with accept/reject demand management is a deterministic profitable multi-trip vehicle routing problem with release and due times (PVRPRDT). It is defined across nodes for the already confirmed and not yet being delivered customer orders, a node representing the current customer request, and nodes for all sampled customers. Additionally, a node c_0 that represents a centrally located depot with coordinates $(x,y)_{c_0} = (0,0)$, is needed. Thus, the corresponding set of nodes $\mathcal{N}^{\omega}$ equals the following union: $\mathcal{C}_t \cup \{c_t\} \cup N^{\omega} \cup \{c_0\}$. For every confirmed customer order and for the current customer request, this set contains information about the customer's location $(x,y)_c$ and their confirmed delivery deadline t_c^{due}. For every sampled customer request, the set contains information about the customer's location $(x,y)_c$, their request time t_c^{req}, the reward their requested shopping-basket r_c will bring, and their utility u_c^i for delivery options $i \in \mathcal{I}$. Further, all sampled customers $c \in N^\omega$ are assigned a preliminary delivery deadline according to the longest available delivery span, $t_c^{due} = t_c^{req} + \max\{l(i) \mid i \in \mathcal{I}\}$. The underlying idea is that if those customers are included in a scenario's solution, it is always possible to offer them at least one, namely the longest, delivery option when their request realizes. This ensures that it is always possible to feasibly reconstruct such a solution with the actual demand management for value approximation.

In the d-SDD-DMTP, a number V of homogeneous vehicles operates a chronologically ordered number of tours $k \in \mathcal{K} = 1..K$. $\zeta_{cc'}$ represents the costs of travelling from the location of customer order c to the location of customer order c'. $\rho_{c'}$ is a customer order individual penalty which equals the value of the shopping basket for all $c' \in N^\omega$ and equals a very high number M for all $c' \in \mathcal{C}_t \cup \{c_t\}$. Since $\mathcal{C}_t$ only contains customer orders for which a feasible solution (without delays and dropped visits) is available, these penalties ensure that no confirmed customer order is dropped when solving the model. The parameter t describes the current decision period. The following decision variables are included in the model:

$$x_{cc'}^{vk} = \begin{cases} 1 & \text{if customer } c' \text{ is served after customer c on tour k by vehicle v} \\ 0 & \text{else} \end{cases} \quad \forall c,\ c' \in \mathcal{N}^\omega : c \neq c',\ v \in \mathcal{V},\ k \in \mathcal{K}$$

$a_{c'}^{vk}$	$\geq$	t	$\forall c' \in \mathcal{N}^\omega,\ k \in \mathcal{K},\ v \in \mathcal{V}$	Delivery time at customerlocation c' on tour k of vehicle v
A^{vk}	$\geq$	t	$\forall k \in \mathcal{K},\ v \in \mathcal{V}$	Departure time of tour k of vehicle v
B^{vk}	$\geq$	t	$\forall k \in \mathcal{K},\ v \in \mathcal{V}$	Time of finishing tour k of vehicle v in the depot

The d-SDD-DMTP can be formulated as the following MIP, which is further explained below:

$$\min \sum_{v\in\mathcal{V}}\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{N}^\omega}\sum_{c'\in\mathcal{N}^\omega} x_{cc'}^{vk}\cdot\zeta_{cc'} + \sum_{c'\in\mathcal{N}^\omega\setminus\{0\}}(1-\sum_{v\in\mathcal{V}}\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{N}^\omega} x_{cc'}^{vk})\cdot\rho_{c'} \tag{C.1}$$

s.t.

$$\sum_{v\in\mathcal{V}}\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{N}^\omega} x_{cc'}^{vk} \leq 1 \quad \forall c' \in \mathcal{N}^\omega \setminus \{c_0\} \tag{C.2}$$

$$A^{vk} \leq a_0^{vk} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K} \tag{C.3}$$

$$t_{c'}^{req}\cdot\sum_{c\in\mathcal{N}^\omega} x_{cc'}^{vk} \leq A^{vk} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K},\ c' \in \mathcal{N}^\omega \setminus \{c_0\} \tag{C.4}$$

$$t_{c'}^{due} + (1-\sum_{c\in\mathcal{N}^\omega} x_{cc'}^{vk})\cdot M \geq a_{c'}^{vk} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K},\ c' \in \mathcal{N}^\omega \setminus \{c_0\} \tag{C.5}$$

$$a_{c'}^{vk} + (1-x_{cc'}^{vk})\cdot M \geq a_c^{vk} + x_{cc'}^{vk}\cdot\tau_{cc'} \quad \forall v \in \mathcal{V}, k \in \mathcal{K}, c \in \mathcal{N}^\omega, c' \in \mathcal{N}^\omega \setminus \{c_0\} \tag{C.6}$$

$$A^{vk} + \sum_{c\in\mathcal{N}^\omega}\sum_{c'\in\mathcal{N}^\omega\setminus\{c_0\}} x_{cc'}^{vk}\cdot\tau_{cc'} \leq B^{vk} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K} \tag{C.7}$$

$$B^{vk} \leq A^{vk+1} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K}\setminus\{K\} \tag{C.8}$$

$$\sum_{c\in\mathcal{N}^\omega} x_{cc'}^{vk} = \sum_{c\in\mathcal{N}^\omega} x_{c'c}^{vk} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K},\ c' \in \mathcal{N}^\omega \setminus \{c_0\} \tag{C.9}$$

$$\sum_{c'\in\mathcal{N}^\omega\setminus\{c_0\}}\sum_{v\in\mathcal{V}} x_{0c'}^{v0} \leq V \tag{C.10}$$

$$\sum_{c'\in\mathcal{N}^\omega\setminus\{c_0\}}\sum_{k\in\mathcal{K}} x_{0c'}^{vk} \leq K \quad \forall v \in \mathcal{V} \tag{C.11}$$

$$\sum_{c'\in\mathcal{N}^\omega} x_{0c'}^{vk} \leq 1 \quad \forall v \in \mathcal{V},\ k \in \mathcal{K} \tag{C.12}$$

$$\sum_{c'\in\mathcal{N}^\omega\setminus\{0\}} x_{c_0c'}^{v+1\ 0} \leq \sum_{c'\in\mathcal{N}^\omega\setminus\{c_0\}} x_{0c'}^{v0} \quad \forall v \in \mathcal{V} \setminus \{V\} \tag{C.13}$$

$$\sum_{c'\in\mathcal{N}^\omega\setminus\{0\}} x_{0c'}^{vk+1} \leq \sum_{c'\in\mathcal{N}^\omega\setminus\{c_0\}} x_{0c'}^{vk} \quad \forall v \in \mathcal{V},\ k \in \mathcal{K} \setminus \{K\} \tag{C.14}$$

The objective function (C.1) minimizes the overall travel costs and the sum of the penalties of all dropped visits. Dropping a sampled customer in the solution of the MIP means rejecting their request. Therefore, (C.1) balances the increase in travel costs for visiting a sampled customer and their shopping basket value – if marginal costs to serve and displacement costs are higher than a customer's shopping basket value, this customer request is rejected. Constraints (C.2) enable dropping visits/rejection of customer requests. Thus, in combination with the objective function, this represents the first-tier demand management. Constraints (C.3)–(C.8) are time restrictions, which ensure that a tour starts neither before t, nor before all allocated customer orders have realized, that all customer orders will be served on time, that the duration of a tour is the sum of all travel times of that tour, and that a vehicle can only start a new tour after having returned to the depot. Constraints (C.9) ensure flow conservation. Constraints (C.10)–(C.12) ensure that the number of available vehicles and the maximum number of tours are not exceeded. Constraints (C.13) and (C.14) are symmetry breaking constraints.

