Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal

Filippo Maria Ottaviani
filippo.ottaviani@polito.it
Politecnico di Torino, Italy

Giovanni Zenezini*
giovanni.zenezini@polito.it
Politecnico di Torino, Italy

Alberto De Marco
alberto.demarco@polito.it
Politecnico di Torino, Italy

Antonio Carlin
antonio.carlin@polito.it
Politecnico di Torino, Italy

Keywords
Last-mile Logistics
Automated Parcel Lockers
Facility Location Problem
MILP
Heuristics algorithm

Publishing history
Submitted: 18 November 2022
Accepted: 14 June 2023
Published: 17 July 2023

Cite as

© 2023 Filippo Maria Ottaviani, Giovanni Zenezini, Alberto De Marco, Antonio Carlin

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

* corresponding author

Logistics Service Providers (LSP) are increasingly adopting Automated Parcel Lockers (APLs) to mitigate the operational pressure of last-mile logistics. The optimal location of APL stations is key for reaching customers’ demand while keeping the investment reasonable. Previous studies developed optimization algorithms and applied them to virtual instances of the problem, lacking applicability to real-life situations encountered by LSPs who aim to serve an urban area with such technology. This study proposes a novel solution to the APLs location problem by combining mixed-integer linear programming and greedy heuristics algorithms. The study tested the propose solution on real customers’ demand data related to Turin, Italy. Results show that covering 90% of the estimated potential demand requires 10 to 11 APLs, on average. The adopted approach enables finding an optimal solution grounded in a real geographical context without requiring time-consuming optimization.
1 Introduction

E-commerce has steadily increased in recent years, with annual growth rates of more than 20% until 2020 (Cramer-Flood, 2023). Even though the growth has been dwindling after the pandemic surge of 2020, e-commerce is still outperforming overall sales growth.

E-commerce home deliveries are characterized by a high frequency of orders, dispersed demand, and the possibility of customers missing the delivery (Seghezzi et al., 2022). Home deliveries are also time-consuming since drivers need to decelerate, look for parking, unload the parcel, and wait for the consignee. Additionally, last-mile logistics generates a significant number of vehicle-kilometres for each delivery route, reducing the profitability of private operators and increasing the emissions of pollutants (e.g., CO2, NOx, PM10) (Arroyo et al., 2020). Therefore, Logistics Service Providers (LSPs) must increase the productivity of their vehicle fleet, measured in the number of stops serviced by each vehicle during a daily tour, to fulfill the increasing request for last-mile deliveries while reducing their negative impacts (Cagliano et al., 2017).

Consolidating users’ demand in APLs is a possible solution to the operational strains of last-mile deliveries. APLs are a type of unattended collection-and-delivery point installed in public and private areas where parcels are retained for a limited amount of time until the customer can retrieve them by using a unique code (e.g., QR code, order reference number, a PIN code) (Lai et al., 2022) or a form of identification (e.g., credit card, government issued ID) (Schnieder et al., 2021). Consolidating the demand in APLs reduces the number of delivery points, increasing vehicle productivity and reducing mean delivery time (Bailey et al., 2013).

However, APLs require investments by private operators and compel space organizers such as public authorities to grant the usage of portions of public space wherein APLs are installed (Zenezini et al., 2018). Moreover, the diffusion of APLs depends on the perception of potential customers and their inclination toward adopting technological innovations that change customer habits. While APLs implementation can instigate co-value generation processes with customers, on the other hand, it can destroy value when failures occur. This results in a vicious cycle that leads to the customer abandoning the service (Vakulenko et al., 2018).

Compatibility with the customers’ lifestyle and convenience are two key factors for accepting the APL solution (Tsai & Tiwasing, 2021). Besides APLs’ ease of use, which might crowd out less tech-savvy customers, one major variable driving the adoption of parcel lockers is location. Traditionally, parcel lockers are installed in controlled places that are easily accessible and close to areas with a high frequency of shipments (service stations, shopping malls, squares) (Janjevic et al., 2013). Using these sites lowers the entry barriers because they solve customer and package safety issues, often encountered in customer surveys (Lachapelle et al., 2018) and helps maximizing the APLs coverage area. Moreover, these sites can be reached by car on the route to or from customers’ homes, thus avoiding a dedicated pickup trip and increasing the convenience of the APL usage process.

Wider adoption of APLs in cities worldwide is expected due to the operational benefits of such a delivery option and factoring in the estimated continuous increase in online purchases. Hence, preferred locations nowadays might not be sufficient to cover all the future demand. Furthermore, understanding the implications of the APL network design on the potential customers’ demand is paramount for avoiding cost overrun while keeping the service level intact (Lin et al., 2022).

