Omega 120 (2023) 102910

Contents lists available at ScienceDirect

Omega

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

Locker box location planning under uncertainty in demand and capacity availability☆

Simona Mancini[a,b,*], Margaretha Gansterer[b], Chefi Triki[c]

[a] University of Palermo, Department of Engineering, Palermo Italy
[b] University of Klagenfurt, Department of Operations, Energy, and Environmental Management, Universitätsstraße 65-67, 9020 Klagenfurt, Austria
[c] University of Kent, Kent Business School, Canterbury, UK

ARTICLE INFO

Article history:
Received 16 November 2022
Accepted 27 May 2023
Available online 2 June 2023

Keywords:
Last-mile delivery
Facility location
Capacity availability
Stochastic
Matheuristics

ABSTRACT

In this paper, we address the location of locker boxes in the last-mile delivery context under uncertainty in demand and capacity. The problem is modeled as an extension of the capacitated facility location problem, in which a fixed number of facilities has to be opened, choosing among a set of potential locations. Facilities are characterized by a homogeneous capacity, but a capacity reduction may occur with a given probability. The uncertainty in demand and capacity is incorporated through a set of discrete scenarios. Each customer can be assigned only to compatible facilities, i.e., to facilities located within a given radius from the individual location. The goal is to first maximize the total number of customers assigned to locker boxes, while, in case of a tie on this primary objective, a secondary objective intervenes aiming at minimizing the average distance covered by customers to reach their assigned locker box. A stochastic mathematical model as well as three matheuristics are presented. We provide an extensive computational study in order to analyze the impact of different parameters on the complexity of the problem. The importance of considering uncertainty in input data is discussed through the usage of general stochastic indicators from the literature as well as of problem specific indicators. A real-world case related to the City of Turin in Italy is analyzed in detail. The benefit achievable by optimizing locker box locations is discussed and a comparison with the current configuration is provided.

1. Introduction and motivation

E-commerce has experienced a considerable growth during the last decade, which has been further accelerated by the pandemic situation that caused long lockdown periods for physical shops.

Online shopping and last-mile delivery of parcels is becoming a huge business that will continue to grow worldwide by 10% per year during the coming decade [1–3]. The numbers concerning last-mile distribution are astonishing. For example, *Amazon* has delivered 3.5 billion parcels in 2019 and is expected to deliver 6.5 billion by 2022 [4]. In order to face such a record growth, practitioners and researchers are continuously trying to develop powerful optimization tools with the aim of reducing delivery costs and increasing the distribution efficiency. In this context, one of the most studied problems is that of defining least-cost distribution plans (see, for example, [5–9]).

The huge number of requests needed to be fulfilled every day makes last-mile delivery a very critical issue for logistics companies. Barenji et al. [10] reported that, in a last-mile delivery context, distribution can cost up to 40% of the price of a product, therefore, it is of crucial importance to efficiently plan and organize delivery operations.

Currently, most of the logistic providers do not allow customers to select a preferred delivery time slot, but just communicate the day on which the delivery will be performed. This system, according to data reported by Morganti et al. [11], can result in up to 40% of missed deliveries, with a consequent huge impact on delivery costs for the companies, who will have to reschedule the visit to the unserved customers on the following days. To overcome this relevant issue, alternative delivery systems have been proposed, such as time window pricing techniques [12–14] or personalized time slot incentives [15–18]. However, although these de-

☆ Area: Production Management, Scheduling and Logistics. This manuscript was processed by Associate Editor Otto.

* Corresponding author.
E-mail addresses: simona.mancini@unipa.it, simona.mancini@aau.at (S. Mancini), margaretha.gansterer@aau.at (M. Gansterer), c.triki@kent.ac.uk (C. Triki).

https://doi.org/10.1016/j.omega.2023.102910

livery systems actually help to reduce the distribution costs, many operational challenges persist, which make them hardly viable for companies handling a huge number of requests per day.

A new last-mile delivery concept involves unattended delivery to shared locations, named locker box stations. These facilities are generally located in widely accessible sites, such as supermarkets, refueling stations, train stations, etc. Each facility consists of several independent locker boxes and a terminal that manages the system. Parcels belonging to the same customer can be stored in the same locker box, if they fit, whereas parcels of different customers cannot be packed together. Once the delivery to the locker box has been performed, the customer receives a notification about the successful delivery and a pick up code.

Typically, customers have few days to pick up their parcel, after which it is returned by the logistics provider to a distribution center, but the advantage is that they have flexible pick up times. Not all orders are suitable for this delivery option, due to the limitations imposed by the parcel size or value, or by customers' willingness. Obviously, customers will not accept this delivery option if the locker boxes are not conveniently accessible for them [7]. Customers who reject the use of locker boxes must be served by the typical home delivery, which is known to be very costly. Hence, location planning is the key to achieve success in this delivery model.

In our study, we focus on the strategic/tactical decision level of locating locker stations. Customers' locations are assumed to be known in advance, whereas the specific demand to be served varies from one day to another. In addition, available capacity is considered uncertain as a locker box can be temporarily unavailable, due to customers not picking up their parcels on time. Since we are addressing a strategic/tactical decision problem, we assume that the exact location of customers per day are not known. Hence, we cannot quantify the routing part, but we know that each customer, who cannot be served by locker stations, must be served by home delivery, and that each additional home delivery to perform, negatively impacts routing costs. Therefore, our goal is to maximize the number of customers serviceable by lockers stations.

Also, technical problems or vandalism can represent further reasons for temporal unavailability. In [19], the authors report that only 70% of customers pick up their parcels within the first 24 hours after the delivery. They also estimate that, on average, 8% of the parcels cannot be delivered because of locker box occupancy. This percentage further increases if we consider reverse flows, in which customers use locker boxes for returning items to their sellers.

We introduce, thus, the facility location problem (FLP) with uncertain demand and uncertain capacity availability (FLP-UDUC). To the best of our knowledge, this problem has never been addressed before in the literature. The main contributions of this paper can be summarized as follows:

1. We formally introduce the FLP-UDUC proposing an Integer Programming formulation and a hierarchical objective function.
2. We design an efficient and effective matheuristic to address large-sized instances.
3. We define a consensus search-based matheuristic that can be generalized to a whole class of two-stages stochastic problems.
4. We present a detailed analysis of the Expected Value of Perfect Information (EVPI) and of the Value of the Stochatic Solution (VSS), two stochastic indicators that are commonly used to determine the importance of considering the uncertainty in the problem under exam.
5. We apply the proposed method to randomly generated instances as well as to a real-world case in the City of Turin, Italy. Potential benefits and further managerial insights are discussed.

The paper is organized as follows. Section 2 presents an analysis of the related literature review. The problem description and the mathematical formulation are reported in Section 3. Solution approaches are presented in Section 4, whereas computational results are discussed and analyzed in Section 5. The consideration of probability of locker box availability depending on utilization rate is discussed in Section 5.4. Section 5.5 is devoted to the real-world case. Finally, conclusions and future developments are reported in Section 6.

2. Literature review

Last-mile logistics providers try to improve their efficiency and to increase their market share with respect to their competitors through expanding their locker stations network and optimizing their configuration and location [20,21]. For example, *Amazon* was committed to install up to 1,000 new digital locker stations in the United States every month along the last few years [22]. Moreover, the leading delivery company *InPost* declared it has already installed more than 3,000 locker stations in the UK and is planning to increase this number to reach 10,000 stations by 2024 [23]. Moreover, the company intends to expand its network to up to 1 million stations worldwide [24]. Similarly, the *DPDHL* group is aiming to install at least 12,500 locker stations in Germany by the end of 2023 [25].

A common challenge faced by all last-mile providers is deciding the location of the locker units. Several studies and surveys highlighted the importance of identifying suitable locker locations to ensure their attractiveness to a wide range of customers [26–31]. Such a problem can be solved following a 2-phase approach: first, a set of potential sites has to be identified and then, an appropriate subset of locations can be selected among the candidate ones. Investigating the first phase is beyond the scope of this study. However, interested readers can be referred, for example, to Lagorio and Pinto [32] or to Faugere and Montreuil [33] who reviewed and analyzed several business models and real-life experiences and identified the most important factors influencing the selection of potential locker station sites (such as availability, accessibility, safety, environmental impact, and costs).

The second phase consists of solving the FLP, which is by far the most popular optimization problem, where a subset of locations has to be selected among the set of potential sites [34,35].

2.1. Deterministic locker location problems

There are several papers addressing the deterministic version of the locker station location problem in the context of last-mile delivery. Wang et al. [36] were the first to consider the viewpoint of a new delivery provider entering a competitive market. The authors develop an optimization model based on a maximal coverage location formulation for locating its p new lockers. They make use of public big data and apply the suggested model to a real-life problem in Singapore.

Deutsch and Golany [37] develop a binary linear model that determines the location and size of locker stations. Their profit-maximization objective function involves even a discount term proposed to incentivize customers to accept the locker box delivery.

Lee et al. [38] identify a set of candidate sites to install the locker stations in the city of Incheon in South Korea based on the concepts of neighbourhood, accessibility, and availability of public facilities. Then they combine the GIS technology, the set covering, and the p-median models within an integrated optimization framework to optimally locate the lockers.

Schwerdfeger and Boysen [39] deal with the variant of dynamic locker stations that can change location over the day. The authors develop three mixed integer models that minimize the number of locker stations while satisfying the customers demand. They develop specialized exact approaches and test them using randomly generated instances.

Lin et al. [40] study the problem of designing a new locker station network and develop an exact approach based on a mixed-integer linear model strengthened by the conditional McCormick inequalities. Moreover, they develop a suggest-and-improve approach to solve large-scale instances. They test both the exact and heuristic methods on randomly generated instances and then on a real-life case related to a pop-locker alliance in Singapore.

Yang et al. [41] solve the problem with specialized modelling approaches based on the bilevel programming paradigm. The upper-level model is devoted for solving the location problem, whereas the lower-level allows each customer to assign the demand to the locker station, minimizing the pick up cost. The authors develop a genetic algorithm approach, embedded with the GIS technique, to solve the problem.

Besides the above-mentioned papers, there are some studies that are predominantly oriented towards real-life applications such as, for example, Simić et al.[42] and Zheng et al. [43]. In addition, few other works combine the location problem with related aspects of the locker station network design. More specifically, Oliveira and dos Santos [44] combine the problem of locating the locker boxes with that of defining a multi-shift routing plan. The authors propose an integer model and develop a Variable Neighborhood Descent-based heuristic to solve the problem. They also test their approach on known instances appropriately adapted to fit their context (a very similar integrated problem is discussed in Veenstra et al. [45], but this latter arises in the field of health care logistics). Likewise, [46] integrate the facility location with the problem of assigning the customers to the locker boxes and apply their model to the locker station network design of the city of Pamplona in Spain. The model, that minimizes both the location and the assignment costs, involves even the cost of decommissioning some of the existing locker boxes.

2.2. Stochastic locker location problems

The lockers location problem is made challenging when uncertainty related to the problems parameters is incorporated into the optimization model. A review on the general facility location model under uncertain data can be found in Snyder [47], and a more updated survey has been recently proposed by Suryawanshi and Dutta [48]. Most of the works available in the literature focus on considering the demand as an uncertain parameter [49,50]. This claim is made even more evident in the context of locating locker units.

More specifically, [51] deal with the problem of selecting the location of movable parcel locker units under stochastic demands and propose a robust optimization model that minimizes the total operating cost. Afterwards, the authors transform the robust formulation into its integer program counterpart and use standard commercial software packages to solve the resulting deterministic equivalent problem.

Rabe et al. [52] and Rabe et al. [53] focus on the multi-period variant of the locker station location problem and develop stochastic simulation-optimization methods for its solution. The stochastic demand is represented through a small set of distinct scenarios. In their approach, the number and location of locker units is defined through exact models whereas the reliability and cost corresponding to each scenario is simulated by the Monte Carlo method. The experimental results in Rabe et al. [53] discuss a real-life case related to the city of Dortmund in Germany.

Kahr [54] deal with locating multi-compartment lockers in the city of Vienna under a discrete stochastic demand representation. The location problem is formulated as an integer program on the basis of which a Benders Decomposition approach is developed. The problem's objective is to maximize the expected utility derived from serving the demand. The problem is constrained by a maximum number of lockers to be installed due to budget restrictions.

Unlike all the above-mentioned works, in this paper, we deal not only with the uncertainty in customers demand but also in the locker box availability. To the best of our knowledge, this is the first study that considers both these features simultaneously.

2.3. Facility location problems with capacity unavailabilities

For the sake of completeness, it is worth summarizing papers covering stochastic FLP with capacity unavailabilities due to disruptions. However, none of these studies is related to the employment of lockers for last-mile delivery. Most of these articles arise either in the context of military applications [55], emergency facility location for disaster preparedness [56], or in resilient supply chain management [57]. The disruption is modelled in these cases through the probability of facilities availability, i.e., the facility can be fully operative with a probability p or is completely unavailable with probability $1-p$.