This MIP formulation is a generalization of a profitable multi-trip vehicle routing problem (PVRPMT), which additionally considers time restrictions. It is an adaption of the PVRPMT formulation of Chbichib et al. (2012) and of a multi-trip team orienteering problem with time windows formulation by Voccia et al. (2019).

Appendix D. Feasibility repair

Algorithm 1 Feasibility repair scenarios.

$\phi \leftarrow$ Heuristic solution of P-VRPTW
$\theta^v(\phi) \leftarrow$ tour of vehicle v according tosolution ϕ
$a_c^v \leftarrow$ Delivery time of customer c with vehicle v according to solution ϕ
for v in $\mathcal{V}$ **do**
    initialize first tour θ^{v1} by adding depot $c=0$ and customer order c with smallest a_c^v according to $\theta^v(\phi)$
    calculate current departure time: $A^{v\ next} \leftarrow a_c^v - \tau_{0c}$
    $\theta^{v\ next} \leftarrow \theta^{v1}$
    $\phi\ ^v \leftarrow \theta^{v\ next}$
    repeat
        $\theta^{v\ curr} \leftarrow \theta^{v\ next}$
        $A^{v\ curr} \leftarrow A^{v\ next}$
        $L(\theta^{v\ curr}) \leftarrow \{\}$
        $X(\theta^{v\ curr}) \leftarrow \{\}$
        add not yet planned customer orders c of $\theta^v(\phi)$ to $L(\theta^{v\ curr})$ with increasing a_c^v until $t_c^{req} \geq A^{v\ curr}$ and amend $X(\theta^{v\ curr})$ accordingly
        cut tour by adding depot return and calculate return time $B^{v\ prev}$
        $\theta^{v\ curr} \leftarrow (A^{v\ curr}, L(\theta^{v\ curr}), X(\theta^{v\ curr}))$
        append $\theta^{v\ curr}$ to ϕ^v
        initialize next tour $\theta^{v\ next}$ by adding depot $c=0$
        add not yet planned customer order c with next smallest a_c^v according to $\theta^v(\phi)$
        calculate latest departure time: $A^{v\ next} \leftarrow \max\{a_c^v - \tau_{0c}, B^{v\ prev}\}$
    until all customer orders c in $\theta^v(\phi)$ are planned to tours
    for *tour* in ϕ^v **do**
        update all $a_c^{v\ tour}$ according to $A^{v\ tour}$ and travel times $\tau_{cc'}$
        if $a_c^{v\ tour} \geq t_c^{due}$ for any sampled customer order c in $\theta^{v\ tour}$ **then**
            remove c from $\theta^{v\ tour}$ and update all left $a_c^{v\ tour}$ according to $A^{v\ tour}$ and travel times $\tau_{cc'}$
        if $a_c^{v\ tour} \geq t_c^{due}$ for any confirmed customer order c in $\theta^{v\ tour}$ **then**
            repeat
                remove sampled customer order c with highest t_c^{req}
                update $A^{v\ tour}$ according to vehicle availability
                update all left $a_c^{v\ tour}$ according to $A^{v\ tour}$ and travel times $\tau_{cc'}$
            until there are no longer any late deliveries
$\phi \leftarrow \{\phi^v : v \in \mathcal{V}\}$

Appendix E. Tour-planning decision and post-processing

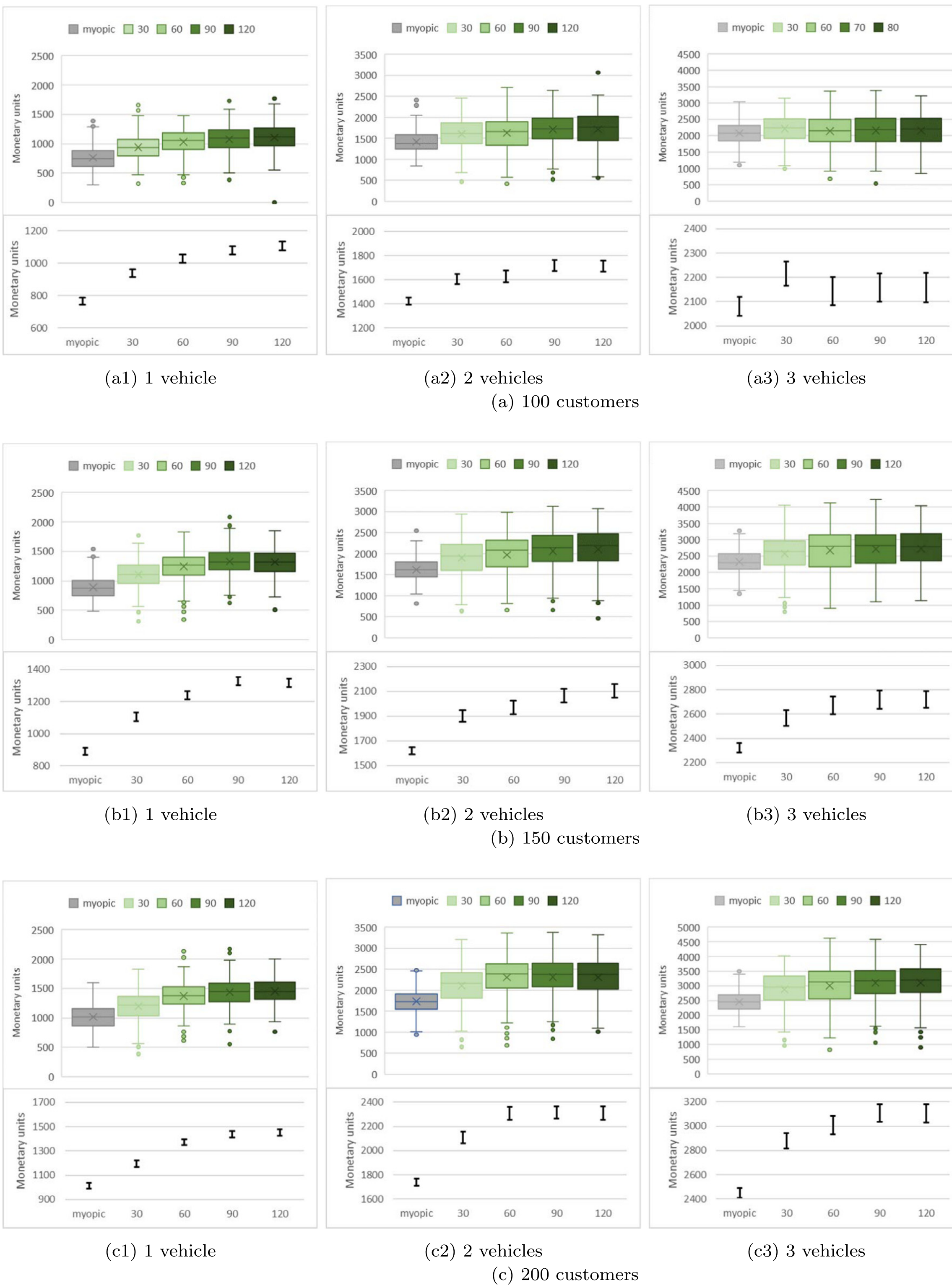
Algorithm 2 Tour-planning decision and post-processing.