Discerning how the demand for APL will drive their diffusion on the urban territory requires an accurate ex-ante appraisal of the variables involved. To this end, this study aims to contribute to the literature by proposing a robust solution to the APL facility location problem. The solution involves simulating multiple instances of a mixed-integer linear CFLP with stochastic demand. The problem parameters are based on a survey submitted to e-consumers. Furthermore, the study strives towards practical applicability by proposing two greedy optimization algorithms that
identify a road map for opening APL locations. Finally, we show a real-life application of both methods to the city of Turin.

The paper is structured as follows. Section 2 reviews the existing literature on the APL location problem. Then, Section 3 explores the research approach and data collection methods. Numerical results are described in Section 4. Lastly, Section 5 provides practical and theoretical implications and draws upon conclusions.

2 Literature review

The APL location problem represents a niche, albeit thriving, stream of research within last-mile logistics. Lachapelle et al. (2018) assessed APL locations in light of customers’ preferences, such as safety, proximity to highways, and accessibility. Thus, the authors draw implications for future locations based on existing ones rather than through an optimization algorithm. Lee et al. (2019) adopted a two-step approach to the APL location problem. Firstly, they identified the potential locations following the preferred attributes by Lachapelle et al. (2018). Then, they used a set-covering model to estimate which potential locations should host an APL. However, the decision criteria for selecting the potential locations are unclear, and the numerical case study comprises a relatively small neighborhood. Deutsch & Golany (2018) proposed a more optimized approach via an un-capacitated facility location problem (UFLP). In their model, demand is lost when customers are not reached by any APL, and their willingness to move decreases with the distance to the parcel locker. A similar approach for the capacitated facility location problem (CFLP) is used by Che et al. (2022).

These works show that the APL location problem can be tackled with a long-term planning horizon. Recent studies looked at the impact of APL locations on a short-term planning horizon, integrating APLs into the Vehicle-Routing Problem (VRP). For example, Orenstein et al. (2019) optimized the assignment of parcels to both the vehicles and APL modules but assumed equal attractiveness of different APL locations for the customers, thus not considering any customers’ preference in the optimization algorithm. Instead, Enthoven et al. (2020) aimed to minimize both customers’ and LSP’s costs, including penalties if customers are not reached via their preferred delivery method. In Jiang et al. (2022), both LSPs and customers incur operating costs for visiting the APL location. The authors aimed to maximize the LSPs’ profits via a Traveling Salesman Problem (TSP) formulation.

The main objective of the APL location problem is to find the APLs locations that minimize the number of APLs to install, and the operational costs connected with traveling to those locations. Previous models considered the distance between the APL and the customer points as a static parameter. However, final customers do not all share the same appreciation and commitment for this last-mile solution. Thus, a location model should consider variability in the customers’ willingness to move from their residence to pick up the parcel and their demand.

Because of the customers’ preferences and the multiple delivery options they can now choose from, failing to cover a customer from an APL location can result in lost demand. Customers’ choices have been recently included in the APL location problem. The attractiveness of a location can be modeled as a decaying function of the distance between the location and the customer (Lin et al., 2022; Luo et al., 2022). Moreover, solutions to the APL facility location problem should be robust to deal with the variability of the demand.

Wang et al. (2020) adopted an integer-linear programming (ILP) model to solve the location problem of movable APLs with stochastic customers’ demand. In this study, all customers within an acceptable distance are assigned to a particular APL location. Robust results can be achieved by simulating multiple runs of an optimized facility location problem, as shown by Rabe et al. (2021), who adopted a Monte-Carlo simulation approach for a CFLP.
<table>
<thead>
<tr>
<th>Work</th>
<th>Candidate Location Set</th>
<th>Customers’ Willingness to Use Parcel Lockers</th>
<th>Capacitated</th>
<th>Objective Function</th>
<th>Demand Covered</th>
<th>Customers’ Demand Distribution</th>
<th>Solution Method</th>
<th>Numerical Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al. (2019)</td>
<td>Selected according to customers’ attributes</td>
<td>All customers within cover radius use the APL</td>
<td>Uncapacitated</td>
<td>Maximize covered demand</td>
<td>Output</td>
<td>Deterministic</td>
<td>Not defined</td>
<td>Real instance</td>
</tr>
<tr>
<td>Deutsch and Golany (2018)</td>
<td>One per each city subdivision</td>
<td>Decreases with the distance to the APL</td>
<td>Uncapacitated</td>
<td>Minimize cost</td>
<td>Variable</td>
<td>Deterministic</td>
<td>Mixed integer linear programming</td>
<td>Real instance</td>
</tr>
<tr>
<td>Che et al. (2022)</td>
<td>Same as customers’ locations</td>
<td>All customers within cover radius use the APL</td>
<td>Capacitated</td>
<td>Maximize covered demand, APL capacity usage</td>
<td>Output</td>
<td>Not defined</td>
<td>Genetic algorithm</td>
<td>Virtual instance</td>
</tr>
<tr>
<td>Lin et al. (2022)</td>
<td>Simulated through uniform distribution</td>
<td>All customers within cover radius use the APL</td>
<td>Uncapacitated</td>
<td>Maximize covered demand, profit</td>
<td>Output</td>
<td>Stochastic</td>
<td>Combinatorial optimization</td>
<td>Virtual instance</td>
</tr>
<tr>
<td>Luo et al. (2022)</td>
<td>Same as potential customers’ locations</td>
<td>Decreases with the distance to the APL</td>
<td>Flexible Size</td>
<td>Minimize cost, maximize covered demand</td>
<td>Output</td>
<td>Stochastic</td>
<td>Evolutionary algorithm</td>
<td>Virtual and real instances</td>
</tr>
<tr>
<td>Wang et al. (2020)</td>
<td>Same as potential customers’ locations</td>
<td>All customers within cover radius use the APL</td>
<td>Capacitated</td>
<td>Minimize cost</td>
<td>Constraint</td>
<td>Stochastic</td>
<td>Integer linear programming</td>
<td>Virtual instance</td>
</tr>
<tr>
<td>Rabe et al. (2021)</td>
<td>One per each city subdivision</td>
<td>According to APL service level and accessibility</td>
<td>Capacitated</td>
<td>Minimize cost</td>
<td>Variable</td>
<td>Stochastic</td>
<td>Not defined</td>
<td>Real instance</td>
</tr>
<tr>
<td>This study</td>
<td>One per each city subdivision</td>
<td>All customers within cover radius use the APL</td>
<td>Flexible Size</td>
<td>Minimize cost</td>
<td>Constraint</td>
<td>Stochastic</td>
<td>Greedy algorithm</td>
<td>Real instance</td>
</tr>
</tbody>
</table>