Partial facility disruption, in which a facility can result to be still available but with a lower capacity, has also been investigated. Rohaninejad et al. [58] consider a probability of full or partial facility capacity failure in a multi-echelon network. In their problem, the authors allow to increase facility capacity, with an additional cost, in order to mitigate the effect of partial capacity failure, while in our case, capacity cannot be modified. Florez et al. [59] study a robust humanitarian facility location in which the capacity of each facility is uncertain and may vary between a minimum and a maximum value. The most related paper to our study is [60], who consider that one or more facilities can be partially operative and apply their approach to facility fortification problems under uncertainties. However, there are some substantial differences with respect to our study. First, [60] formulate the problem as a robust optimization approach rather than a stochastic model. Also, unlike our work, their goal is to minimize the total costs, given by the sum of facilities to be opened, customers-facilities assignment, and the cost of not satisfying part of the demand. Unknown facility capacity availability has been addressed also in Ulmer and Streng [61], where the authors study the problem of same-day delivery with pick-up stations. They address a decision problem arising at the operational level. A set of customer requests dynamically arrives and the company has to decide whether to accept or reject the request and to which pick-up station to deliver the order. The time incurring between the delivery of the order to the station and the pick-up of the order by the customer, is considered stochastic and it impacts the available capacity at the station. Although this problem shows some similarity to ours, the two problems are clearly different. Our problem arises on the strategic/tactical level and involves both facility location and assignment decisions, while the problem tackled in Ulmer and Streng [61] is on the operational level, in a dynamic setting, and only deals with the assignment decision. Capacity uncertainty is also experienced in different humanitarian logistics. In these cases, the capacity of a facility is the quantity of resource it can supply to the population. This quantity may be not known in advance, such as in the case of blood and medicines donations, since it depends on donors availability which cannot be controlled by the organization, being based on a voluntary action. However, these problems are very different respect to ours, in which instead, the maximum capacity is known, but a capac-

Table 1
Summary of the related contributions.

	STOCHAST.		LOCKERS	MODELING	SOLUTION	
ARTICLE	DEM.	CAP.	LOCATION	APPROACH	APPROACH	CASE
[36]			√	max. cov.	realloc. constr.	Singapore
[37]			√	IP	exact	Toronto
[38]			√	set cov.		
				p-median	exact	Incheon
[39]			√	MILP	exact	
[40]			√	MILP	exact/heuristic	Singapore
[41]			√	bilevel progr.	Genetic Algorithms	Changsha
[44]			√	IP	VNS	
[45]			√	MILP	exact/heuristic	Netherlands
[46]			√	MILP	simul.-opt.	Pamplona
[51]	√		√	IP	robust opt.	
[52]	√		√	IP	simul.-opt.	Dortmund
[53]	√		√	IP	simul.-opt.	Dortmund
[54]	√		√	IP	Benders Decomp.	Vienna
[58]	√	√		MILP	Benders Decomp.	
					samp. avg. approx.	
[59]	√	√		IP	stoch. multi-scen.	Peru
[60]	√	√		robust opt.	column-constr. gen.	
This article	√	√	√	IP	matheur.	Turin

ity reduction can occur. For this reason, we do not explicitly cover this topic in our literature review but we refer interested readers to Dönmez et al. [62].

On the basis of the above analysis and the overview of contributions that we report in Table 1, it becomes obvious that this is the first study dealing with the stochastic FLP while incorporating the uncertainty not only in customer demands, but also in available capacity. We believe that our study closes a relevant gap and brings research on locker stations closer to real-life settings.

3. Problem definition and Mathematical Formulation

3.1. Problem Definition

The problem addressed in this work is a strategic/tactical problem, where the goal is to determine the optimal location for a fixed number of homogeneous lockers containing the same number of locker boxes. The decision is based on forecasts of future customers' demand and uncertain capacity availability. Customers can be accepted, and assigned to one of the lockers within a maximum radius, or can be rejected. The overall goal is to first maximize the number of customers accepted, and then, in case of a tie on this primary objective, to minimize the average travel distance between a customer and the locker to which it has been assigned. Available capacity is considered uncertain as a locker box can be temporarily unavailable, due to, e.g., customers not picking up their parcels on time or due to technical problems or vandalism acts. To represent the two sources of uncertainty we consider a set of demand scenarios and a set of capacity scenarios. We evaluate all the possible combinations. Each customer belongs to one and only one demand scenario, which means that demand scenarios are disjoint. All the customers are evaluated in all the capacity scenarios. For each demand and capacity scenario combination, we have to assign customers to lockers, trying to maximize the number of customers served, while respecting the available capacity. The novel aspect of this problem, with respect to a classical stochastic facility location, is twofold. Firstly, we consider two sources of uncertainty: customers' demand and facilities' available capacity, while in the classical problem only uncertainty on customers' demand is handled. Secondly, we use a hierarchical function, in which the primary goal is to maximize the served customers across all the scenarios, while the secondary one aims at minimizing the average travel distance between a customer and the facility to which it has been assigned. The classical problem, instead, deals with a single objective.

In Figure 1 we depict, for the same customer scenario, the customer-assignment in two different capacity scenarios (*a* and *b*). The instance contains 22 customers, represented by small blue circles, and 5 locker locations, (*A, B, C, D,* and *E*). Among those 5 locations (facilities), 3 have to be selected. Open facilities are depicted in green, whereas closed ones are depicted in red. Near each facility we report the available capacity in the specific scenario. In the first scenario, available capacity allows to assign all the customers to their nearest facility. Instead, in the second one, a capacity reduction occur on facility E. Hence capacity is reduced from 10 to 8. Since the numbers of customers, for which E was the nearest facility, is 10, two of them have been allocated to D. The two customers which have been reallocated are those with the lowest difference of distance from D and from E. This solution implies the minimum increment of averaged traveled distance, and therefore it is preferable with respect to the others, according to the secondary objective. In fact, the primary objective is not affected by any allocation change, but only by changes in the acceptance/rejection decisions. Facility A also experiences a capacity reduction from 9 to 6. This means that two of the 8 customers previously assigned to it, need to be reallocated. Unfortunately, none of those customers can be assigned to another open facility, since both D and E are too far from them. Therefore, two customers must be rejected in scenario b. The rejected customers are the two farthest from A. This, in fact, reduces the average traveled distance for customers. Any other solution, in which another pair of customers, previously assigned to A, are rejected, would show the same primary objective value, but a worse value for the secondary objective, and therefore, it would be suboptimal. Rejected customers are depicted in the figure with a dark-red circle. Note that, while in the depicted example the optimal solution can be reached with a very small and straight reallocation, in other cases complex reassignment chains could be needed, potentially involving all the open facilities.

In the following we provide a formal description of the problem and introduce the notation used in the mathematical formulation. The problem aims at determining the best location for a fixed number of locker stations, P, chosen among a set of candidate locations J. Each locker station is composed of C locker boxes. A set of demand scenarios S is considered. In each scenario s, where $s \in S$, a set of potential customers I^s out of the set of all customers I has to be served. All the demand scenarios are disjoint, meaning that each customer belongs to one and only one scenario. Note that this

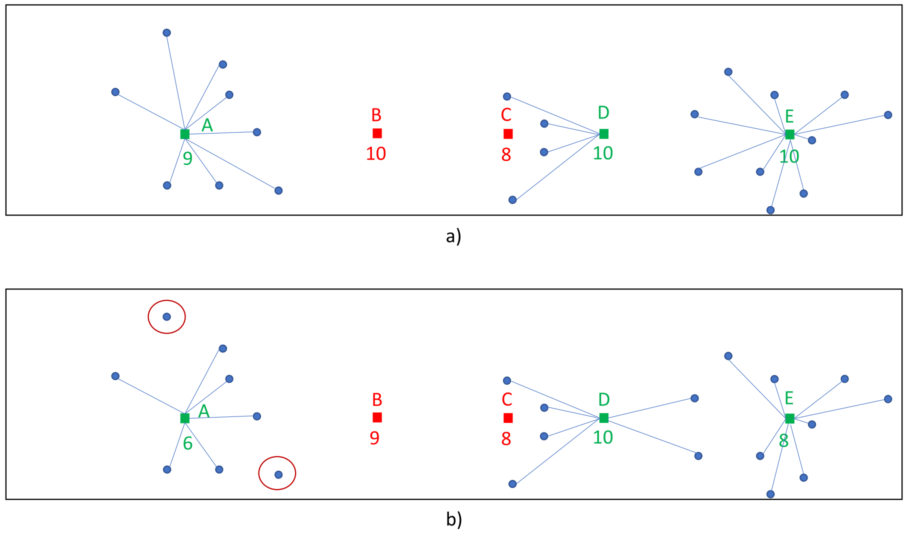

Fig. 1. An illustrative example of customer assignments in two different capacity scenarios.

does not prevent having customers sharing exactly the same location. It only means that scenarios are uncorrelated among each other, i.e., we do not have customers simultaneously belonging to several scenarios. Consequently, allocation decisions in one scenario do not impact allocation decisions in other scenarios, while location decisions impact all the allocation decisions in the different scenarios. A customer may be accepted or rejected. If accepted, it must be assigned to a compatible locker station, j, where $j \in J$. Only locker stations located within a radius ρ from the customer's delivery address are considered as compatible. We indicate with ϕ_{ij} the compatibility between customer i ($i \in I$) and a locker station j (ϕ_{ij}=1 indicates that they are compatible and ϕ_{ij}=0 that they are not).

The number of customers assigned to a locker station, within the same demand scenario, cannot exceed its capacity. A set of capacity reduction scenarios Ω ($\omega \in \Omega$) is defined. In each scenario, the capacity of each locker station is reduced by a quantity δ_j^{ω} that represents a temporal unavailability of the capacity of a set of locker boxes. The primary objective of the problem is to determine the locker station locations which maximize the average number of customers served by locker box delivery, over all the demand and capacity availability scenarios. A secondary objective, which intervenes only in case of a tie, aims at minimizing the average travel distance from customers' locations to the locker stations to which the customers have been assigned.

3.2. Mathematical Formulation

We introduce the following sets of decision variables exploited in the mathematical model.

- Y_{ij}^{ω}: binary variable indicating whether customer i is assigned to facility j in scenario ω or not
- Z_j: binary variable indicating whether a locker station is placed in location j or not

The problem is modeled as a two-stage stochastic model, where the Z_j are the first-stage decision variables while the Y_{ij}^{ω} are associated to second-stage decisions, that depend on the specific scenario that materialized. The mathematical model can be formulated as follows.

$$\max \sum_{\omega\in\Omega}\sum_{i\in I}\sum_{j\in J} Y_{ij}^{\omega} + \sum_{\omega\in\Omega}\sum_{i\in I}\sum_{j\in J} \frac{d_{min}}{d_{ij}} Y_{ij}^{\omega} \frac{1}{|I||\Omega|} \tag{1}$$

$$\sum_{i\in I_s} Y_{ij}^{\omega} \le C - \delta_j^{\omega} \quad \forall \omega \in \Omega \;\; \forall s \in S \;\; \forall j \in J \tag{2}$$

$$\sum_{j\in J} Y_{ij}^{\omega} \le 1 \quad \forall i \in I \;\; \forall \omega \in \Omega \tag{3}$$

$$Y_{ij}^{\omega} = 0 \quad \forall \omega \in \Omega \;\; \forall (i \in I, j \in J | \phi_{ij} = 0) \tag{4}$$

$$\sum_{\omega\in\Omega}\sum_{i\in I} Y_{ij}^{\omega} \le |I||\Omega| Z_j \quad \forall j \in J \tag{5}$$

$$\sum_{j\in J} Z_j = P \tag{6}$$

$$Z_j \in \{0, 1\} \quad \forall j \in J \tag{7}$$

$$Y_{ij}^{\omega} \in \{0, 1\} \quad \forall i \in I \;\; \forall j \in J \;\; \forall \omega \in \Omega \tag{8}$$

The hierarchical objective function is reported in (1). It primarily aims at maximizing the number of customers assigned to locker stations, and secondly, to minimize the average distance covered by a customer to pick up her parcel. The secondary objective is formulated such that it can assume only values between 0 and 1. In fact, for each served customer, i, we compute the ratio between the potential minimum distance (d_{min}) between the customer and any locker station and the actual distance (d_{ij}) covered by customer i to reach locker j, i.e. the one she was assigned to. Note that, in case of $d_{min} = 0$, we automatically set d_{min} to be 50 meters. Otherwise, the secondary objective would be equal to 0 for all solutions. For a similar reason, very small values of d_{min} (i.e., $\le$ 50 meters), are rounded to 50 meters. Given the above assumption, the ratio between d_{min} and d_{ij} can only take values between 0 and 1. It can be 1 only if d_{ij} and d_{min} coincide, and it can never be equal to 0. We also exclude here the special case in which $d_{ij} = 0$, that is locker station j is installed in the same location as customer i). Since each ratio is lower or equal to 1, and each customer can be assigned to at most one locker station in each capacity scenario ω, the sum of the ratio in the secondary objective cannot be larger than the number of customers multiplied by the number of capacity scenarios, $|I||\Omega|$. Consequently, the second term of the objective function is always between 0 and 1, while the primary objective is always an integer number. Thus, the secondary objective only intervenes in case of a tie on the primary objective. In other words, it allows us to further discern among the set of solutions that are optimal for the primary objective, selecting the one (or the subset of them) with the highest value for the secondary objective.