$i \leftarrow$ customer choice
$S_{t+1}^i \leftarrow$ regarding successor state
$\phi^* \leftarrow \arg\max_{\{\phi^\omega : \omega \in \Omega_t^i\}} \tilde{V}^\omega(S_{t+1}^i)$

for $v \in \mathcal{V}$ **do**
    for $k \in \mathcal{K}$ **do**
        Remove all sampled customers $c \in \theta^{vk}$
        for remaining customers $c \in \theta^k$ **do**
            Update a_c^{vk} according to A^{vk} and travel and service times $\tau_{cc'}$
        $\theta^{*vk} \leftarrow \theta^{vk}$
$\phi^* \leftarrow \{\theta^{*vk} : v \in \mathcal{V}, k \in \mathcal{K}\}$
$\phi_{t+1} \leftarrow \phi^*$

Appendix F. Basic setting - Further numerical results

F1. Mean contribution margins - Boxplots and confidence intervals



(a1) 1 vehicle (a2) 2 vehicles (a3) 3 vehicles

(a) 100 customers

(b1) 1 vehicle (b2) 2 vehicles (b3) 3 vehicles

(b) 150 customers

(c1) 1 vehicle (c2) 2 vehicles (c3) 3 vehicles

(c) 200 customers

Fig. F1. Mean contribution margins across 300 simulation runs: Boxplots and 95%-confidence intervals.

F2. Averaged results on customer segments, purchase choice, and prices paid

look-ahead	myopic	30	60	90	120
# segment 1 customers	5.01	7.14	8.71	9.85	10.67
# segment 2 customers	14.16	12.36	10.53	8.77	7.27
# 90 minutes choice	1.25	1.60	1.44	1.21	1.67
# 300 minute choice	17.92	17.90	17.80	17.41	16.26
average price 90 minutes	8.00	8.49	8.55	8.69	8.78
average price 300 minutes	5.00	5.36	5.45	5.52	5.53
active vehicle minutes	809	760	734	704	686

Table 6: 1 vehicle – 100 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	9.93	12.76	13.65	15.47	16.13
# segment 2 customers	25.19	21.08	18.82	16.12	14.36
# 90minutes choice	3.33	4.13	3.35	2.75	2.95
# 300 minute choice	31.78	29.71	29.13	28.84	27.53
average price 90 minutes	8.00	8.31	8.34	8.42	8.50
average price 300 minutes	5.00	5.24	5.31	5.34	5.37
active vehicle minutes	1569	1447	1364	1317	1288

Table 7: 2 vehicle – 100 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	15.61	17.68	17.98	19.05	19.61
# segment 2 customers	33.94	29.58	26.14	23.18	21.49
# 90minutes choice	6.11	6.93	5.16	3.50	3.94
# 300 minute choice	43.44	40.32	38.96	38.74	37.16
average price 90 minutes	8.00	8.32	8.31	8.40	8.44
average price 300 minutes	5.00	5.17	5.21	5.22	5.27
active vehicle minutes	2285	2048	1917	1842	1819

Table 8: 3 vehicle – 100 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	5.69	8.63	10.54	12.39	13.06
# segment 2 customers	15.82	13.17	11.33	8.72	6.59
# 90minutes choice	1.06	1.52	1.49	1.34	1.56
# 300 minute choice	20.45	20.28	20.38	19.77	18.08
average price 90 minutes	8.00	8.43	8.58	9.00	8.90
average price 300 minutes	5.00	5.39	5.55	5.61	5.61
active vehicle minutes	822	773	741	703	673

Table 9: 1 vehicle – 150 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	10.82	14.73	16.49	18.82	20.24
# segment 2 customers	28.55	24.15	20.72	16.72	13.95
# 90minutes choice	3.21	4.17	3.63	2.82	3.17
# 300 minute choice	36.16	34.71	33.57	32.72	31.02
average price 90 minutes	8.00	8.36	8.43	8.59	8.54
average price 300 minutes	5.00	5.31	5.40	5.46	5.46
active vehicle minutes	1605	1484	1407	1337	1299

Table 10: 2 vehicle – 150 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	15.87	19.45	22.07	24.26	25.23
# segment 2 customers	40.17	34.67	29.83	25.02	22.47
# 90minutes choice	5.84	6.92	5.22	3.65	4.36
# 300 minute choice	50.20	47.19	46.68	45.62	43.34
average price 90 minutes	8.00	8.29	8.38	8.42	8.48
average price 300 minutes	5.00	5.23	5.29	5.32	5.34
active vehicle minutes	2354	2121	2011	1917	1891

Table 11: 3 vehicle – 150 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	6.46	9.15	11.74	13.58	14.58
# segment 2 customers	17.25	14.14	11.82	8.46	6.35
# 90minutes choice	1.02	1.58	1.69	1.14	1.60
# 300 minute choice	22.70	21.72	21.87	20.89	19.34
average price 90 minutes	8.00	8.53	8.78	9.11	9.00
average price 300 minutes	5.00	5.43	5.61	5.68	5.70
active vehicle minutes	828	773	748	702	674