Table 1. Review of the APL facility location literature
Theoretical applications of the APL facility location problem should be easy to implement in a real-life scenario by LSPs and practitioners alike. However, most APL network design solutions are applied to virtual instances rather than real urban settings, except for Deutsch & Golany (2018), Luo et al. (2022), and Rabe et al. (2021).

This study proposes novel algorithms to solve the capacitated APL facility location problem. The algorithms aim to minimize the APL network cost while satisfying a predetermined demand share. The study applies the proposed method to a real instance assuming stochastic customers’ demand.

3 Research methodology

This study presents two methods for determining the optimal number, location, and capacity of APLs to install. The first method involves optimizing a mixed-integer linear programming (MILP) model (Model 1), simultaneously determining the aforementioned variables. In contrast, the second method employs greedy algorithms (Algorithms 1 and 2) that iteratively select the next APL and determine its capacity.

Both methods entail running multiple simulations to ensure robust results. In each simulation, Procedure 1 randomizes the users’ variables.

Procedure 1. Randomize Users

```python
def RandomizeUsers():
    for u ← 1 to nUsers do
        area_u ← f^{-1}[0,1]
        x_u←U(x_{area_u}min, x_{area_u}max)
        y_u←U(y_{area_u}min, y_{area_u}max)
        dem_u ← g^{-1}[0,1]
        tol_u ← h^{-1}[0,1]
        sat_u ← 0
        if APLs locations are not fixed then
            return x, y, sat, tol, dem
        else if APLs locations are fixed then
            for a ← 1 to nAPLs do
                D_{ua} ← [(x_u - x_a)^2 + (y_u - y_a)^2]^{1/2}
            return x, y, sat, tol, dem
```

For all users (for u ← 1 to nUsers), the procedure randomizes their location (x_u, y_u), the maximum distance they are willing to travel to reach the APL (tol_u), their daily demand (dem_u), and sets the users as not served by any facility (sat_u ← 0).

A user location is represented in Cartesian coordinates (x_u, y_u), with the meter being the unit of measurement, and is randomized in two steps. The procedure first randomizes the area to which the user belongs (area_u) by following the empirical distribution (f), which can be determined based on the city under analysis. Then, the procedure randomizes the user coordinates through a uniform distribution within the area limits (i.e., x_{area_u}min, x_{area_u}max, y_{area_u}min, and y_{area_u}max).

Concerning the APLs, if their locations are not predetermined, the procedure evaluates the Euclidean distance (D_{ua}) between the u-th user and all the APLs (for a ← 1 to nAPLs).

3.1 MILP model

This subsection introduces the equations that make up the complete MILP model. Then, it provides a simplified version of the MILP model to allow faster simulations without compromising the reliability of the results.
Complete model

The complete MILP model consists of Equations 1-20.