In our specific case, the primary goal is company-oriented. In fact, it aims at maximizing the number of customers serviced by locker box delivery. The secondary goal, however, is customer-oriented, since it aims at minimizing the average travel distance. Such hierarchical objective functions are particularly useful for multi-objective optimization problems in which a ranking of the objectives' priority is known [63,64]. The role of the secondary objective is to help us to further distinguish among solutions having the same primary objective value. This is particularly useful in problems with a primary objective covering only discrete values, which might lead to a large amount of equal solutions. One can argue that the importance of the second objective is limited, since the average travel distance will be small due to compatibility matrix anyway. However, if we compare two solutions serving exactly the same number of customers, we should prefer the one with the shorter travel distance, even if the difference is rather small. Note that our approach is not a classical weighted sum objective, where a large gain on one objective can balance a small loss on the other one. In fact, in our case, solutions having a better value on the primary objective will be always preferred, whichever is the value of the secondary one.

Constraints (2) ensure that locker stations' available capacity is respected in all demand and capacity scenarios. A customer can be assigned to at most one locker station in each capacity scenario, as expressed in constraints (3). Furthermore, customers can be assigned only to compatible locations, as ensured by constraints (4), only among those in which a locker station has been installed (constraints (5)). The compatibility among customers and lockers is computed in a preprocessing phase, in which we assign value $\Phi_{ij} = 1$ if the distance between customer i and locker j is lower than the maximum allowed, and $\Phi_{ij} = 0$ otherwise. The assignment variables related to an infeasible matching (where $\Phi_{ij} = 0$), are forced to take value 0 due to constraints (4). Finally, the number of locker stations to be installed, which is known to the decision maker in advance to be equal to P, is ensured by constraints (6). All the variables involved in the model are binary. The problem is NP-hard since it can be seen as an extension of the facility location problem which was proven to be NP-hard itself.

4. A matheuristic framework for the FLP-UDUC

As only small instances of the presented problem can be solved to optimality within acceptable amount of time (see Section 5), we propose a new matheuristic framework for solving large and challenging instances. In this framework we initially select a set of P facilities to be added to an initial *core*. Keeping fixed the facilities to open, the decision problem turns into an assignment problem, in which each combination of demand and capacity scenario can be solved separately. The optimal solution for the global problem is then obtained merging the optimal solutions of each single combined scenario, where the current best solution is. Then, all the facilities belonging to the core are marked, while all the others are unmarked. After this preliminary phase, a local search procedure is run. At each iteration, one of the unmarked facilities is added to the core and it is marked as already processed. The resulting restricted optimization problem is solved by running the model with a commercial MIP solver. This subproblem is easier to solve with respect to the original one, since we have to choose to open P facilities out of $P+1$. The non-opened facility is then removed from the core. If the solution obtained so far is better then the current best, then it is kept as current best. In this case, all the facilities belonging to the core are marked and all the others are unmarked. The procedure terminates when all the facilities are marked, i.e. when all the possible single insertions in the core have been tested without obtaining any improvement. This means that no further improvements can be achieved with the local search procedure. A flowchart depicting the procedure is reported in Figure 2.

The performance of the algorithm strictly depends on two key algorithmic decisions: 1) How to select the initial core and 2) which criteria to use for choosing the next facility to be processed (i.e. to be added to the core). We designed a version of this algorithm, named Consensus Search (CS), which exploits ad hoc strategies, for both 1) and 2), specifically tailored for this problem, and compare it with two versions in which more classical strategies are applied. Since these standard strategies yield to a premature convergence toward local minima, we embed them in a metaheuristic framework, equipped with diversification mechanisms, such as Iterated Local Search (ILS) and Variable Neighborhood Search (VNS). Note that this diversification is not necessary if the newly proposed framework is used. The three different version of the algorithms (CS, ILS, and VNS) are described in detail in the following.

4.1. The consensus search

In this section, we describe a new framework based on the idea of searching for consensus among scenarios. For the sake of consistency with existing literature, we denote a potential location of a locker station as *facility*. Each facility is considered open if a locker is installed therein, and closed otherwise. For the sake of clarity, we first introduce three key concepts that are the basis for our newly proposed method.

- *Score:* defined as the total number of scenarios in which facility j is open, denoted as p_j.
- *Interchangeability*: represents a measure of the proximity of facility j to the nearest other facility in any specific scenario ω, denoted as γ_j^ω.
- *Attractiveness*: is a measure of the importance of opening facility j in scenario ω, that we denote as α_j^ω.

The algorithm starts by solving each capacity scenario ω separately to obtain an *ideal* set of facilities to open for specific scenario (F_ω). We introduce a parameter σ_j^ω, which is equal to 1 if j belongs to F_ω and 0 otherwise. Combining all the ideal sets may yield an infeasible global solution since more facilities than the maximum allowed number, have to be opened. If this is not the case, we already have a global consensus and the solution, in terms of set of locations selected, is optimal. Otherwise, we search for consensus among scenarios.

Besides computing the score p_j for each facility j, we also calculate the *isolation degree*, which is a normalized parameter taking a value between 0 and 1. The higher the shorter distance between j and any other facility, the greater the value of isolation. This is due to the fact that solutions with a high degree of isolation are most difficult to replace and therefore, they are more likely to stay in the global optimal solution.

Additionally, we use a normalized value of the interchangeability for each facility in each scenario ω. The interchangeability parameter γ_j^ω can assume values between 0 and 1. If j belongs to F_ω, γ_j^ω is set equal to 0, otherwise it is fixed equal to $d_{min}/\tilde{d}_j^\omega$, where $\tilde{d}_j^\omega = min_{l \in J} d_{jl}$. We also define the attractiveness of facility j in scenario ω as $\alpha_j^\omega = \sigma_j^\omega + \gamma_j^\omega$. We then create a set of candidate facilities Λ^ω, which contains all the facilities ordered by a non-increasing value of α_j^ω.

The role of the interchangeability and attractiveness parameters is crucial in the success of the algorithm. They both help to better rank the alternatives, avoiding ties among two or more of them, which would render the ranking, and consequently the selection of the alternatives, almost random.

To generate a first feasible solution S^0, we create an initial *core* of facilities to open, picking the P facilities with the highest score.

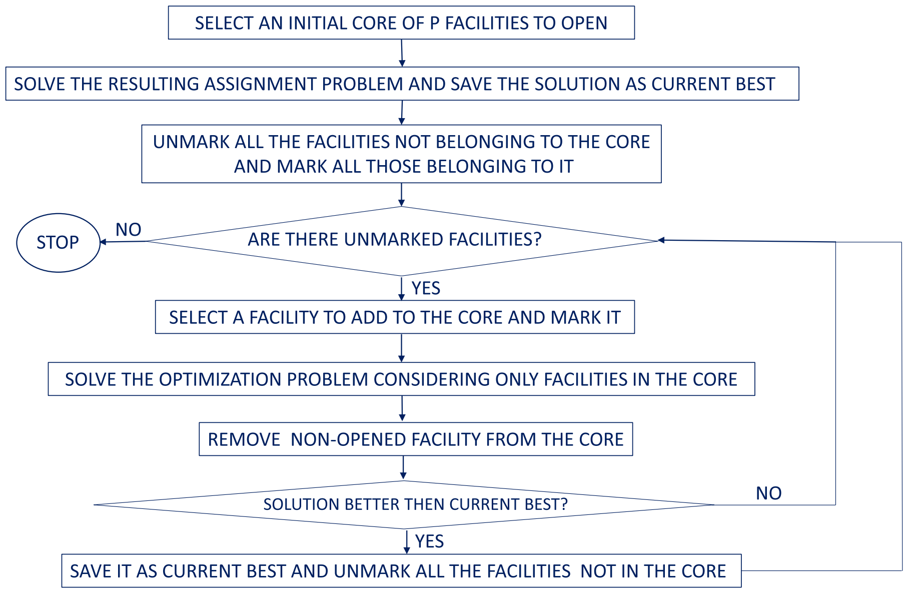

Fig. 2. Flowchart of the matheuristic framework.

We solve the original problem by opening only the facilities in the initial core. The objective function value associated to S^0 is denoted as O^0. The improvement phase then starts and in each iteration, we compute the level of satisfaction for each scenario, H^ω as the sum of customers served in that scenario minus the average percentage increment of distance between a customer location and the facility to which it has been assigned, with respect to the nearest facility. The scenario with the lowest satisfaction, denoted as ω_{worst}, is further investigated. The first facility in the α-based preference list, which has not been marked as *tested* yet, is marked as tested and is added to the core. The problem is solved again with the updated core. Since most P facilities can be used but the core contains $P+1$ facilities, one of them will be discarded by the model. This facility is removed from the core. Every time an improvement is found, all the tested facilities are marked again as untested. The algorithm terminates after a maximum number of iterations (*itermax*) is reached in case all the facilities, not belonging to the current core, are already marked as tested (i.e., if no further improvement, according to our search strategy, is possible). The pseudocode of the CS matheuristic is reported in Algorithm 1.

The concept of consensus searching is not entirely new in the literature, since it has been introduced in Bent and Van Hentenryck [65]. Nevertheless, the way in which this concept is applied here represents one of the contributions of this paper. In [65], the authors propose a general solution framework for dynamic stochastic vehicle routing problem. At each timestep in which a decision must be made, they generate a set of sampling scenarios, representing future requests, optimize each scenario separately and then make the decision which resulted to be the best performing on the largest number of scenarios. The same algorithm has been generalized to all online stochastic problems in Van Hentenryck and Bent [66].

The main drawback of this strategy is that with a high probability, a tie could occur among two or more alternatives, and the method is not able to further discern among them. Furthermore, the method search for a global consensus, without looking at the local consensus. This way, an alternative which is very convenient in the 60% of the sampled scenarios, but very inconvenient for the other 40%, would be always preferred respect to another alternative which is quite convenient in all scenarios. Differently from them, in our consensus based matheuristic we first look for global

Algorithm 1 CS pseudocode

Require: O^0;
Require: set of facilities open in the initial solution (*core*);
Require: for each scenario ω: O_ω (ideal objective function's value for scenario ω);
Require: for each facility j: p_j (score of facility j);
Require: α_j^ω;
  mark all the facilities as untested
  solve the problem opening all the facilities belonging to the core
  $best \leftarrow O^0$
  compute satisfaction of each scenario H^ω
  $\omega_{worst} = argmin_{\omega\in\Omega} H^\omega$
  candidate $\leftarrow$ the first facility within the list of untested facilities ordered by $\alpha_j^{\omega_{worst}}$
  add candidate to the core
  $iter \leftarrow 1$
  while $iter \le itermax$ and at least one facility is marked as untested **do**
    $O^{iter} \leftarrow$ solve the problem allowing to open only facilities belonging to core
    compute the happiness of each scenario H^ω
    $\omega_{worst} = argmin_{\omega\in\Omega} H^\omega$
    candidate $\leftarrow$ the first facility scrolling the list of untested facilities ordered by $\alpha_j^{\omega_{worst}}$
    add candidate to the core
    $iter \leftarrow iter + 1$
    if $O^{iter} \ge best$ **then**
      $best \leftarrow O^{iter}$
      mark all facilities as untested
      remove from the core the facility which has not been opened
    else
      remove candidate from the core
      mark candidate as tested
    end if
  end while
  return *best*

consensus and then, starting from it, we try to increase the local one. Moreover, we introduce two parameters, interchangeability and attractiveness, which helps us ranking alternatives which obtained a tie in the consensus score.

The concept of consensus is studied also in Crainic et al. [67], in which the authors present a progressive-hedging based metaheuristic for stochastic network design. This method starts from an overall design based on all the scenarios and iteratively modify variables selection fixed costs in the objective function in order to push local design to converge to an overall one, achieving a better global consensus. Although this method share some similarity to our, the two methods present strong differences. In [67], the method act on the local consensus, trying to perturb fixed costs in order to derive a stronger global consensus, while our algorithm works exactly in the opposite direction, aiming at modifying the global solutions, exploiting information coming from the local solutions of the scenarios with the lower degree of consensus, also listening to *voices outside the choir*.

4.2. Iterated Local Search

ILS is a very well known metaheuristic framework for combinatorial optimization problems [68]. The underlying idea is that a local search mechanism is run several times starting from different initial solutions. Local search is a powerful tool to explore solutions spaces but its main flaw is that it tends to be trapped in local optima. To overcome this issue, a re-start at another initial solution can be conducted. At the end of the process, a set of local - and possibly the global - optima are available. The selection of the starting solution plays a crucial role. On the one hand, it is important to choose a solution sufficiently far from the current local optimum to allow to avoid being trapped in a local optimum. On the other hand, it should not be too from a region which has been shown to be promising. Therefore, a diversification mechanism used to generate new starting solutions plays a crucial role in the performance of the algorithm. If the local search mechanism is carried out by means of an exact approach, the ILS becomes a matheuristic, as in our case.