Table 12: 1 vehicle – 200 customers

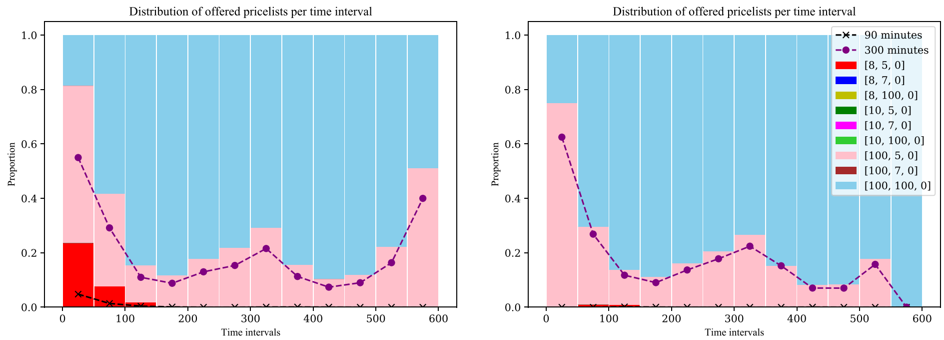
look-ahead	myopic	30	60	90	120
# segment 1 customers	11.34	15.85	19.48	21.37	22.64
# segment 2 customers	30.47	26.78	22.02	16.87	13.46
# 90minutes choice	3.11	4.05	3.69	2.69	3.23
# 300 minute choice	38.70	38.58	37.82	35.56	32.87
average price 90 minutes	8.00	8.36	8.47	8.62	8.65
average price 300 minutes	5.00	5.37	5.49	5.54	5.53
active vehicle minutes	1624	1503	1444	1363	1315

Table 13: 2 vehicle – 200 customers

look-ahead	myopic	30	60	90	120
# segment 1 customers	15.73	21.32	24.79	27.91	29.12
# segment 2 customers	44.19	38.56	31.93	26.21	22.64
# 90minutes choice	5.91	6.84	5.45	4.16	4.54
# 300 minute choice	54.01	53.04	51.27	49.96	47.22
average price 90 minutes	8.00	8.33	8.40	8.50	8.53
average price 300 minutes	5.00	5.28	5.34	5.40	5.41
active vehicle minutes	2388	2189	2061	1968	1909

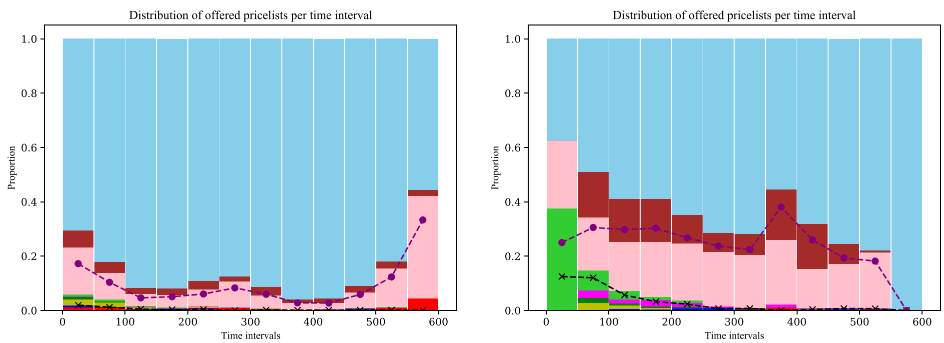
Table 14: 3 vehicle – 200 customers

F3. Offers and choices per customer segment



(a1) Segment 2 customers (a2) Segment 1 customers

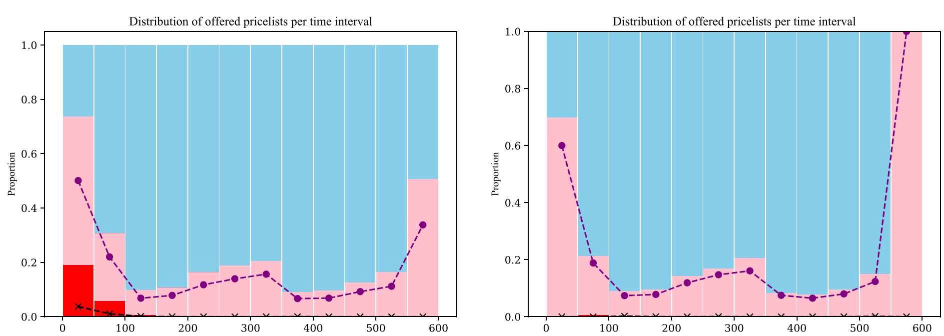
(a) myopic approach



(b1) Segment 2 customers (b2) Segment 1 customers

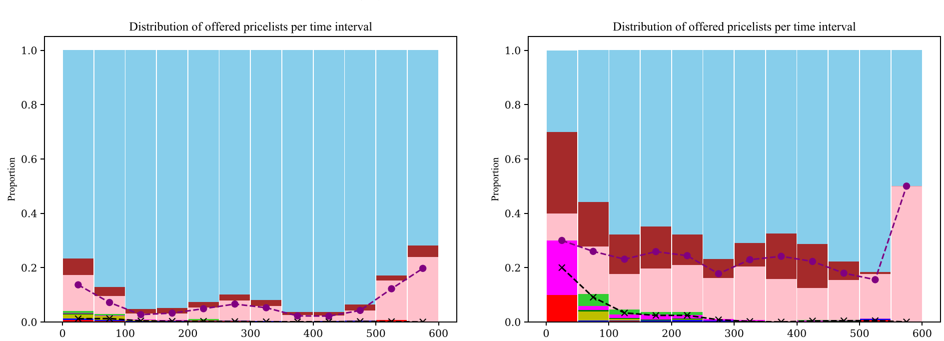
(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F2. Offers and choices per customer segment - myopic and anticipatory - 1V_150.



(a1) Segment 2 customers (a2) Segment 1 customers

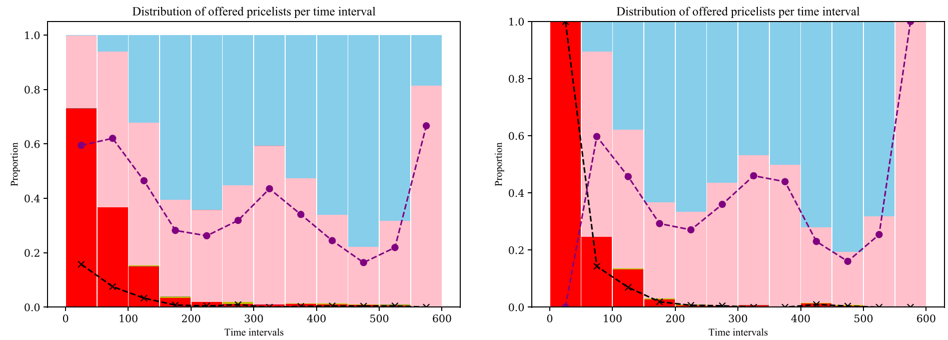
(a) myopic approach



(b1) Segment 2 customers (b2) Segment 1 customers

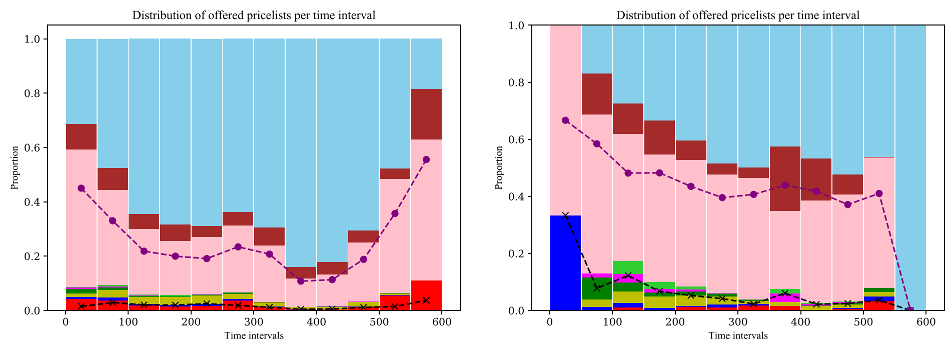
(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F3. Offers and choices per customer segment - myopic and anticipatory - 1V_200.