\[
\text{minimize} \quad c^{\text{APL}} \sum_{a \in A} \text{act}_a + c^{\text{Module}} \sum_{a \in A} \text{mod}_a \\
\text{subject to} \quad \sum_{u \in U} \text{sat}_u \cdot \text{dem}_u / \sum_{u \in U} \text{dem}_u \geq \text{SL} \\
D^{\text{ua}} = D_{\text{ua}} \cdot \text{act}_a + M(1 - \text{act}_a) \quad \forall u \in U, \forall a \in A \\
D_{\text{ua}} = [(x_u - x_a)^2 + (y_u - y_a)^2]^{1/2} \quad \forall u \in U, \forall a \in A \\
mD_u \leq D^{\text{ua}} \quad \forall u \in U, \forall a \in A \\
mD_u \geq D^{\text{ua}} - M(1 - \text{ass}_{\text{ua}}) \quad \forall u \in U, \forall a \in A \\
\sum_{a \in A} \text{ass}_{\text{ua}} = 1 \quad \forall u \in U \\
tol_u - mD_u + \epsilon - M \cdot \text{sat}_u \leq 0 \quad \forall u \in U \\
mD_u - \text{tol}_u - M(1 - \text{sat}_u) \leq 0 \quad \forall u \in U \\
\sum_{u \in U} \text{ass}_{\text{ua}} \cdot \text{dem}_u \leq \text{cap}_{\text{APL}} + \text{cap}_{\text{Module}} \cdot \text{mod}_a \quad \forall a \in A \\
\text{mod}_a \leq \text{mod}_{\text{max}} \cdot \text{act}_a \quad \forall a \in A \\
\sum_{a \in A} \text{act}_{\text{ua}} \leq \max_{\text{a} \in A} \quad \forall a \in A \\
\text{act}_a \in \{0, 1\} \quad \forall a \in A \\
\text{sat}_u \in \{0, 1\} \quad \forall u \in U \\
\text{ass}_{\text{ua}} \in \{0, 1\} \quad \forall u \in U, \forall a \in A \\
\text{mod}_a \in \mathbb{N} \quad \forall a \in A \\
\text{D}_{\text{ua}} \in \mathbb{R}^+ \quad \forall u \in U, \forall a \in A \\
\text{D}_{\text{ua}}^* \in \mathbb{R}^+ \quad \forall u \in U, \forall a \in A \\
mD_u \in \mathbb{R}^+ \quad \forall u \in U, \forall a \in A \\
x_u, y_u \in \mathbb{R}^+ \quad \forall a \in A
\]

As in Procedure 1, the subscript \(a\) denotes the \(a\)th APL facility, while the subscript \(u\) refers to the \(u\)th user. In each simulation, users’ variables are treated as parameters.

Equation 1 is the objective function to minimize and consists of the sum of the total cost of the APLs \(c^{\text{APL}} \sum_{a \in A} \text{act}_a\) and the total cost of the additional modules \(c^{\text{Module}} \sum_{a \in A} \text{mod}_a\). The variable \(\text{act}_a\) is Boolean (as per Equation 13) and indicates whether the \(a\)th APL is installed \((\text{act}_a = 1)\) or not \((\text{act}_a = 0)\). Instead, the variable \(\text{mod}_a\) is natural (as per Equation 16) and refers to the number of modules to install in the \(a\)th APL. Equation 11 limits the maximum number of modules per APL to \(\text{mod}_{\text{max}}\) due to volumetric constraints and prevents any module from being installed if an APL is not active \((\text{act}_a = 0)\). To ensure that the optimization routine prefers expanding existing APLs by installing additional modules over activating new APLs, the cost of installing an APL \(c^{\text{APL}}\) must be greater than the cost of an installing an additional module \(c^{\text{Module}}\).

Equation 2 expresses the service level constraint, which ensures that the percentage of expected satisfied demand \(\sum_{u \in U} \text{sat}_u \cdot \text{dem}_u / \sum_{u \in U} \text{dem}_u\) meets or exceeds the target service level \(\text{SL}\). The variable \(\text{sat}_u\) is also Boolean (as per Equation 14) and determines whether the \(u\)th user is considered covered \((\text{sat}_u = 1)\) when its demand \((\text{dem}_u)\) is assigned to any APL.

Equations 3-7 pertain to the distances \(D_{\text{ua}}\) (Equation 17) and minimum distances \(mD_u\) (Equation 18) between the users and the APLs. Equation 4 computes the Euclidean distance \((D_{\text{ua}})\) between the users and the APLs. Equation 3 prevents an APL from being assigned to any user if the APL is not active \((\text{act}_a = 0)\), as the distance between them would be set to \(M\). Equations 5-6 determine the
distance between the $u$th user and its nearest active APL ($mD_u$), while Equation 7 assigns the user’s demand ($dem_u$) to that APL ($ass_{ua} = 1$). Equations 10 and 12 relate to calculating the expected demand to be satisfied by each APL. Specifically, Equation 10 ensures that no APL is assigned more demand than its capacity, which is the sum of its standard capacity ($c^{APL}$) and the capacity of each additional module ($cap^{\text{Module}} \cdot mod_a$). Equation 11 fixes the maximum number of modules to be installed to 10, which is reasonable given the amount of space each module requires for its location, while Equation 12 prevents the module from being installed ($mod_a$) to an APL facility whenever such APL is not activated ($act_a = 0$). Equations 8-9 ensure that whenever the distance between a user and its nearest APL ($mD_u$) is less than or equal to the user’s tolerated travel distance ($tol_u$), the user is assigned to that APL (Equation 7), and its demand is considered satisfied ($sat_u = 1$). While multiple APLs may serve one user’s demand, this assumption simplifies the problem and reduces computation time. This formulation assigns a user’s demand to its nearest APL.