In the following, we describe the ILS we use to solve the FLP-UDUC. First, we compute an initial solution following a classical procedure for FLP, where we compute for each potential location the number of customers that can be covered by it. This set of customers is denoted as Cov_j and includes all customers within a compatibility radius j. We then open the P facilities with the highest value of Cov_j and add them to the core. The optimal solution obtained by solving the model presented in Section 3.2 with the fixed set of open facilities is then kept as initial solution S^0 with a corresponding objective function value O^0.

Afterwards, a classical local search operator is used. Here we consider as neighborhood all the solutions that can be obtained by changing only one facility to open. The neighborhood is explored following a first improvement strategy. The procedure works as follows: In each iteration, a candidate facility is selected and added to the current core. Then, the model is solved, allowing to open only facility belonging to the core. If the solution obtained is better than the current best, it is kept as current best, the neighborhood exploration is restarted, and the facility belonging to the core and not opened in the new optimal solution is removed from the core. Otherwise, the candidate facility is removed from the core. Once no further improvements can be achieved, a perturbation is applied, according to which two facilities are randomly removed from the core and substituted with two other randomly selected facilities from the set of currently closed ones. The overall procedure is repeated n_{pert} times.

4.3. Variable Neighborhood Search

VNS is a very broadly used metaheurstic framework for combinatorial optimization problems. After its introduction in 1997 [69], VNS has been successfully applied to a wide range of problems. The core idea of this method consists of a systematic change of neighborhoods within a local search procedure. Although several different versions have been proposed in the literature, the most common practice is to exploit concentric neighborhoods of increasing size. Every time a local optimum is reached, a random solution is generated in the new neighborhood and the local search is restarted. When all the neighborhoods are tested without further improvement, the procedure is terminated. We design a VNS based algorithm specific for the FLP-UDUC. It starts from an initial core computed in the same way as in the case of the ILS presented previously. The corresponding objective function's value is again denoted as O^0. A set of neighborhoods $\mathcal{N}$ with a size of $Nmax$ is defined, where the n^{th} neighborhood consists of changing n elements of the core. Neighborhoods are explored in a size-increasing order following a first improvement strategy. At each step, a solution is randomly selected in the current neighborhood and the local search is applied. Every time a local optimum is reached, the next neighborhood in the list is analyzed. The local search operator is the same as in the ILS.

The CS mathueristics introduces two main aspects of novelty. The first one concerns the construction of the initial core of facilities to open, which is determined by a mechanism aiming at achieving consensus among the different scenarios. Although the idea of consensus search among scenarios is not completely new (see [70]), it so far has been applied by simply counting the number of scenarios in which a variable is active and then using the variables with the highest score. Conversely, we exploit a more complex mechanism to evaluate the score of each facility, which also takes into account the so called *isolation degree*. This yields to fairer consensus decisions since the method does not consider only the most preferred facilities but also try to take into account that facilities within short distance can be easily interchangeable. This way, the method tries to propose valid alternatives for the scenarios for which the preferred facilities have not been included in the initial core, increasing, thus, the global consensus. The second element of novelty is the strategy according to which the local search neighborhood is explored. In fact, CS, identifies the *worst* scenario (i.e., the one which has been penalized the most by the consensus achievement mechanism), and tries to increase its consensus by inserting in the core the most preferred facility for that scenario (among those who are not yet part of the core and who have not been analyzed yet). This exploration strategy aims at converging towards near-optimal solutions more quickly. This is in contrast to classical exploration strategies in which facilities are analyzed in a predetermined sequential order, which does not take into account the level of satisfaction of the scenarios in the current solution. Furthermore, a smart neighborhood exploration strategy, such the one we propose, allows to avoid a premature convergence towards local minima. As it is shown in the computational study (Section 5, the exploration strategy is actually so successful in avoiding local minima, such that CS does not need any diversification mechanism, which, instead, are necessary for both ILS and VNS to avoid premature convergence towards low quality solutions. This makes the CS not only more effective than ILS and VNS but also considerably more efficient. Note that, while we suggest to start from the worst scenario, the proposed framework is flexible enough to allow to pick other scenarios. Concluding, although the CS algorithm seems very easy and simple at first sight, it is smarter than classical approaches. Its apparent simplicity is therefore a strength of the algorithm, since it makes it faster without loosing in accuracy, following the paradigm of *less is more* (see [71] and [72]).

5. Computational study

Our computational experiments analyze several aspects. First, we want to analyze how instances' parameters affect the difficulty of the problem and how they impact the computational times required to solve it to optimality. Second, in order to assess the performance of the newly proposed CS, we compare it against ILS and VNS, and the performance of a solver applied to the exact model. Third, we analyze the importance of considering uncertainty in this problem, rather than solving its deterministic counterpart. This is performed by calculating problem specific indicators as well as two well known uncertainty indicators in this problem: (1) the Expected Value of Perfect Information (EVPI) and (2) the Value of the Stochastic Solution (VSS). Fourth, we conduct experiments regarding unavailability probabilities. Finally, we present experiments based on real-world data from the city of Turin.

For the first analysis, we randomly generate 20 sets of instances (S1-S20), each one composed of 10 instances. The first set, S1, covers (i) 5 demand scenarios with 20 customers each, (ii) 5 capacity scenarios, (iii) 5 facilities (locker stations) with 5 locker boxes each to be selected among 10 candidates, and (iv) an unavailability probability of 10% for each locker box. Sets S2 and S3 are generated based on S1, but with unavailability probability increased to 20% and 30%, respectively. Sets S4 and S5 are also based on S1 but with increased number of potential locations (20 and 30, respectively). S6 and S7 are based on S5. Hence, the number of potential locations is 30, but the number of facilities to open is increased to 10 and 15, respectively. In S8, S9, and S10 we consider the same parameters as in S1, but we increase the number of demand scenarios to 10, 20, and 50, respectively. In S11 and S12, we also increase the number of scenarios to 10 and 20, but we consider a higher number of potential locations (30), such as in S5. In S13 and S14, we again start from S5, but we increase the number of capacity scenarios to 10 and 20. In S15, S16, S17, and S18, we consider an increment of the number of customers per scenarios, keeping the number of scenarios fixed. In S15, S16, and S17, this number is increased to 40, 100, and 200, respectively. However, the capacity of the facilities in terms of number of locker boxes (5 locker boxes each) is kept constant. In S18, we consider 200 customers per demand scenario, but with a larger capacity at the facilities (50 locker boxes). Finally, in S19 and S20, we start from S5 (30 potential locations), and we simultaneously increase the number of customers per scenario and the capacity of locker stations. Namely, we have 100 customers and a capacity of 25 for S19, and 200 customers and a capacity of 50 for S20.

To analyze the impact of each single instance's parameter on the computational times, we keep all the other parameters to very small values, letting vary only the one under study. This allows to exclude mutual interactions among parameters which might yield to incorrect interpretations. Once we can determine the most influencing parameter, we perform further experiments by assuming a large value for this parameter and vary the other parameters one at a time. The aim of this analysis is to evaluate the impact of parameters variation on instances which are already challenging. The computational study (see Section 5) reveals that some parameters that seem to be not influential if analyzed singularly, turn out to have a considerable impact on already challenging instances.

To improve readability of the paper, we resume the instance characteristics for each set in Table 2. All data including detailed results are publicly available in Mancini et al. [73].

5.1. Impact of instance parameters

Averaged results, obtained by solving the proposed mathematical model (Section 3.2) by means of a commercial solver with a time limit of 3,600 seconds are reported in Table 3.

Table 2
Overview of instance set characteristics: number of demand scenarios (#DS), number of capacity scenarios (#CS), number of customers in each demand scenario (#CUST), number of locker facilities to open (#F), number of potential facility locations (#PL), facility capacity (C), and probability of locker box unavailability (P_UNAV). The latter is given as percentage value.

Set	#DS	#CS	#CUST	#F	#PL	C	P_UNAV
S1	5	5	20	5	10	5	10
S2	5	5	20	5	10	5	20
S3	5	5	20	5	10	5	30
S4	5	5	20	5	20	5	10
S5	5	5	20	5	30	5	10
S6	5	5	20	10	30	5	10
S7	5	5	20	15	30	5	10
S8	10	5	20	5	10	5	10
S9	20	5	20	5	10	5	10
S10	50	5	20	5	10	5	10
S11	10	5	20	5	30	5	10
S12	20	5	20	5	30	5	10
S13	5	10	20	5	30	5	10
S14	5	20	20	5	30	5	10
S15	5	5	40	5	10	5	10
S16	5	5	100	5	10	5	10
S17	5	5	200	5	10	5	10
S18	5	5	200	5	10	50	10
S19	5	5	100	5	30	25	10
S20	5	5	200	5	30	50	10

Table 3
Averaged exact results obtained within a time limit of 3,600 seconds. The table reports average value of the objective function (OF), upper bound (UB), percentage of customers served (%SERVED), and computational time (time).

Set	OF	UB	%SERVED	TIME (SEC.)
S1	462.92	462.92	92.56	2.22
S2	444.41	444.41	88.86	0.88
S3	421.22	421.23	84.22	0.63
S4	485.92	485.92	97.16	11.75
S5	493.37	493.38	98.66	70.92
S6	499.13	499.14	99.80	0.35
S7	500.08	500.10	100.00	0.20
S8	951.37	951.37	95.13	2.68
S9	1,782.35	1,782.35	89.12	2.82
S10	4,584.43	4,584.47	91.69	54.72
S11	981.99	982.00	98.19	454.01
S12	1,960.83	1,964.25	98.04	1,944.73
S13	987.09	987.82	98.70	816.56
S14	1,951.37	1,958.98	97.57	1,961.33
S15	578.27	578.27	57.82	0.42
S16	580.32	580.32	58.03	0.42
S17	592.91	592.91	59.29	0.60
S18	4,738.24	4,738.24	94.76	39.32
S19	2,458.22	2,477.13	98.33	3,177.33
S20	4,822.92	4,962.31	96.46	3,600.00

The objective of these experiments is to analyze the impact of instance features on computational times. The first analysis concerns the impact of the locker box unavailability probability. By comparing S1, S2, and S3, it can be seen that this parameter does not significantly influence computational times, as it always remains very short. Indeed, it even tends to decrease with the increase of the unavailability probability. It is also interesting to note that the number of customers served does not linearly decrease with the availability of locker boxes. Even in set S3, where the probability of unavailability is 30%, 84.22% of customers can be served.

The number of potential locations has a strong impact on computational times. While with 10 potential locations, the computational time is very low (see S1-S3 in Table 3), it rises considerably when the number of potential locations increases (S4 and S5). Obviously, the higher the number of potential locations, the higher is the percentage of customers who can be served.

Table 4
Comparison of CS against MODEL, ILS, and VNS on challenging instances. We report average values for the objective value of MODEL (OF) and run times (TIME). For CS, ILS, and VNS the average gaps of the objective value compared against MODEL are also reported (negative gaps indicate better solutions compared to MODEL).

SET	MODEL		CS		ILS		VNS	
	OF	TIME	GAP	TIME	GAP	TIME	GAP	TIME
S5	493.37	70.92	0.0%	20.99	0.0%	24.11	-0.2%	26.38
S11	981.99	454.01	-0.1%	40.89	-0.2%	143.30	-0.2%	143.16
S12	1,960.83	1,944.72	-0.1%	80.39	-0.1%	371.67	-0.1%	367.54
S13	2,458.22	3,177.33	-0.2%	87.68	-0.2%	343.42	-0.2%	384.49
S14	4,822.92	3,600.00	1.3%	129.74	1.2%	658.26	1.1%	713.20
S19	987.09	816.56	0.0%	56.12	0.0%	128.12	0.0%	136.95
S20	1,951.37	1,961.41	-0.1%	169.35	-0.2%	360.89	-0.2%	395.43
AVG	1,950.83	1,717.85	0.11%	83.56	0.07%	277.14	0.03%	298.28

If, while keeping fixed the number of potential locations, we increase the number of facilities to be opened, the problem becomes very easy to solve. In fact, with more facilities open, the total capacity increases, and almost all of the customers can be served. Therefore, finding the optimal solution becomes trivial. With 20 facilities open (S7), all of the customers can be served. Therefore, no acceptance or rejection decisions but only assignment decisions are required.

The number of demand scenarios influences the computational times, but only in combination with an increase in the potential locations. This can be observed by the surge between S8-S10 compared against S1-S12.

Furthermore, computational times strongly increase with the number of capacity scenarios as can be seen by comparing S4 and S5 against S13 and S14.