(a1) Segment 2 customers (a2) Segment 1 customers

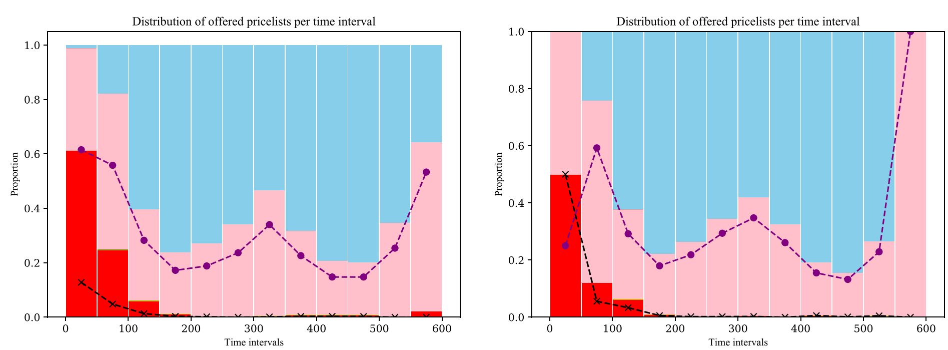
(a) myopic approach



(b1) Segment 2 customers (b2) Segment 1 customers

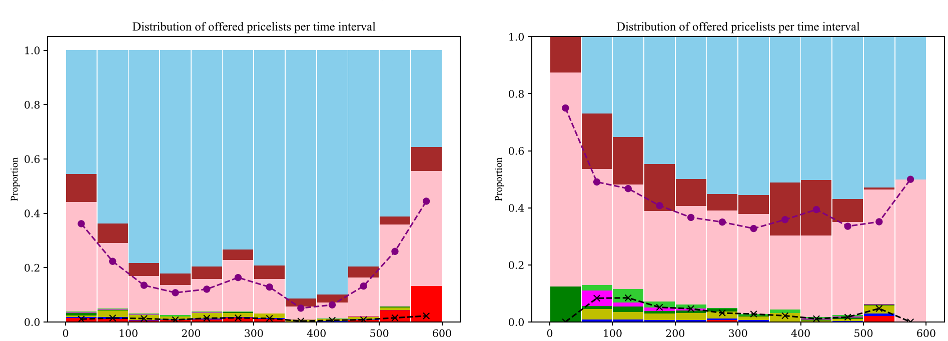
(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F4. Offers and choices per customer segment - myopic and anticipatory - 2V_100.



(a1) Segment 2 customers (a2) Segment 1 customers

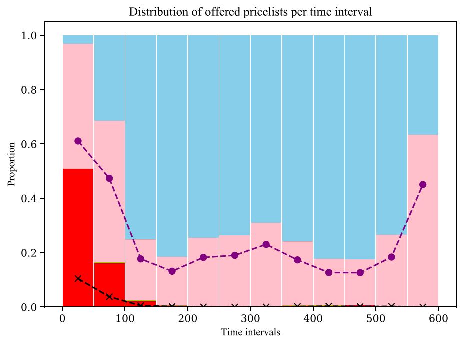
(a) myopic approach



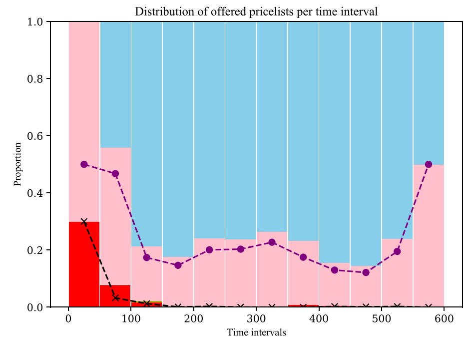
(b1) Segment 2 customers (b2) Segment 1 customers

(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F5. Offers and choices per customer segment - myopic and anticipatory - 2V_150.

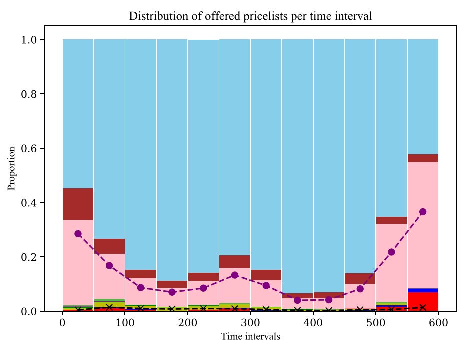


(a1) Segment 2 customers

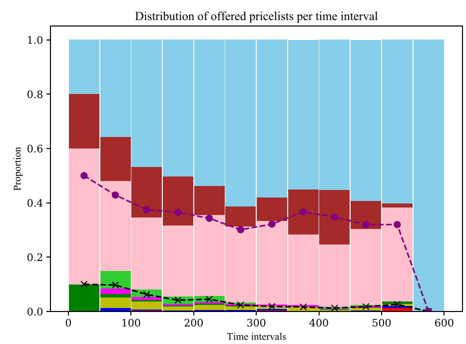


(a2) Segment 1 customers

(a) myopic approach



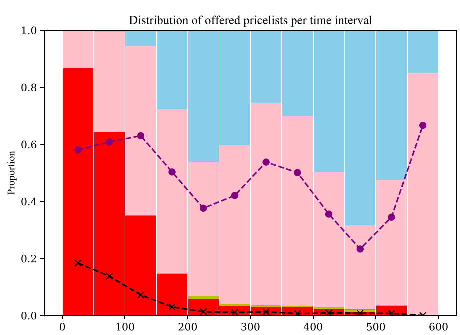
(b1) Segment 2 customers



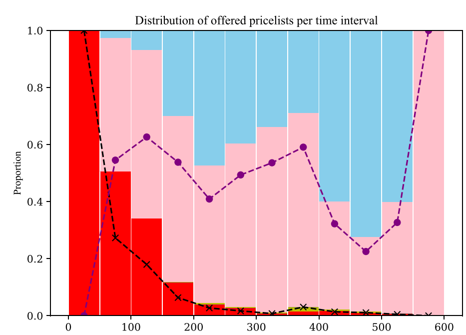
(b2) Segment 1 customers

(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F6. Offers and choices per customer segment - myopic and anticipatory - 2V_200.