**Simplified model**

The simplified MILP model (Equations 21-35) shows key differences from the complete model, summarized as follows.

\[
\begin{align*}
\text{minimize} & \quad (mod^{\text{max}} \cdot c^{\text{Module}} + c^{\text{Module}}) \sum_{a \in A} act_a + c^{\text{Module}} \sum_{a \in A} mod_a \\
\text{subject to} & \quad \sum_{u \in U} sat_u/nUsers \geq SL \\
& \quad D_{ua}^* = D_{ua} \cdot act_a + M(1 - act_a) \quad \forall u \in U, \forall a \in A \\
& \quad mD_u \leq D_{ua} \quad \forall u \in U, \forall a \in A \\
& \quad mD_u \geq D_{ua} - M(1 - ass_{ua}) \quad \forall u \in U, \forall a \in A \\
& \quad \sum_{a \in A} ass_{ua} = 1 \quad \forall u \in U \\
& \quad tol_u - mD_u + \epsilon - M \cdot sat_u \leq 0 \quad \forall u \in U \\
& \quad mD_u - tol_u - M(1 - sat_u) \leq 0 \quad \forall u \in U \\
& \quad \sum_{u \in U} ass_{ua} \cdot SF \cdot \bar{dem} \leq cap^{APL} + cap^{\text{Module}} \cdot mod_a \quad \forall a \in A \\
& \quad mod_a \leq mod^{\text{max}} \cdot act_a \quad \forall a \in A \\
& \quad act_a \in \{0, 1\} \quad \forall a \in A \\
& \quad sat_u \in \{0, 1\} \quad \forall u \in U \\
& \quad ass_{ua} \in \{0, 1\} \quad \forall u \in U, \forall a \in A \\
& \quad mod_a \in \mathbb{N} \quad \forall a \in A \\
& \quad mD_u \in \mathbb{R}^+ \quad \forall u \in U
\end{align*}
\]

Firstly, the maximum number of APLs ($max_{a \in A}$) is limited to the number of postcodes used by postal and logistics providers (mAPLs). To achieve this, each postcode area is mapped onto a coordinate system, and the APL locations are fixed to the centers of gravity of the respective postcode area. The relevant postcode data can be found in Table 3.

Secondly, the users are aggregated into fewer demand collection points, and, to account for the same demand, a scaling factor (SF) is introduced.

Thirdly, the cost of each APL is set to $mod^{\text{max}} \cdot c^{\text{Module}} + c^{\text{Module}}$. These values guide the optimization routine to prefer saturating APLs, when possible, instead of activating new ones.

Finally, the model uses the mean values of demand ($\bar{dem}$) and tolerated travel distance ($tol$) for each demand point.
For the first assumption, assuming an average demand approach is a justifiable measure for two key reasons. Firstly, the literature on the impact of socio-demographic factors on e-commerce consumer behavior has presented conflicting findings (Agudo-Peregrina et al., 2016). Secondly, the COVID-19 pandemic and the ongoing trend towards digitalization have led to a narrowing in the behavior of e-commerce consumers, irrespective of their socio-demographic characteristics.

The second assumption stems from the survey analyses presented in previous papers (Mitrea et al., 2020; Ottaviani et al., 2020). The studies found that the average tolerated distance does not depend on socio-demographic factors. Tolerance is a highly variable factor, both from individual to individual and from the same individual. In other words, the tolerated travel distance is a function of the means the individual uses to collect their order and how long they are willing to divert their daily commute to collect their order. These factors could vary depending on the user, the current traffic, the value of the order content, the urgency to collect the order, and other factors unrelated to the socio-demographic characteristics.

For these reasons, using mean values is consistent with the assumption of homogeneity of the population within each area. This approach allows for a simpler, computationally feasible model while still capturing the essential aspects of the problem. However, using mean values may result in some loss of information and potential inaccuracies, especially if the actual demands or travel distances vary significantly within the area.

3.2 Algorithms

This study presents two greedy algorithms that determine the number of APLs, their location, and their capacity. Doing so allows for evaluating the expected satisfied customers’ demand curve depending on the number of APLs. In contrast, the MILP formulation does not provide a sequence of APLs to be installed because it determines the optimal solution by solving a mathematical optimization problem that considers all possible scenarios.

These MILP and greedy algorithms approaches have different advantages and disadvantages. The MILP model provides an optimal solution but has high computational complexity. Instead, the greedy algorithms are computationally efficient but may not guarantee an optimal solution. However, both algorithms implement the following procedures, which are consistent with findings from Luo et al. (2022) and Wang et al. (2020).