Increasing the number of customers per demand scenario, while keeping constant the locker stations' capacity does not have any impact on computational times. The problem remains very easy to solve (less than 1 second) even if we increase the number of customers to 200 (S17). However, if we also increase the locker station capacity from 5 to 50, which would allow to serve 250 customers if no capacity failure occurs, the problem becomes more difficult to solve (S18). Hence, as long as the capacity is rather very small compared to the number of customers, the latter parameter does not impact computational times, and the problem is very easy to solve. But the number of customers per scenario has a considerable impact if the capacity also grows.

5.2. Comparison of solution methods

To assess the performance of the newly proposed CS, we compare it against the solution of: (i) mathematical model presented in Section 3.2 (MODEL), (ii) ILS, and (iii) VNS, as described above. It is worth noting that in the local search phase of ILS and VNS, we exploit the same neighborhood we designed for CS. The differences among CS and the other two matheuristics lies in the following: (i) the procedure adopted to compute the initial solution, (ii) the order in which solutions belonging to the neighborhood are analyzed, and (iii) the diversification mechanism, which is not needed in CS, while it is a fundamental component for both ILS and VNS. Such a comparison has been carried out only on the most challenging sets of instances: S5, S11, S12, S13, S14, S19, and S20. In fact, the other sets can be solved to optimality so quickly that they do not require the usage of heuristic methods. The results are reported in Table 4, where for each set, the average results are considered. For the MODEL. a time limit of 3,600 seconds is imposed.

Negative gaps indicate that the algorithm obtains better solutions with respect to MODEL. This may happen because it was not possible to solve all the instances to optimality within the imposed time limit, and, therefore, some of the best solutions found by the model are suboptimal.

Table 5
Average iteration number, for each set of challenging instances, on which the best solution is found by CS, ILS, and VNS.

SET	CS	ILS	VNS
S5	6	154	161
S11	11	280	283
S12	12	257	266
S13	13	262	279
S14	20	272	288
S19	9	275	282
S20	7	268	288
AVG	11.14	252.57	263.86

The results reveal that all of the proposed matheuristic approaches show an excellent performance, obtaining near optimal solutions for instances which are solved to optimality by the MODEL. They strongly outperform the MODEL on the most challenging instance set (S14) for which the MODEL cannot provide an optimal solution. Although from a solutions' quality point of view, all three of the algorithms are effective, CS strongly outperforms ILS and VNS in terms of efficiency. On average, CS requires about 80% shorter average computational times than the other algorithms. This is due to the high quality of the initial solution and a smarter neighborhood exploration strategy. The fact that all three of the algorithms perform very well highlights the effectiveness of the newly proposed matheuristic local search routine. The idea of searching consensus among scenarios allows us to start from good quality solutions. Further, the consensus-searching mechanism that we use to select the candidate facility for the core seems to be highly effective to search the solution space.

In Table 5 we report for each set of instances tested with the matheuristics, the average iteration in which the best solution has been found by CS, ILS, and VNS. The trend is very clear and confirms that CS quickly converges towards very good solutions, while the improvement path of both ILS and VNS is much slower. This is due to the smart and effective strategy used by CS to select candidates to be part of the core.

In Figure 3 we plot, for CS, ILS, and VNS, the average iteration number, where a solution within 10%, 5%, 2%, and 1%, respecitvely, has been found. The graphic shows that, while CS reaches a gap of 1% in very few iterations, both ILS and VNS, quickly reach a gap of 5%, but require a very large number of iterations to pass from good to excellent solutions (within 1%).

Finally, we analyze in Figure 4 a single instance, namely I9 of set S5, and plot the evolution of the solution across iterations, for all the three methods. It is clear how, at each iteration along the solving process, the best solution obtained by CS is constantly better than those obtained by ILS and VNS.

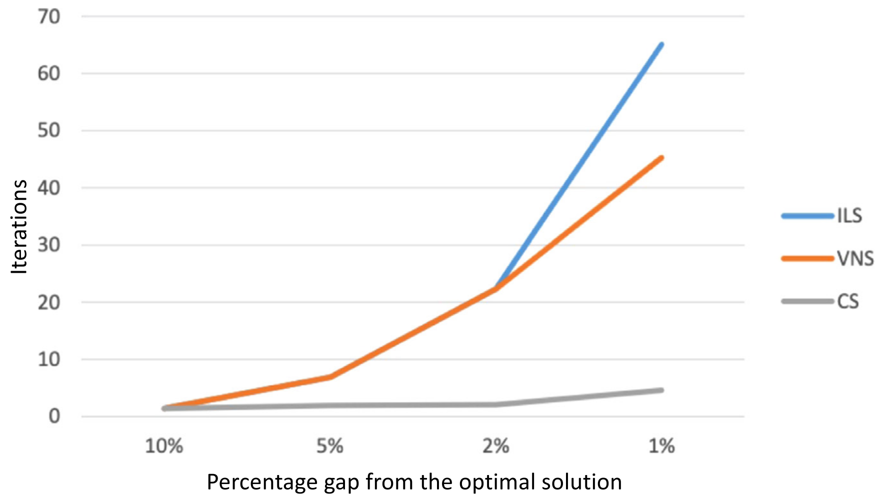

Fig. 3. Average iteration in which a solution within 10%, 5%, 2% and 1% from the optimum is reached by CS, ILS and VNS.

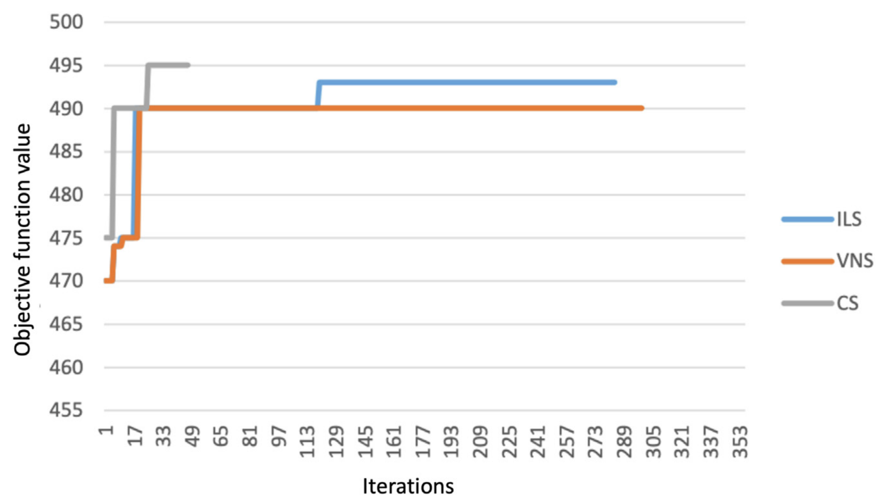

Fig. 4. Evolution of the current best solution of instance I9 from set S5, with CS, ILS and VNS.

5.3. Analysis of stochastic indicators

We conduct an analysis of the importance of recognizing the stochastic version of the problem through the use of two well-known and broadly used stochastic indicators, i.e., EVPI and VSS [74]. EVPI measures the expected cost of uncertainty. It represents the profit/loss due to uncertainty of data. The higher this value, the higher the importance of having precise information on input data. In a two-stage stochastic problem, where tactical and operational decisions have to be taken, a low value of EVPI means that the operational level does not have a strong impact on the fist, i.e., the tactical stage. In the opposite case (i.e., when EVPI is high), input data variation strongly impacts both stages.

In the context of the locker stations for last mile-delivery, a low EVPI value would mean that it would be reasonable to invest in having permanent lockers. On the contrary, a high EVPI would mean that, in the case of high variability of demand and uncertainty of input data, the company would do better to invest in mobile lockers, which can be easily moved every day around the city to meet customer demand. The EVPI can be calculated, in a maximization problem, as the difference between the objective value of the solution of the wait-and-see strategy (WS) and the objective value of the solution of the stochastic problem (SP). The WS strategy represents the utopian case in which perfect information about the input data is available a priori to make the first stage decisions. In other words, the first-stage variables are allowed to take different values for each scenario. In our case, in the WS strategy, we split the problem into several single-scenario problems, and each one is solved separately. This way, the set of facilities to be opened may vary among the different scenarios. The global objective function of the WS strategy is computed by summing up the objective values obtained in each scenario. We do not report the actual EVPI value, but we report the percentage gain that is obtainable after having perfect information about input data, computed as $EVPI = (WS - SP)/SP$.

VSS represents the gain of profit achievable by solving the SP with respect to solving a reference scenario, and then fixing the first stage variables such that they are equal to the value assumed in the optimal solution of the reference scenario. Generally, the reference scenario is an average scenario in which the values of uncertain parameters are substituted with their means. However, while this procedure perfectly suits those problems where the input data is distributed around a mean value, it does not apply to problems in which there is uncertainty about, say, the reduction of capacity or budget, such as in our case. In fact, considering an average scenario in which the capacity of the lockers is 10% less with respect to the actual capacity, we will automatically exclude the cases in which we have a full availability of capacity at some lockers. Hence, we adopt, as a reference scenario, the ideal scenario in which all of the boxes at a locker are available. The expected ob-

Table 6
Average values of stochastic indicators.

SET	SP	WS	EVPI	U^{EVPI}	EEV	VSS	U^{VSS}
S5	493.37	495.87	0.51%	2.5	485.56	1.63%	7.8
S11	981.99	984.79	0.29%	2.8	962.38	2.12%	19.6
S12	1,960.83	1,967.43	0.34%	6.6	1,923.61	1.96%	37.2

jective function value obtainable by fixing the first stage variables to the value, they would assume in the deterministic problem that is solved while replacing the uncertain parameters with their expected value, is referred to as the EEV [74]. VSS can be defined as $\text{VSS} = (SP - \text{EEV})/SP$.

We also compute two problem-specific parameters, i.e., U^{EVPI} and U^{VSS}. The former represents the number of home deliveries that can be avoided by having perfect information on input data, and it is computed as $\lfloor WS \rfloor - \lfloor SP \rfloor$, while U^{VSS} indicates the number of home deliveries avoided by solving the SP, instead of its deterministic version in which uncertain parameters are substituted by their expected values, and they are computed as $\lfloor SP \rfloor - \lfloor EEV \rfloor$.

In Table 6 we report the results for the stochastic indicators, for instance, for Sets S5, S11, and S12, which are of particular interest, as they reflect variations in demand scenarios.

The percentage value of EVPI is very low for all of the three sets. This means that having perfect information would imply only a little increment in customers served by locker boxes and consequently in home deliveries being avoided. Note that each customer that cannot be served by lockers, must be served by home delivery. This result supports the choice of locating permanent facilities rather than mobile ones.

Although, the percentage value of VSS is also quite low, it is very important to address the SP, and uncertainty in capacity availability cannot be neglected. In fact, solving the SP instead of solving its deterministic counterpart, in which uncertain parameters are substituted by their expected value, we can serve, by locker stations, up to 37 more customers (see U^{VSS} column) and consequently save up to 2% of home deliveries every day, which represents a huge reduction in routing costs. However, it is interesting to note that, increasing the number of demand scenarios from 5 (S5) to 20 (S12), the VSS value does not significantly change. This means that analyzing a small number of scenarios is already sufficient to capture the stochasticity of the problem. Also the number of capacity scenarios does not need to be particularly large, since several capacity reduction scenarios do not directly impact the optimal solution. For instance, if in the optimal solution of the case with full capacity availability, a facility j is used at the 80% of its actual capacity, all the scenarios in which its capacity is between 80% and 100% and the capacity of the other facilities open in the optimal solution are not reduced, will not impact at all the optimal solution. However, when the other facilities in the optimal solution experience a capacity reduction, the reduction of the capacity of j may play a role, since the additional capacity, initially not needed, could be helpful to mitigate the impact of the unavailability in the other facilities. However, the impact in these cases is generally not very relevant. The scenarios which actually impact the solution, are those in which the capacity reduction directly affects a facility fully (or almost fully) exploited in the optimal solution. In these cases, the impact of the unavailability can be strongly relevant and can generate a cascade effect on all the other facilities. Therefore, it is not necessary to analyze a large number of capacity scenarios, but it is sufficient to analyze those who really impact the decision problem. Hence, if we carefully select the scenarios to analyze, we can strongly limit their number without loosing relevant information. For these reasons, we consider only 5 demand scenarios and 5 capacity scenarios for the analysis of the real-world case presented in the following section.

In Figure 5 we depict, for the sample instance shown in Figure 1, the optimal solution of the reference scenario (EEV), in which all the facilities are assumed to have full capacity, and of the SP. Customers are depicted by blue circles, while facilities are indicated with squares, which are red if the facility is closed in the optimal solution and green if it is open. Each facility has a capacity of 10. Dotted orange circles represent facilities' coverage area. The number of facilities to open is 3. Five capacity reduction scenarios are considered: [2,0,0,0,0], [0,2,0,0,0], [0,0,2,0,0], [0,0,0,2,0], [0,0,0,0,2]. A single demand scenario is analyzed. In the optimal solution of EEV, the open facilities are A, C and E, which are able to cover all the 22 customers in the first four capacity scenarios, while in the fifth, where E has a capacity of 8, only 20 customers are covered. In the optimal solution of SP, the open facilities are A, D and E. This configuration allows to cover all the 22 customers in every scenario. In fact, the partial overlapping of coverage areas of D and E allows to remedy to the capacity reduction in E, by covering with D the 2 customers located in the intersection of the two coverage areas. This example emphasizes the importance of considering the stochasticity in the capacity availability.