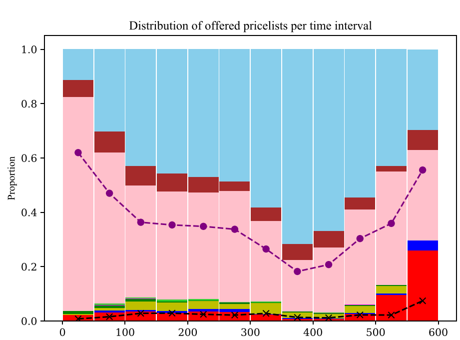


(a1) Segment 2 customers

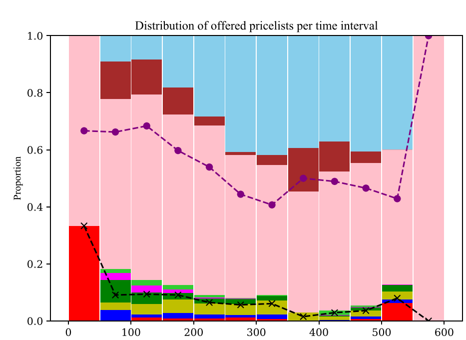


(a2) Segment 1 customers

(a) myopic approach



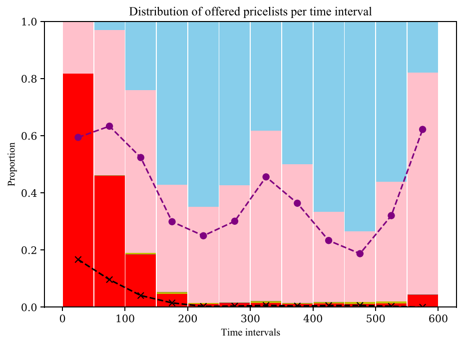
(b1) Segment 2 customers



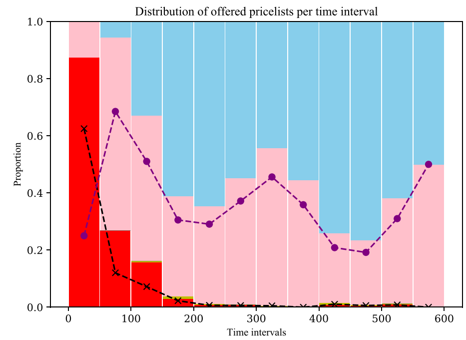
(b2) Segment 1 customers

(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F7. Offers and choices per customer segment - myopic and anticipatory - 3V_100.

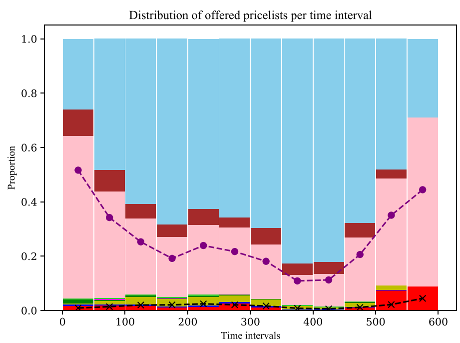


(a1) Segment 2 customers

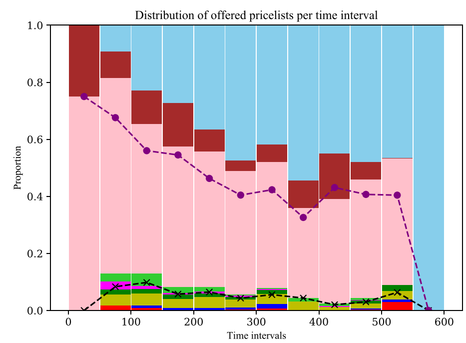


(a2) Segment 1 customers

(a) myopic approach



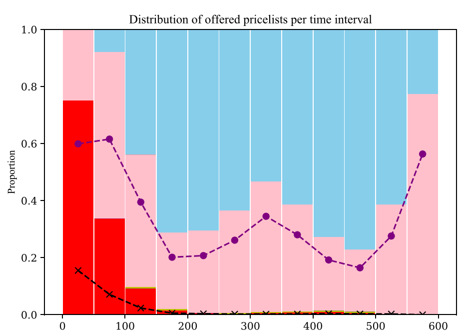
(b1) Segment 2 customers



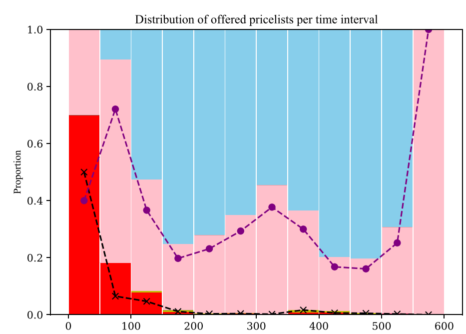
(b2) Segment 1 customers

(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F8. Offers and choices per customer segment - myopic and anticipatory - 3V_150.