Procedure 2 loads the previous APLs. It first takes as input variables the ordered sequence of APLs (APL_order), the next APL candidate (APL*), the array of satisfied users (sat), and the user/APL distances matrix (D). Then, it evaluates the potential number of users (pot_users) and respective demand (pot_demand) that each APL in the ordered sequence can cover. Next, the procedure determines the number of modules (mod_a) that each APL can accommodate and assigns users accordingly, up to mod_{max}. Finally, it marks the users as covered by their assigned APL (sat_u ← 1).

### Procedure 2. Load Previous APLs

```python
def LoadPreviousAPLs(APL_order, sat, D, APL*):
    for a ∈ APL_order do
        if a ≠ APL* then
            pot_users ← 0
            pot_dem ← 0
        for u ← 1 to nUsers do
            if sat_u = 0 and D_{au} ≤ tol then
                pot_users ← pot_users + 1
                pot_dem ← pot_dem + tol
            mod_a ← 0
            cap_a ← capAPL
            t1 ← 0
            while t1 ≤ mod_{max} and pot_users > 0 do
                if t1 > 0 then
```

31

Ottaviani, Zenezini, De Marco & Carlin
Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal
Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal

Procedure 3 evaluates the next APL candidate ($APL^*$). To accomplish this, it takes the ordered sequence of APLs ($APL_{order}$), the array of satisfied users ($sat$), and the user/APL distances matrix ($D$) as input variables. The procedure then loops through the APLs that are not already in $APL_{order}$ and selects the APL that could cover the highest potential number of users or demand.

```python
def EvaluateNextAPL(APL_order, sat, D):
    APL* ← 0
    max_dem ← 0
    for a ← 1 to nAPLs do
        if a ∉ APL_order then
            pot_dem ← 0
            for u ← 1 to nUsers do
                if sat_u = 0 and D_ua ≤ tol then
                    pot_dem ← pot_dem + 1
            if pot_dem ≥ max_dem then
                max_dem ← pot_dem
                APL* ← a
    return APL*
```

Procedure 4 evaluates the potential demand of the set APL, increases its modules if they do not reach their maximum ($mod_{max}$), and sets all covered users as satisfied ($sat_u = 1$).

```python
def LoadNextAPL(sat, D, APL*):
    pot_users ← 0
    pot_dem ← 0
    for u ← 1 to nUsers do
        if sat_u = 0 and D_ua ≤ tol then
            pot_users ← pot_users + 1
            pot_dem ← pot_dem + 1
            mod_a ← 0
            it₁ ← 0
            while it₁ ≤ mod_{max} and pot_users > 0 do
                if it₁ > 0 then
                    mod_a ← mod_a + 1
                    cap_a ← cap_a + cap_{Module}
                if min(cap_a, pot_dem) = cap_a then
                    sat_temp ← cap_a/dem
                else
                    sat_temp ← pot_users/dem
                for it₂ ← 0 to sat_temp do
                    for u ← 0 to nUsers do
                        if sat_u = 0 and D_ua ≤ tol then
                            sat_u ← 1
return sat
```

The procedure then loops through the APLs that are not already in $APL_{order}$ and selects the APL that could cover the highest potential number of users or demand.
Ottaviani, Zenezini, De Marco & Carlin
Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal

Algorithm 1

Algorithm 1 involves two main steps. In the first step (A1S1), the users are randomized S times, and the sequence of APLs is determined (\textit{APL\_order}). The percentage of covered demand is evaluated for each APL in the sequence. Once A1S1 is completed, the APLs are sorted based on the average percentage of covered demand across all S simulations, thus providing an ordered sequence of APLs for the second step (A1S2). The second step consists of another round of S simulations, which allows for the determination of the average covered demand by each APL (\overline{pc\_ts_a}).

Algorithm 1. Sequence-provided algorithm

| Data: x, y |
| Result: APLs location and capacity evaluation given fixed sequence initialization |

/* Step 1 (Facultative) */

\textbf{init:} pc\_ts

\textbf{for} s ← 1 to S \textbf{do}

\textbf{APL} ← 0

x, y, sat, tol, dem, D ← RandomizeUsers()

\textbf{for} a ← 1 to nAPLs \textbf{do}

sat ← LoadPreviousAPLs(APL\_order, sat, D, APL)

APL ← EvaluateNextAPL(APL\_order, sat, D)

sat, nSats ← LoadNextAPL(sat, D, APL)

pc\_ts\_s\_APL ← nSats/nUsers

\textbf{for} \textbf{a} ∈ APL\_order \textbf{do}

\ \overline{pc\_ts_a} ← \sum_{s=1}^{S} pc\_ts\_s\_APL/S

Sort APLs by descending order of \overline{pc\_ts_a} OR provide APLs sequence manually

/* Step 2 (Mandatory) */

\textbf{init:} pc\_ts

\textbf{for} s ← 1 to S \textbf{do}

D, sat ← RandomizeUsers()