5.4. Impact of facility utilization rate on unavailability probability

In this work we assume to have an estimation of the probability of locker box unavailability. This obviously holds for boxes, which were used in the previous delivery time slot (i.e., the day before), while for boxes, which were already empty the day before, the probability of finding them occupied is zero. This implies that if a facility is always underutilized the probability to find an unavailable box in this facility is very low. Hence, the probability that, when a capacity failure occurs, this actually impacts the optimal solution, is also almost zero. In fact, the only cases in which it can have an impact are those where a facility is underutilized on average, but is fully utilized in a specific scenario, and a capacity failure occurs exactly in that scenario, which is very unlikely. However, we modify the model in order to take into account that the probability of unavailability depends on the utilization rate, and compare the obtained results with those of the original model. The goal of this experiment is twofold. Firstly, we would like to analyze the impact of this dependency on the optimal solution, regarding customers serviceable by lockers. Secondly, we quantify the impact on computational times and difficulty to solve the problem.

The capacity dependency can be modeled substituting constraints (2) with the following ones.

$$\sum_{i \in I_s} Y_{ij}^{\omega} \leq C - \delta_j^{\omega} u_j / C \quad \forall \omega \in \Omega \;\; \forall s \in S \;\; \forall j \in J, \tag{9}$$

where u_j is a variable representing the number of boxes used in facility j:

$$u_j \geq \sum_{\omega \in \Omega} \sum_{i \in I} Y_{ij}^{\omega} \qquad \forall j \in J. \tag{10}$$

We test the new model on the most challenging set of instances: S5, S11 and S12. Results are reported in Table 7.

As we can notice from the table, the benefit of considering the unavailability dependency on utilization is negligible. In set S5, the number of customers served is exactly the same. In S11, the gain achievable is very low (0.2 additional customers served on average over 1,000), while in S12 results obtained considering dependency are even slightly worse. This happens because, while with the original model we solve to the optimality all instances, the new model is more complex to solve and therefore a lower number of instances can be solved. In fact, computational times and also runtime memory required by the new model are much larger.

Resuming, we observe that, even if availability probability actually depends on the utilization rate, considering this dependency

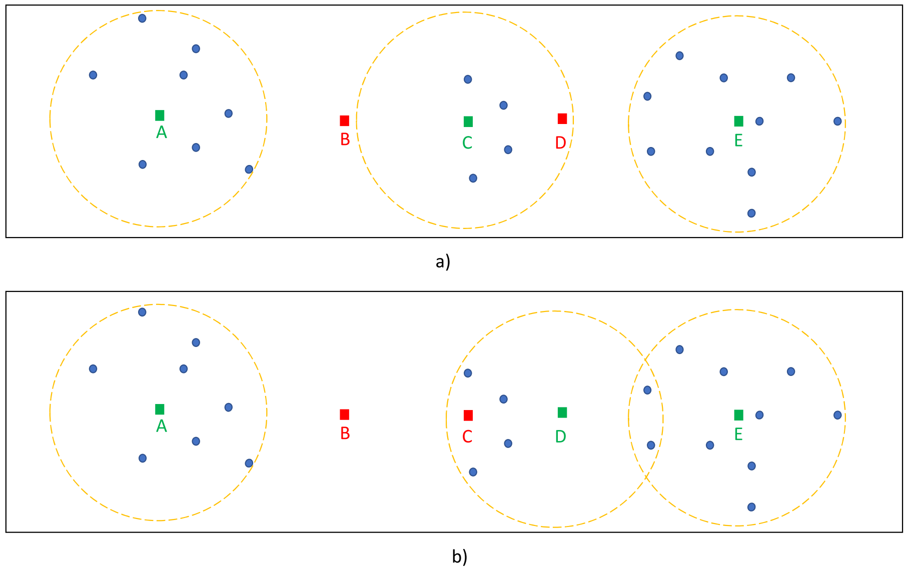

Fig. 5. Optimal solution for the EEV (a) and the SP (b) for a sample instance. If the capacity of E, which is assumed to be 10, is reduced, not all customers can be served by the locker station.

Table 7
Comparison of results considering unavailability probability independent or dependent from utilization rate. We compare the number of customers served and the computational time elapsed (in seconds) with a time limit of 3,600 seconds (if the time limit is reached, the best solution found is considered).

	#CUSTOMERS SERVED		COMPUTATIONAL TIME (s)	
	INDEPENDENT	DEPENDENT	INDEPENDENT	DEPENDENT
S5	493.3	493.3	71	91
S11	981.9	982.1	454	853
S12	1,960.8	1,959.1	1,945	3,079

does not yield any significant benefit, but makes the problem considerably more complex to solve.

5.5. A real-world case in the city of Turin

To investigate a real-world case, we considered Turin, a city of 900 thousand inhabitants, located in the north west of Italy. The city is divided into 34 districts and grouped into 10 wards as depicted in Figure 6.

We used Google maps[1] to derive the addresses of 80 locker stations, spread all over the city. However, in our analysis, we focused only on the southern part of the city, which comprises wards 1, 2, 3, 8, 9, and 10, which cover the districts from 1 to 9 and from 28 to 34, adding up to a total of 16 districts. In this area, 40 locker stations were available, but we considered 40 additional potential locations at refueling stations, large super markets, and shopping malls. All of these 80 lockers (40 existent and 40 potential) that were covered by this study are illustrated in Figure 7.

For each district, the city of Turin provides, in the public domain, information about the number of inhabitants for certain age classes[2]. Five such classes were considered (0-17, 18-30, 31-45, 46-65, and > 65). The number of inhabitants per class and the distribution of population per age class, are reported in Table 9 in the Appendix. As can be noted from the table, the distribution of population significantly varies from one district to another. For instance, while in District 2 (Crocetta), a significant number of the inhabitants are aged over 65 (34%), in Cavoretto and Borgo Po only 5.14% are aged over 65, while the largest part of the population belongs to Classes 31-45 and 46-65 (37.6% and 39.9%, respectively). Furthermore, only 1.85% are aged under 18 years. In District 3 (Santa Rita), the largest quote is for the Class 46-65 (40%), while in District 4 (Mirafiori Nord), young adults are almost absent (only 0.16%).

The age distribution of the population is of crucial importance for this study, as the willingness to select the option of delivery to locker stations, strongly varies among age classes. We used the data from a survey conducted in Austria, which was used to determine the probability of acceptance of this delivery option and cities of different sizes. In Figure 8, we report data for the 4 typologies of cities. Turin belongs to the large city category (upper-left graphic in the figure).

The data from the the survey indicates that, for this category, the probability of using locker stations is quite homogeneous for the first 4 age classes (from 67% to 71%), whereas it considerably decreases for older people (only 47%). It is worth noting that for different city typologies, remarkable differences in customers' behaviors can be observed. Combining data about the population distribution and the willingness to use lockers, we calculated the probability of a customer being located in a specific district as reported in Table 10 in the Appendix. We used these probabilities, as determined, to create 5 demand scenarios. The number of customers per scenario was 200. The number of capacity scenarios was 5. Each of the 80 locker stations being considered has a capacity of 10 locker boxes. We assumed that 40 locker stations have to be opened. We considered two probability levels for capacity unavailability (10% and 30%) and two compatibility thresholds for the customers being assigned to the locker stations (500m and 1km).

The layout of the instance is reported in Figure 9. The customers are marked in green, existent locker stations in red, and potential locations in blue. Distances among points have been computed using the Manhattan distance formula. Since Turin's map is very similar to the Manhattan one, i.e., it is composed of a set of orthogonal streets, using Manhattan distances seem to be perfectly suitable.

The results are presented in Table 8. For each combination of P_unav and compatibility threshold, we report the optimal number of customers served using only the 40 locker stations actually

[1] www.google.it/maps/.
[2] http://aperto.comune.torino.it/.

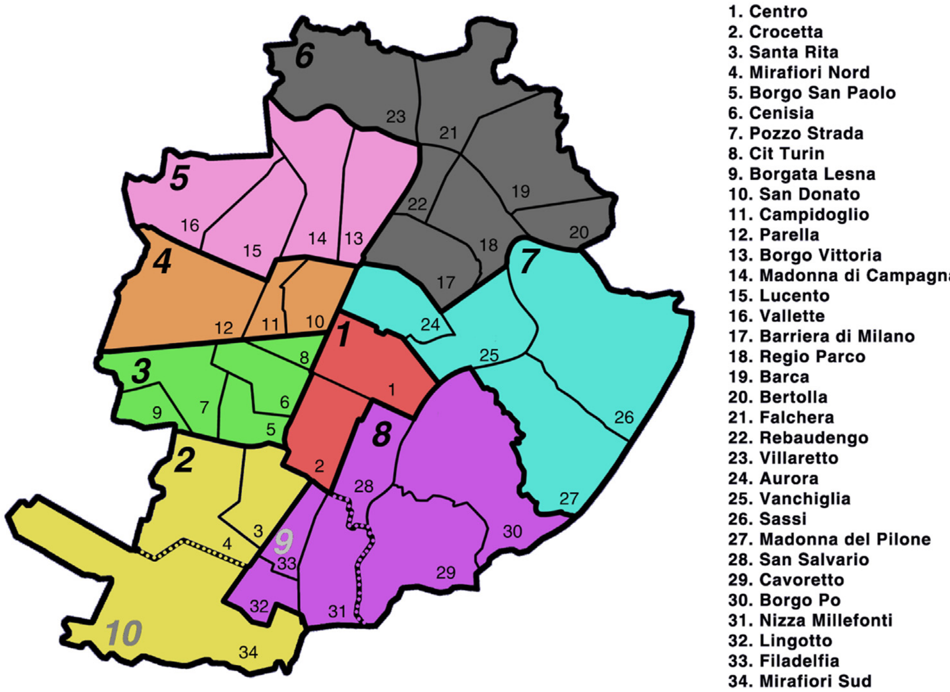

Fig. 6. Turin's districts.[a] [a] https://commons.wikimedia.org/wiki/File:Circoscrizioni_torino_2016.png is licensed under CC BY-SA 4.0

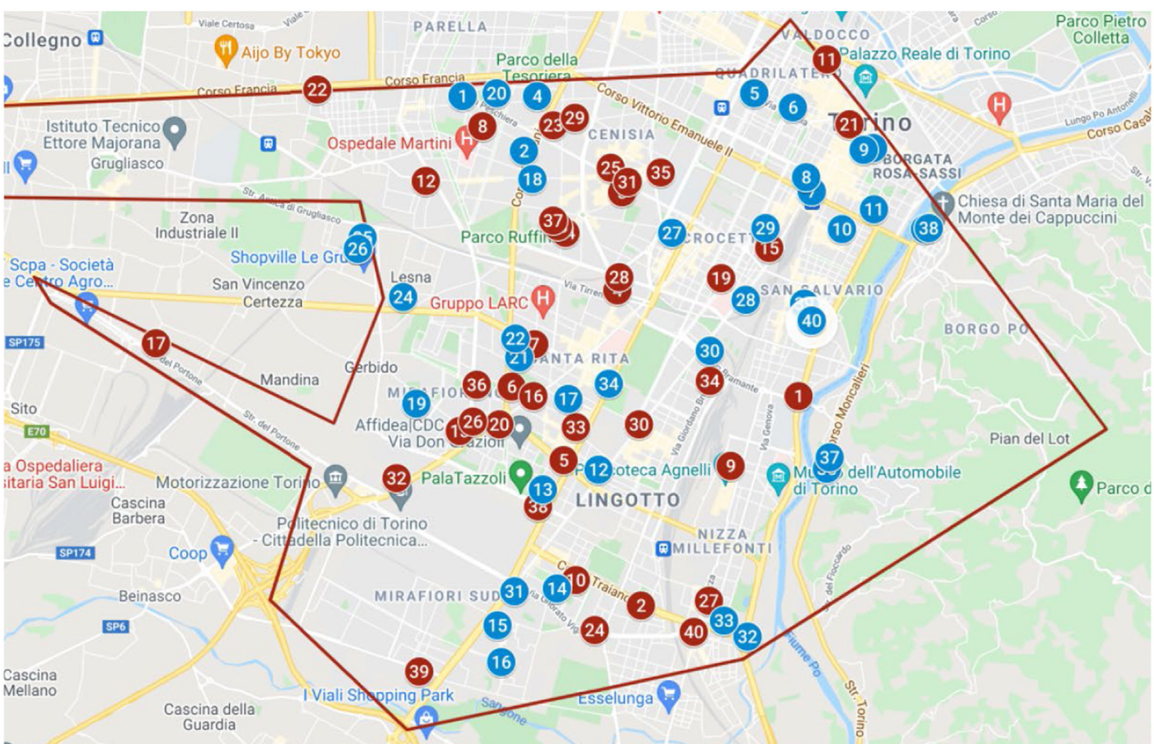

Fig. 7. Available and potential locker stations in certain areas of Turin; potential locations are marked in blue, while existent ones are in red.[b] [b] www.maps.google.com

Table 8
Results of the Turin case with two compatibility thresholds and two reduction of capacity scenarios.