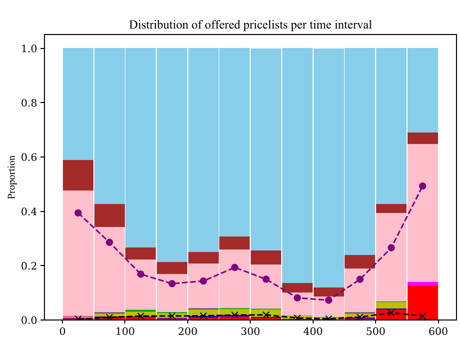


(a1) Segment 2 customers

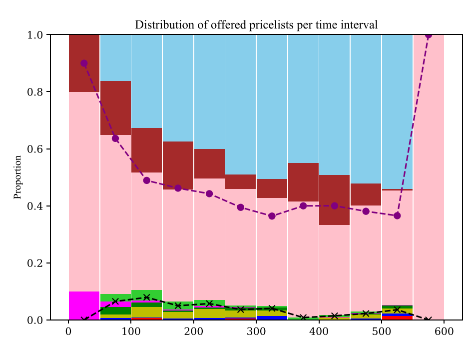


(a2) Segment 1 customers

(a) myopic approach



(b1) Segment 2 customers

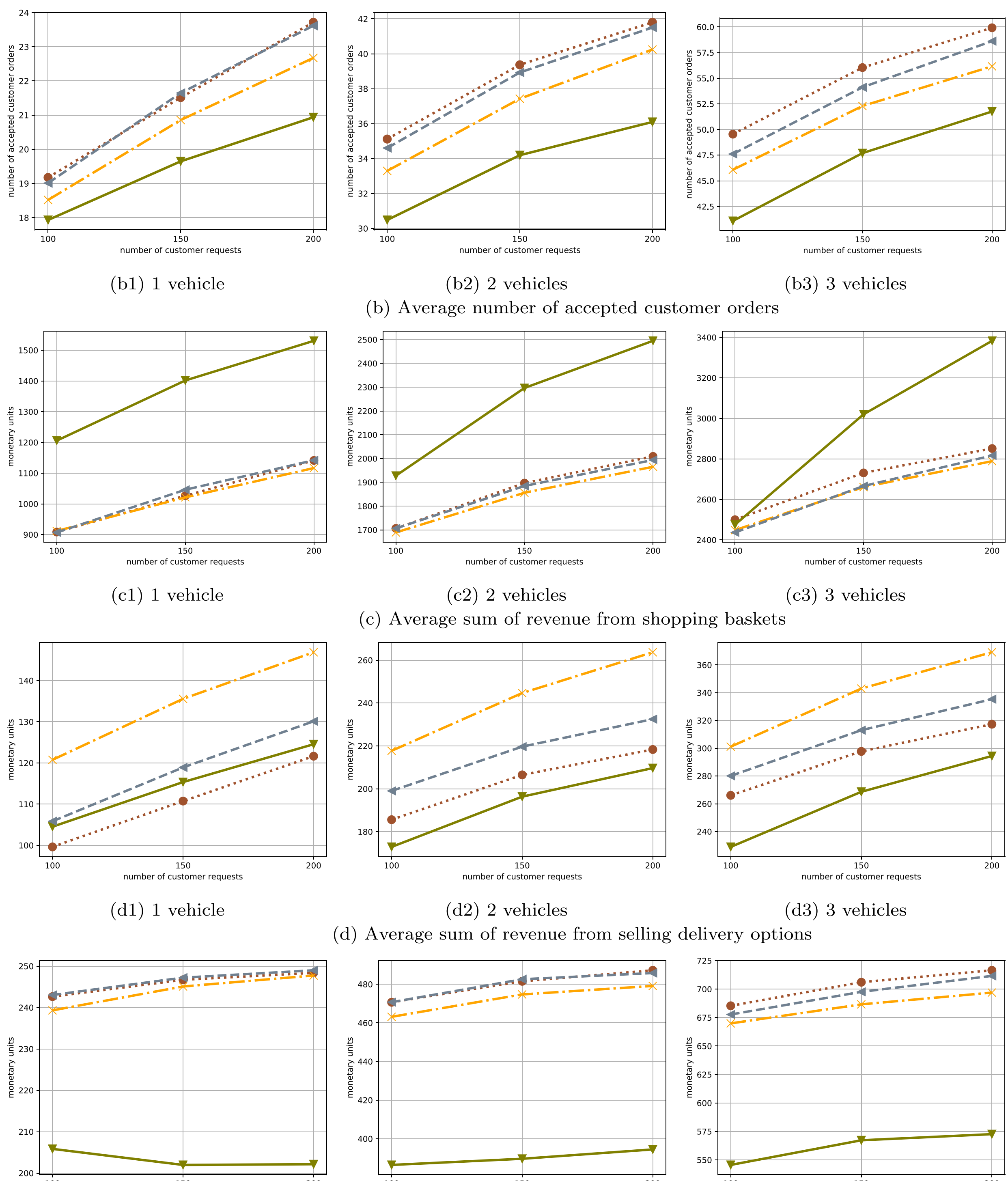


(b2) Segment 1 customers

(b) Anticipatory approach - 15 samples, 120 minutes sampling horizon

Fig. F9. Offers and choices per customer segment - myopic and anticipatory - 3V_200.

F4. Rule-based benchmark results



(b1) 1 vehicle (b2) 2 vehicles (b3) 3 vehicles

(b) Average number of accepted customer orders

(c1) 1 vehicle (c2) 2 vehicles (c3) 3 vehicles

(c) Average sum of revenue from shopping baskets

(d1) 1 vehicle (d2) 2 vehicles (d3) 3 vehicles

(d) Average sum of revenue from selling delivery options

(e1) 1 vehicle (e2) 2 vehicles (e3) 3 vehicles

(e) Average delivery costs

Fig. F10. Rule-based benchmark.

References

Agatz, N., Campbell, A. M., Fleischmann, M., Van Nunen, J., & Savelsbergh, M. (2013). Revenue management opportunities for internet retailers. *Journal of Revenue and Pricing Management, 12*(2), 128–138.

Agatz, N. A., Fleischmann, M., & Van Nunen, J. A. (2008). E-fulfillment and multi--channel distribution–a review. *European Journal of Operational Research, 187*(2), 339–356.

Angelelli, E., Archetti, C., Filippi, C., & Vindigni, M. (2021). A dynamic and probabilistic orienteering problem. *Computers & Operations Research, 136*, 105454.

Archetti, C., & Bertazzi, L. (2021). Recent challenges in routing and inventory routing: E-commerce and last-mile delivery. *Networks, 77*(2), 255–268.

Asdemir, K., Jacob, V. S., & Krishnan, R. (2009). Dynamic pricing of multiple home delivery options. *European Journal of Operational Research, 196*(1), 246–257.

Azi, N., Gendreau, M., & Potvin, J. Y. (2012). A dynamic vehicle routing problem with multiple delivery routes. *Annals of Operations Research, 199*(1), 103–112.

Bent, R., & Van Hentenryck, P. (2007). Waiting and relocation strategies in online stochastic vehicle routing. *IJCAI, 7*, 1816–1821.

Bent, R. W., & Van Hentenryck, P. (2004). Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research, 52*(6), 977–987.

Berbeglia, G., Cordeau, J.-F., & Laporte, G. (2010). Dynamic pickup and delivery problems. *European Journal of Operational Research, 202*(1), 8–15.

Boysen, N., Fedtke, S., & Schwerdfeger, S. (2021). Last-mile delivery concepts: A survey from an operational research perspective. *Or Spectrum, 43*(1), 1–58.

Campbell, A. M., & Savelsbergh, M. W. (2005). Decision support for consumer direct grocery initiatives. *Transportation Science, 39*(3), 313–327.

Chbichib, A., Mellouli, R., & Chabchoub, H. (2012). Profitable vehicle routing problem with multiple trips: Modeling and variable neighborhood descent algorithm. *American Journal of Operational Research, 2*(6), 104–119.

Chen, X., Ulmer, M. W., & Thomas, B. W. (2019). Deep q-learning for same-day delivery with a heterogeneous fleet of vehicles and drones. arXiv preprint 1910.11901.

Chen, X., Wang, T., Thomas, B. W., & Ulmer, M. W. (2020). Same-day delivery with fairness. ArXiv preprint arXiv:2007.09541.

Company, M. a. (2016). Parcel delivery the future of last mile. https://www.mckinsey.com/~/media/mckinsey/industries/travel%20transport%20and%20logistics/our%20insights/how%20customer%20demands%20are%20reshaping%20last%20mile%20delivery/parcel_delivery_the_future_of_last_mile.ashx (accessed 06/01/2022).

Côté, J.-F., de Queiroz, T. A., Gallesi, F., & Iori, M. (2021). Dynamic optimization algorithms for same-day delivery problems. In Bureau de montreal, université de montreal. https://www.semanticscholar.org/paper/Dynamic-Optimization-Algorithms-for-Same-Day-C%C3%B4t%C3%A9-Queiroz/5e8b03a60f3f033f6dea533ef7d9f1ff5ccfabce#citing-papers.