\textbf{for} a ∈ APL\_order \textbf{do}

APL ← a

sat, nSats ← LoadNextAPL(sat, D, APL)

pc\_ts\_s\_APL ← nSats/nUsers

\textbf{for} \textbf{a} ∈ APL\_order \textbf{do}

\ \overline{pc\_ts_a} ← \sum_{s=1}^{S} pc\_ts\_s\_APL/S

Sort APLs by descending order of \overline{pc\_ts_a}

Algorithm 2

The output of Algorithm 2 consists of a sequence of APLs, evaluated as follows. Firstly, S simulations are conducted nAPLs times to evaluate and select the next APL to install. At each step, the algorithm considers the presence of previously installed APLs that already cover a portion of users. Once the best APL has been selected most often out of the S simulations, another round of S simulations is conducted to calculate its mean expected covered demand. This process is repeated for nAPLs times, resulting in a unique sequence of APLs expected to provide the highest level of demand coverage.

Algorithm 2. Sequence-evaluation algorithm

| Data: x, y |
| Result: APLs location and capacity evaluation with greedy evaluation sequence initialization |
init: \( APL\_\text{order}, pcts \)

\[ \text{for } it \leftarrow 0 \text{ to } \text{nAPLs do} \]

init: \( APL\_\text{maxs}, pctss \)

\[ APL^* \leftarrow 0 \]

\[ \text{for } s \leftarrow 1 \text{ to } S \text{ do} \]

\[ x, y, sat, tol, dem, D \leftarrow \text{RandomizeUsers()} \]

\[ sat \leftarrow \text{LoadPreviousAPLs}(APL\_\text{order}, sat, D, APL^*) \]

\[ APL^* \leftarrow \text{EvaluateNextAPL}(APL\_\text{order}, sat, D) \]

\[ APL\_\text{maxs}_s \leftarrow APL^* \]

\[ APL\_\text{order}_{it} \leftarrow APL^* \]

\[ \text{for } s \leftarrow 1 \text{ to } S \text{ do} \]

\[ x, y, sat, tol, dem, D \leftarrow \text{RandomizeUsers()} \]

\[ sat, nSats \leftarrow \text{LoadNextAPL}(sat, D, APL^*) \]

\[ pcts_s \leftarrow nSats/nUsers \]

\[ \text{for } s \leftarrow 1 \text{ to } S \text{ do} \]

\[ \text{for } i \leftarrow 1 \text{ to } \text{nAPLs do} \]

\[ \text{for } s \leftarrow 1 \text{ to } S \text{ do} \]

\[ pcts_{it} \leftarrow \sum_{s=1}^{S} pcts_s / S \]

4 Numerical experiments

This section presents the results of the simplified MILP model and Algorithms 1 and 2 applied to the case study under analysis.

Equation 36 provides the empirical distribution of the user population, where \( z \sim U(0,1) \).

\[
f(z) = \begin{cases} 
1 & z \leq .1548 \\
2 & .1548 < z \leq .2999 \\
3 & .2999 < z \leq .4416 \\
4 & .4416 < z \leq .5825 \\
5 & .5825 < z \leq .7031 \\
6 & .7031 < z < .8129 \\
7 & .8129 < z \leq .9101 \\
8 & z > .9101 
\end{cases} \quad (36) 
\]

The empirical distribution was based on the data collected from the city municipality office, as of Table 2 (Comune di Torino - Ufficio di Statistica, n.d.).

<table>
<thead>
<tr>
<th>Area</th>
<th>( x^\text{min} )</th>
<th>( x^\text{max} )</th>
<th>( y^\text{min} )</th>
<th>( y^\text{max} )</th>
<th>( pct )</th>
<th>( pct_\text{cum} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4200</td>
<td>-500</td>
<td>-6300</td>
<td>-1800</td>
<td>.1548</td>
<td>.1548</td>
</tr>
<tr>
<td>2</td>
<td>-1100</td>
<td>2200</td>
<td>-5300</td>
<td>1600</td>
<td>.1451</td>
<td>.2999</td>
</tr>
<tr>
<td>3</td>
<td>-6000</td>
<td>-1600</td>
<td>-2300</td>
<td>500</td>
<td>.1417</td>
<td>.4416</td>
</tr>
<tr>
<td>4</td>
<td>-6000</td>
<td>-1600</td>
<td>2500</td>
<td>5200</td>
<td>.1409</td>
<td>.5825</td>
</tr>
<tr>
<td>5</td>
<td>-1600</td>
<td>1600</td>
<td>2800</td>
<td>5000</td>
<td>.1206</td>
<td>.7031</td>
</tr>
<tr>
<td>6</td>
<td>-6000</td>
<td>-1600</td>
<td>-500</td>
<td>1600</td>
<td>.1098</td>
<td>.8129</td>
</tr>
<tr>
<td>7</td>
<td>-1600</td>
<td>3000</td>
<td>1600</td>
<td>3000</td>
<td>.0972</td>
<td>.9101</td>
</tr>
<tr>
<td>8</td>
<td>-1500</td>
<td>500</td>
<td>-2200</td>
<td>1600</td>
<td>.0899</td>
<td>1</td>
</tr>
</tbody>
</table>