	THRESHOLD	$\delta^{\omega 1}$=10%		$\delta^{\omega 1}$=30%	
		OLD	OLD+NEW	OLD	OLD+NEW
# cust. served	0.5 km	4445	4535	4405	4524
	1 km	4955	4975	4955	4975
avg. dist. (m)	0.5 km	389	340	405	340
	1 km	531	475	573	604

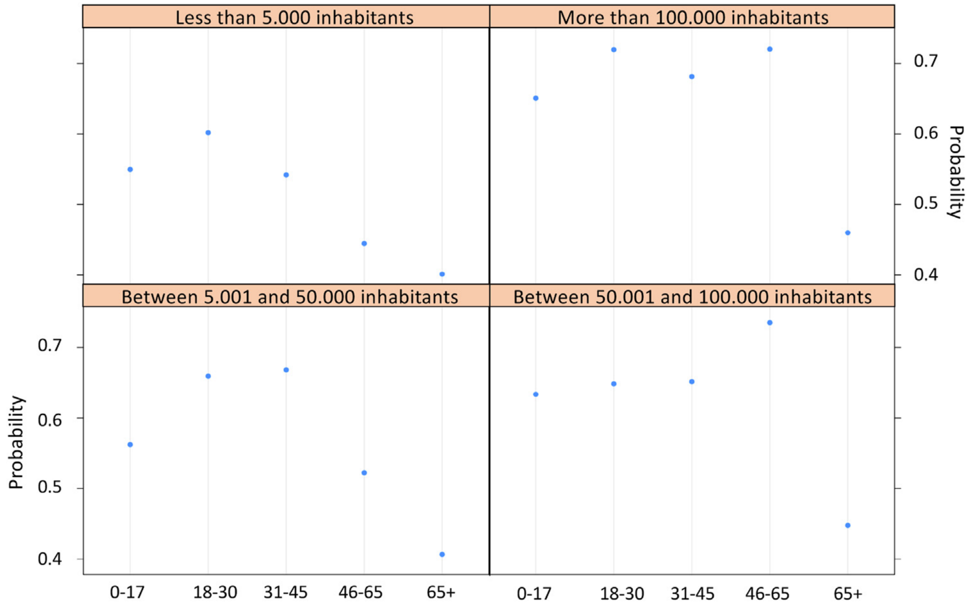

Fig. 8. Customers' willingness to use locker stations per age class.

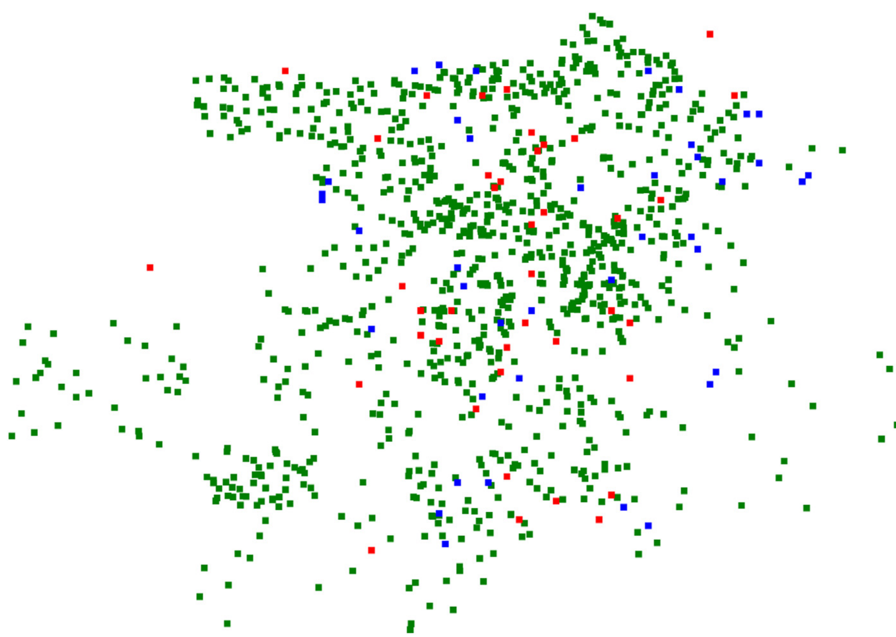

Fig. 9. Instance layout for the case study. Customers are depicted in green, existent locker stations in red, and potential locations in blue.

in use (OLD) and those serviceable choosing the best 40 locations among the 80 available (OLD+NEW). We also report the average distance covered by customers to reach a locker station.

When the compatibility threshold is larger, the location of the locker station becomes a less critical issue. Although the percentage improvement achievable, considering more potential locations is quite small (0.4%), it corresponds to a total saving of 20 customers (i.e., 4 per demand scenario on average). In terms of the routing costs, this increment in the number of home deliveries required is not necessarily costly, but it might still impact routing costs. The probability of unavailability does not impact the number of served customers at all, but it does yield a considerable increment in the distance covered by customers. When the compatibility threshold is smaller, the location of locker stations plays a crucial role. In fact, exploiting a larger number of potential locations allows to save 119 home deliveries, which can presumably effect a huge reduction in routing costs. In this case, the probability of unavailability impacts the number of serviceable customers. When P_Unav grows to 30%, the number of serviceable customers is reduced to 40 for the OLD configuration, while with the OLD+NEW configuration, we have a reduction of only 11 customers. This means that the impact of unavailability is smaller if the locker station distribution is smarter.

For what concerns the covered distance, the OLD+NEW configuration guarantees the same average covered distance independent of the probability of unavailability, while with the OLD configuration shows a slight increment in distances. In terms of covered distance, we observe that the average distance does not grow linearly. For instance, by allowing a maximum covered distance of 500 meters, the average covered distance is 340, while by allowing a maximum of 1 km, the average covered distance is only 475 meters. This seems to be based on a saturation level of the required covered distance.

6. Conclusions and future developments

In this paper, we studied the problem of locker station location under conditions of uncertain demand and capacity availability. The problem is modeled as an extension of the capacitated facility location problem in which a fixed number of facilities have to be open, choosing among a set of potential locations. The facilities have homogeneous capacities; however, a capacity reduction can occur for a given probability. A set of different demand and capacity scenarios was considered. Each customer can be assigned only to compatible facilities, i.e., those facilities located within a given radius. The primary objective of solving the problem is to maximize the total number of customers assigned to locker stations; however, in case of a tie on the primary objective, the secondary objective is to minimize the average distance covered by customers to reach the locker station to which they have been assigned.

To solve the problem, we developed a mathematical model and three matheuristics, two of which are based on established frameworks, i.e., ILS and VNS, while the third one, CS, is a newly proposed approach. All of the three models exploit the same local search mechanism designed for this specific problem, but they utilize different procedures to generate an initial solution, different neighborhood exploration strategy, and different diversification schemes.

All of the proposed matheuristics show excellent performance in terms of the solution's quality, but CS strongly outperforms the other two in terms of computational times. A detailed analysis of the impact of instance features on the difficulty of the problem is provided through an extensive computational campaign. The importance of considering the SP, instead of solving its deterministic equivalent variant, is discussed by means of standard stochastic indicators, namely EVPI and VSS, as well as some problem-specific stochastic indicators. This analysis showed that it is important to address the stochastic version of the problem, as considering uncertainty of input data may allow to serve several additional customers and, consequently, to avoid several costly home deliveries.

A real-world case related to the city of Turin (Italy) is discussed. We have shown that smartly located locker stations can save a considerable number of home deliveries.

Future developments in this field could focus, from a methodological point of view, on the adaptation of the CS framework to a broad class of two-stage SPs as well as a specific class of bi-level problems, with one first-level decision maker and several second-level ones. From an application point of view, it would be interesting to explicitly include routing aspects, addressing a location-routing version of the problem, as well as to consider a distribution system involving more delivery options, such as roaming delivery, cargo bikes, and pedestrian and occasional drivers. Other interesting research avenues could be the fortification techniques that could mitigate the impact of uncertainty in input data, as well as the application of the SC reduction to other facility location problems, arising in different fields of application. It would be also interesting to perform an ex-post analysis on how the reduction of the number of delivery stops actually impacts on the reduction on routing costs. Another interesting development could cover the definition of a chance-constraint model, where in each scenario, a minimum number of satisfied customers must be guaranteed with a given probability.

Author statement

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication.

Declaration of Competing Interest

None.

Data availability

The instances data set has been made publicly available in mendeley

Acknowledgments

The authors would like to thank Hanna Gnatchuk, who kindly provided the results of the survey conducted in Austria. The authors would like to thank Andrea Isabello as well for the relevant information he shared regarding city statistics. Margaretha Gansterer appreciates the support of the Austrian Science Fund (FWF): P 34502-N and P 34151. Simona Mancini appreciates the support of the Fund for University Research of the University of Palermo: FFR 2023.

Appendix

This Appendix includes some additional input data related to our real-world case study. More specifically, Table 9 reports the number of inhabitants in each district of the city of Turin together with the population distribution per age classes and Table 10 includes the probability of a customer to be located in each of these districts.

Table 9
Number of inhabitants and population distribution for age classes in Turin.

		# INHABITANTS						POPULATION DISTRIBUTION				
ID	NAME	0-17	18-30	31-45	46-65	≥ 66	Total	0-17	18-30	31-45	46-65	≥ 66
1	Centro	5374	5755	8760	13499	8929	42317	12.70%	13.60%	20.70%	31.90%	21.10%
2	Crocetta	4681	4287	9625	10106	15050	43749	10.70%	9.80%	22.00%	23.10%	34.40%
3	Santa Rita	7087	6555	4097	16510	6719	40968	17.30%	16.00%	10.00%	40.30%	16.40%
4	Mirafiori Nord	5679	47	7395	12436	8306	33863	16.77%	0.14%	21.84%	36.72%	24.53%
5	Borgo San Paolo	4545	4280	6515	10076	12462	37878	12.00%	11.30%	17.20%	26.60%	32.90%
6	Cenisia	2163	2102	4511	4837	6797	20410	10.60%	10.30%	22.10%	23.70%	33.30%
7	Pozzo Strada	7367	6441	477	16353	7244	37882	19.45%	17.00%	1.26%	43.17%	19.12%
8	Cit Turin	2644	2569	5512	5911	8306	24943	10.60%	10.30%	22.10%	23.70%	33.30%
9	Borgata Lesna	4050	3541	262	8991	3983	20827	19.45%	17.00%	1.26%	43.17%	19.12%
28	San Salvario	4744	4398	8051	1012	13002	31207	15.20%	14.09%	25.80%	3.24%	41.66%
29	Cavoretto	121	1015	2465	2613	337	6552	1.85%	15.49%	37.63%	39.89%	5.14%
30	Borgo Po	161	1346	3269	3465	446	8686	1.85%	15.49%	37.63%	39.89%	5.14%
31	Nizza Millefonti	3735	3572	7762	8119	9288	32476	11.50%	11.00%	23.90%	25.00%	28.60%
32	Lingotto	528	4609	882	11380	8090	25489	2.07%	18.08%	3.46%	44.65%	31.74%
33	Filadelfia	109	949	181	2343	1665	5247	2.07%	18.08%	3.46%	44.65%	31.74%
34	Mirafiori Sud	4671	4327	5298	10315	6741	31352	14.90%	13.80%	16.90%	32.90%	21.50%

Table 10
Customers' distribution probability per district in Turin.