Dumouchelle, J., Frejinger, E., & Lodi, A. (2021). Can machine learning help in solving cargo capacity management booking control problems? ArXiv preprint arXiv:2102.00092.

eMarketer (2022). Retail e-commerce sales worldwide from 2014 to 2025 (in billion u.s. dollars). https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/ (accessed 06/01/2022).

Fleckenstein, D., Klein, R., & Steinhardt, C. (2022). Recent Advances in Integrating Demand Management and Vehicle Routing: A Methodological Review. European Journal of Operational Research.

Giallombardo, G., Guerriero, F., & Miglionico, G. (2020). Profit maximization via capacity control for distribution logistics problems. ArXiv preprint arXiv:2008.03216.

Hildebrandt, F. D., Thomas, B., & Ulmer, M. W. (2021). Where the action is: Let's make reinforcement learning for stochastic dynamic vehicle routing problems work! ArXiv preprint arXiv:2103.00507.

Klapp, M. A., Erera, A. L., & Toriello, A. (2018). The dynamic dispatch waves problem for same-day delivery. *European Journal of Operational Research, 271*(2), 519–534.

Klapp, M. A., Erera, A. L., & Toriello, A. (2020). Request acceptance in same-day delivery. *Transportation Research Part E: Logistics and Transportation Review, 143*, 102083.

Klein, R., Koch, S., Steinhardt, C., & Strauss, A. K. (2020). A review of revenue management: Recent generalizations and advances in industry applications. *European Journal of Operational Research, 284*(2), 397–412.

Klein, R., Mackert, J., Neugebauer, M., & Steinhardt, C. (2018). A model-based approximation of opportunity cost for dynamic pricing in attended home delivery. *OR Spectrum, 40*(4), 969–996.

Koch, S., & Klein, R. (2020). Route-based approximate dynamic programming for dynamic pricing in attended home delivery. *European Journal of Operational Research, 287*(2), 633–652.

Lang, M. A., Cleophas, C., & Ehmke, J. F. (2021a). Anticipative dynamic slotting for attended home deliveries. *Operations Research Forum, 2*, 1–39.

Lang, M. A., Cleophas, C., & Ehmke, J. F. (2021b). Multi-criteria decision making in dynamic slotting for attended home deliveries. *Omega, 102*, 102305.

Lebedev, D., Goulart, P., & Margellos, K. (2021). A dynamic programming framework for optimal delivery time slot pricing. *European Journal of Operational Research, 292*(2), 456–468.

Lebedev, D., Margellos, K., & Goulart, P. (2020). Approximate dynamic programming for delivery time slot pricing: a sensitivity analysis. ArXiv preprint arXiv:2008.00780.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. John Wiley & Sons.

Mackert, J. (2019). Choice-based dynamic time slot management in attended home delivery. *Computers & Industrial Engineering, 129*, 333–345.

Powell, W. B., Simao, H. P., & Bouzaiene-Ayari, B. (2012). Approximate dynamic programming in transportation and logistics: A unified framework. *EURO Journal on Transportation and Logistics, 1*(3), 237–284.

Prokhorchuk, A., Dauwels, J., & Jaillet, P. (2019). Stochastic dynamic pricing for same-day delivery routing. ArXiv preprint arXiv:1912.02946.

PwC (2018). Global consumer insights survey 2018. https://www.pwc.com/gx/en/retail-consumer/assets/delivery-expectations-global-consumer-insights-survey.pdf (accessed 06/01/2022).

Snoeck, A., Merchan, D., & Winkenbach, M. (2020). Revenue management in last-mile delivery: state-of-the-art and future research directions. *Transportation Research Procedia, 46*, 109–116.

Soeffker, N., Ulmer, M. W., & Mattfeld, D. (2017). On fairness aspects of customer acceptance mechanisms in dynamic vehicle routing. *Proceedings of logistikmanagement, 2017*, 17–24.

Soeffker, N., Ulmer, M. W., & Mattfeld, D. C. (2021). Stochastic dynamic vehicle routing in the light of prescriptive analytics: A review. *European Journal of Operational Research*.

Strauss, A. K., Gülpınar, N., & Zheng, Y. (2021). Dynamic pricing of flexible time slots for attended home delivery. *European Journal of Operational Research, 294*(3), 1022–1041.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Talluri, K. T., & Van Ryzin, G. J. (2006). *The theory and practice of revenue management*: 68. Springer Science & Business Media.

Toth, P., & Vigo, D. (2014). *Vehicle routing: Problems, methods, and applications*. SIAM.

Ulmer, M. W. (2017). *Approximate dynamic programming for dynamic vehicle routing*. Springer.

Ulmer, M. W. (2019). Anticipation versus reactive reoptimization for dynamic vehicle routing with stochastic requests. *Networks, 73*(3), 277–291.

Ulmer, M. W. (2020a). Dynamic pricing and routing for same-day delivery. *Transportation Science, 54*(4), 1016–1033.

Ulmer, M. W. (2020b). Horizontal combinations of online and offline approximate dynamic programming for stochastic dynamic vehicle routing. *Central European Journal of Operations Research, 28*(1), 279–308.

Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., & Hennig, M. (2019). Offline–online approximate dynamic programming for dynamic vehicle routing with stochastic requests. *Transportation Science, 53*(1), 185–202.

Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., & Thomas, B. W. (2020). On modeling stochastic dynamic vehicle routing problems. *EURO Journal on Transportation and Logistics*, 100008.

Ulmer, M. W., Mattfeld, D. C., & Köster, F. (2018). Budgeting time for dynamic vehicle routing with stochastic customer requests. *Transportation Science, 52*(1), 20–37.

Ulmer, M. W., & Thomas, B. W. (2020). Meso-parametric value function approximation for dynamic customer acceptances in delivery routing. *European Journal of Operational Research, 285*(1), 183–195.

Vinsensius, A., Wang, Y., Chew, E. P., & Lee, L. H. (2020). Dynamic incentive mechanism for delivery slot management in e-commerce attended home delivery. *Transportation Science, 54*(3), 567–587.

Voccia, S. A., Campbell, A. M., & Thomas, B. W. (2019). The same-day delivery problem for online purchases. *Transportation Science, 53*(1), 167–184.

Waßmuth, K., Köhler, C., Agatz, N., & Fleischmann, M. (2022). Demand management for attended home delivery–a literature review. In ERIM report series reference forthcoming. https://madoc.bib.uni-mannheim.de/62840/.

Yang, X., & Strauss, A. K. (2017). An approximate dynamic programming approach to attended home delivery management. *European Journal of Operational Research, 263*(3), 935–945.

Yang, X., Strauss, A. K., Currie, C. S., & Eglese, R. (2016). Choice-based demand management and vehicle routing in e-fulfillment. *Transportation science, 50*(2), 473–488.