ID	NAME	PROB
1	Centro	9.72%
2	Crocetta	9.56%
3	Santa Rita	9.57%
4	Mirafiori Nord	7.65%
5	Borgo San Paolo	8.33%
6	Cenisia	4.48%
7	Pozzo Strada	8.78%
8	Cit Turin	5.47%
9	Borgata Lesna	4.83%
28	San Salvario	6.60%
29	Cavoretto	1.59%
30	Borgo Po	2.12%
31	Nizza Millefonti	7.25%
32	Lingotto	5.69%
33	Filadelfia	1.17%
34	Mirafiori Sud	7.19%

References

[1] Melacini M, Perotti S, Rasini M, Tappia E. E-fulfilment and distribution in omni-channel retailing: a systematic literature review. International Journal of Physical Distribution & Logistics Management 2018.
[2] Olsson J, Hellström D, Pålsson H. Framework of last mile logistics research: A systematic review of the literature. Sustainability 2019;11(24):7131.
[3] Nieto-Isaza S, Fontaine P, Minner S. The value of stochastic crowd resources and strategic location of mini-depots for last-mile delivery: a benders decomposition approach. Transportation Research Part B: Methodological 2022;157:62–79.
[4] Charged Retail. 2019a. Last accessed on 12 March 2022; https://www.chargedretail.co.uk/2019/12/23.
[5] Wen J, Li Y. Vehicle routing optimization of urban distribution with self-pick-up lockers. In: 2016 International Conference on Logistics, Informatics and Service Sciences (LISS). IEEE; 2016. p. 1–6.
[6] Sitek P, Wikarek J. Capacitated vehicle routing problem with pick-up and alternative delivery (cvrppad): model and implementation using hybrid approach. Annals of Operations Research 2019;273(1):257–77.
[7] Mancini S, Gansterer M. Vehicle routing with private and shared delivery locations. Computers & Operations Research 2021;133:105361.
[8] Dumez D, Lehuede F, Peton O. A large neighborhood search approach to the vehicle routing problem with delivery options. Transportation Research Part B: Methodological 2021;144:103–32.
[9] Buzzega G, Novellani S. Last mile deliveries with lockers: formulations and algorithms. Soft Computing 2022:1–19.
[10] Barenji AV, Wang W, Li Z, Guerra-Zubiaga DA. Intelligent e-commerce logistics platform using hybrid agent based approach. Transportation Research Part E: Logistics and Transportation Review 2019;126:15–31.
[11] Morganti E, Seidel S, Blanquart C, Dablanc L, Lenz B. The impact of e-commerce on final deliveries: Alternative parcel delivery services in France and Germany. Transportation Research Procedia 2014;4:178–90.
[12] Campbell AM, Savelsbergh M. Incentive schemes for attended home delivery services. Transportation Science 2006;40(3):327–41.
[13] Bruck BP, Cordeau J-F, Iori M. A practical time slot management and routing problem for attended home services. Omega 2018;81:208–19.
[14] Köhler C, Ehmke JF, Campbell AM. Flexible time window management for attended home deliveries. Omega 2020;91:102023.
[15] Agatz N, Campbell A, Fleischmann M, Savelsbergh M. Time slot management in attended home delivery. Transportation Science 2011;45(3):435–49.
[16] Yang X, Strauss A. An approximate dynamic programming approach to attended home delivery management. European Journal of Operational Research 2017;263:935–45.
[17] Hernandez F, Gendreau M, Potvin J-Y. Heuristics for tactical time slot management: a periodic vehicle routing problem view. International Transactions in Operational Research 2017;24(6):1233–52.
[18] Florio AM, Feillet D, Hartl RF. The delivery problem: Optimizing hit rates in e-commerce deliveries. Transportation Research Part B: Methodological 2018;117:455–72.
[19] Schnieder M, Hinde C, West A. Combining parcel lockers with staffed collection and delivery points: An optimization case study using real parcel delivery data (london, uk). Journal of Open Innovation: Technology, Market and Complexity 2021;183(7).
[20] Moroz M, Polkowski Z. The last mile issue and urban logistics: choosing parcel machines in the context of the ecological attitudes of the y generation consumers purchasing online. Transportation Research Procedia 2016;16:378–93.
[21] Seghezzi A, Siragusa C, Mangiaracina R. Parcel lockers vs. home delivery: a model to compare last-mile delivery cost in urban and rural areas. International Journal of Physical Distribution & Logistics Management 2022.
[22] Charged Retail. 2019b. Last accessed on 12 March 2022; https://www.chargedretail.co.uk/2019/09/18.
[23] Charged Retail. 2021a. Last accessed on 12 March 2022; https://www.chargedretail.co.uk/2021/12/20.
[24] Charged Retail. 2021b. Last accessed on 12 March 2022; https://www.chargedretail.co.uk/2021/04/22.
[25] Infinium Logistics. 2021. Last accessed on 12 March 2022; https://dev.mashuni.com/infinium/parcel-lockers-bridging-the-final-mile-in-urban-deliveries/.
[26] Lemke J, Iwan S, Korczak J. Usability of the parcel lockers from the customer perspective–the research in polish cities. Transportation Research Procedia 2016;16:272–87.
[27] Kedia A, Kusumastuti D, Nicholson A. Acceptability of collection and delivery points from consumers perspective: A qualitative case study of christchurch city. Case Studies on Transport Policy 2017;5(4):587–95.
[28] Lachapelle U, Burke M, Brotherton A, Leung A. Parcel locker systems in a car dominant city: Location, characterisation and potential impacts on city planning and consumer travel access. Journal of Transport Geography 2018;71:1–14.
[29] Yuen KF, Wang X, Ng LTW, Wong YD. An investigation of customers intention to use self-collection services for last-mile delivery. Transport Policy 2018;66:1–8.
[30] Iannaccone G, Marcucci E, Gatta V. What young e-consumers want? forecasting parcel lockers choice in rome. Logistics 2021;5(3):57.
[31] Schaefer JS, Figliozzi MA. Spatial accessibility and equity analysis of amazon parcel lockers facilities. Journal of Transport Geography 2021;97:103212.
[32] Lagorio A, Pinto R. The parcel locker location issues: An overview of factors affecting their location. In: Proceedings of the 8th International Conference on Information Systems, Logistics and Supply Chain: Interconnected Supply Chains in an Era of Innovation, ILS; 2020. p. 414–21.
[33] Faugere L, Montreuil B. Hyperconnected city logistics: smart lockers terminals & last mile delivery networks. In: Proceedings of the 3rd International Physical Internet Conference, Atlanta, GA, USA, vol. 29; 2016.
[34] Ulukan Z, Demircioğlu E. A survey of discrete facility location problems. International Journal of Industrial and Manufacturing Engineering 2015;9(7):2487–92.
[35] Jalal AM, Toso EAV, Morabito R. Integrated approaches for logistics network

planning: a systematic literature review. International J of Production Research 2021:1–29.
[36] Wang Y, Ong T, Lee LH, Chew EP. Capacitated competitive facility location problem of self-collection lockers by using public big data. In: 2017 13th IEEE Conference on Automation Science and Engineering (CASE). IEEE; 2017. 1344–1344
[37] Deutsch Y, Golany B. A parcel locker network as a solution to the logistics last mile problem. International Journal of Production Research 2018;56(1-2):251–61.
[38] Lee H, Chen M, Pham HT, Choo S. Development of a decision making system for installing unmanned parcel lockers: Focusing on residential complexes in korea. KSCE Journal of Civil Engineering 2019;23(6):2713–22.
[39] Schwerdfeger S, Boysen N. Optimizing the changing locations of mobile parcel lockers in last-mile distribution. European Journal of Operational Research 2020;285(3):1077–94.
[40] Lin YH, Wang Y, He D, Lee LH. Last-mile delivery: Optimal locker location under multinomial logit choice model. Transportation Research Part E: Logistics and Transportation Review 2020;142:102059.
[41] Yang G, Huang Y, Fu Y, Huang B, Sheng S, Mao L, Huang S, Xu Y, Le J, Ouyang Y, et al. Parcel locker location based on a bilevel programming model. Mathematical Problems in Engineering 2020;2020.
[42] Simić V, Lazarević D, Dobrodolac M. Picture fuzzy waspas method for selecting last-mile delivery mode: a case study of belgrade. European Transport Research Review 2021;13(1):1–22.
[43] Zheng Z, Morimoto T, Murayama Y. A gis-based bivariate logistic regression model for the site-suitability analysis of parcel-pickup lockers: A case study of guangzhou, china. ISPRS International Journal of Geo-Information 2021;10(10):648.
[44] Oliveira WJP, dos Santos AG. Last mile delivery with lockers: Formulation and heuristic. In: ICEIS (1); 2020. p. 460–7.
[45] Veenstra M, Roodbergen KJ, Coelho LC, Zhu SX. A simultaneous facility location and vehicle routing problem arising in health care logistics in the netherlands. European Journal of Operational Research 2018;268(2):703–15.
[46] Serrano-Hernandez A, Martinez-Abad S, Ballano A, Faulin J, Rabe M, Chicaiza-Vaca J. A hybrid modeling approach for automated parcel lockers as a last-mile delivery scheme: a case study in pamplona (spain). In: 2021 Winter Simulation Conference (WSC). IEEE; 2021. p. 1–12.
[47] Snyder LV. Facility location under uncertainty: a review. IIE transactions 2006;38(7):547–64.
[48] Suryawanshi P, Dutta P. Optimization models for supply chains under risk, uncertainty, and resilience: A state-of-the-art review and future research directions. Transportation Research Part E: Logistics and Transportation Review 2022;157:102553.
[49] Bieniek M. A note on the facility location problem with stochastic demands. Omega 2015;55:53–60.
[50] Pagès-Bernaus A, Ramalhinho H, Juan AA, Calvet L. Designing e-commerce supply chains: a stochastic facility–location approach. International Transactions in Operational Research 2019;26(2):507–28.
[51] Wang Y, Bi M, Lai J, Chen Y. Locating movable parcel lockers under stochastic demands. Symmetry 2020;12(12):2033.
[52] Rabe M, Chicaiza-Vaca J, Tordecilla RD, Juan AA. A simulation-optimization approach for locating automated parcel lockers in urban logistics operations. In: 2020 Winter Simulation Conference (WSC). IEEE; 2020. p. 1230–41.
[53] Rabe M, Gonzalez-Feliu J, Chicaiza-Vaca J, Tordecilla RD. Simulation-optimization approach for multi-period facility location problems with forecasted and random demands in a last-mile logistics application. Algorithms 2021;14(2):41.
[54] Kahr M. Determining locations and layouts for parcel lockers to support supply chain viability at the last mile. Omega 2022;113:102721.
[55] Bell JE, Griffis SE. Military applications of location analysis. In: Applications of location analysis. Springer; 2015. p. 403–33.
[56] Aydin N. A stochastic mathematical model to locate field hospitals under disruption uncertainty for large-scale disaster preparedness. An International Journal of Optimization and Control: Theories & Applications (IJOCTA) 2016;6(2):85–102.
[57] Yu G, Haskell WB, Liu Y. Resilient facility location against the risk of disruptions. Transportation Research Part B: Methodological 2017;104:82–105.
[58] Rohaninejad M, Sahraeian R, Tavakkoli-Moghaddam R. An accelerated benders decomposition algorithm for reliable facility location problems in multi-echelon networks. Computers & Industrial Engineering 2018;124:523–34.
[59] Florez JV, Lauras M, Okongwu U, Dupont L. A decision support system for robust humanitarian facility location. Engineering Applications of Artificial Intelligence 2015;46:326–35.
[60] Cheng C, Adulyasak Y, Rousseau L-M. Robust facility location under demand uncertainty and facility disruptions. Omega 2021;103:102429.
[61] Ulmer M, Streng S. Same-day delivery with pickup stations and autonomous vehicles. Computers & Operations Research 2019;108:1–19.
[62] Dönmez Z, Bahar Y, Karsu O, Saldanha-da Gama F. Humanitarian facility location under uncertainty: Critical review and future prospects. Omega 2021;102:102393.
[63] Aringhieri R, Duma D, Landa P, Mancini S. Combining workload balance and patient priority maximisation in operating room planning through hierarchical multi-objective optimisation. European Journal of Operational Research 2022;298(32):627–43.
[64] Ceschia S, Gansterer M, Mancini S, Meneghetti A. The on-demand warehousing problem. International Journal of Production Research 2023;61(10):3152–70. DOI: 10.1080/00207543.2022.2078249
[65] Bent R, Van Hentenryck P. Scenario-based planning for partially dynamic vehicle routing with stochastic customers. Operations Research 2004;52(6):977–87.
[66] Van Hentenryck P, Bent R. Online stochastic combinatorial optimization. The MIT press; 2006.
[67] Crainic T, Fu X, Gendreau M, Rei W, Wallace S. Progressive hedging-based metaheuristics for stochastic network design. Networks 2011;58(2):114–24.
[68] Lourenço HR, Martin OC, Stützle T. Iterated Local Search. Boston, MA: Springer US; 2003. p. 320–53.
[69] Mladenović N, Hansen P. Variable neighborhood search. Computers & Operations Research 1997;24(11):1097–100.
[70] Ausseil R, Pazour J, Ulmer M. Supplier menus for dynamic matching in peer-to-peer transportation platforms. Transportation Science 2022;56(5):1111–408.
[71] Mladenović N, Pei J, Pardalos P, Urosević D. Less is more approach in optimization: a road to artificial intelligence. Optimization Letters 2022;16(1):409–20.
[72] Turkes R, Sorensen K, Hvattum L. Meta-analysis of metaheuristics: Quantifying the effect of adaptiveness in adaptive large neighborhood search. European Journal of Operational Research 2021;292(2):423–42.
[73] Mancini S, Gansterer M, Triki C. Instances and results for locker box location planning under uncertainty in demand and capacity availability. Mendeley Data 2022;v2. doi:10.17632/33v84jrtk2.2.
[74] Birge JR, Louveaux F. Introduction to stochastic programming. Springer; 2011.