The demand data used were collected via a survey submitted to e-commerce users. The survey results, presented in Mitrea et al. (2020) and Ottaviani et al. (2020), were used to estimate the model parameters, namely the overall APL daily demand and the average tolerated distance by the APL users. The APL daily demand (\( dem_u \)) is randomized through the exponential distribution with shape parameter \( \lambda = \text{tol} = 0.027 \) (Ottaviani et al., 2020) as a function of the share of APL potential users and the overall demand for parcels by e-commerce users (Fig. 1).
Figure 1.  APL daily demand probability density function

The users’ tolerated travel distance ($tol_u$) is randomized through the beta distribution, extended to the domain $[0, 6670]$, with shape parameters $\alpha = 1.23$ and $\beta = 3.27$ (Ottaviani et al., 2020). The distribution is portrayed in Figure 2, where $1800$ indicates the mean $\bar{tol} = 6670 \cdot \alpha/(\alpha + \beta) = 1.23/(1.23 + 3.27) \sim 6670 \cdot 0.27 \sim 1800$.

Figure 2.  Tolerated travel distance probability density function

For simplicity, the APL network is assumed to be able to serve the entire city demand. This assumption would equate to a scenario where a single LSP delivers all parcels to APL customers. Hence, to set the estimated number of e-commerce users, the maximum number of users surveyed is multiplied by the percentage of the whole population based on the survey, resulting in $n_{Users} = 102000$. We then aggregated the number of users to 1020 demand collection points, where each point accounts for 1000 users.

The number of potential APL locations was determined based on the number of postcodes in the urban area of Turin, which resulted in a maximum of $n_{APLs} = 33$ APL locations.

Table 3.  Areas bounds, inhabitants’ percentages, and inhabitants’ cumulative percentages

<table>
<thead>
<tr>
<th>Postcode</th>
<th>$x_{min}$</th>
<th>$x_{max}$</th>
<th>$y_{min}$</th>
<th>$y_{max}$</th>
<th>$x_a$</th>
<th>$y_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10121</td>
<td>-1450</td>
<td>0</td>
<td>0</td>
<td>700</td>
<td>-725</td>
<td>350</td>
</tr>
<tr>
<td>10122</td>
<td>-1450</td>
<td>0</td>
<td>700</td>
<td>1570</td>
<td>-725</td>
<td>1135</td>
</tr>
<tr>
<td>10123</td>
<td>0</td>
<td>1150</td>
<td>0</td>
<td>1000</td>
<td>575</td>
<td>500</td>
</tr>
<tr>
<td>10124</td>
<td>0</td>
<td>1600</td>
<td>980</td>
<td>1600</td>
<td>800</td>
<td>1290</td>
</tr>
<tr>
<td>10125</td>
<td>0</td>
<td>780</td>
<td>-1000</td>
<td>0</td>
<td>390</td>
<td>-50</td>
</tr>
<tr>
<td>10126</td>
<td>0</td>
<td>1200</td>
<td>-4000</td>
<td>-1000</td>
<td>600</td>
<td>-2500</td>
</tr>
<tr>
<td>10127</td>
<td>-1000</td>
<td>900</td>
<td>-6000</td>
<td>-4000</td>
<td>-50</td>
<td>-5000</td>
</tr>
<tr>
<td>10128</td>
<td>-700</td>
<td>0</td>
<td>-1900</td>
<td>0</td>
<td>-350</td>
<td>-950</td>
</tr>
<tr>
<td>10129</td>
<td>-1400</td>
<td>-700</td>
<td>-1950</td>
<td>-280</td>
<td>-1050</td>
<td>-1115</td>
</tr>
<tr>
<td>10131</td>
<td>1200</td>
<td>5000</td>
<td>-1400</td>
<td>1500</td>
<td>3100</td>
<td>50</td>
</tr>
<tr>
<td>10132</td>
<td>1800</td>
<td>4000</td>
<td>1250</td>
<td>5000</td>
<td>2900</td>
<td>3125</td>
</tr>
<tr>
<td>10133</td>
<td>1100</td>
<td>5000</td>
<td>-4500</td>
<td>0</td>
<td>3050</td>
<td>-2250</td>
</tr>
<tr>
<td>10134</td>
<td>-1100</td>
<td>0</td>
<td>-4100</td>
<td>-1800</td>
<td>-550</td>
<td>-2950</td>
</tr>
</tbody>
</table>
As the population distribution for the different postcodes was unavailable, the demand collection points were assumed to be uniformly distributed across the 33 postcodes.

Figure 3 displays eight city areas and the APLs predetermined locations.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10135</td>
<td>-3600</td>
<td>0</td>
<td>-6800</td>
<td>-4100</td>
<td>-1800</td>
<td>-5450</td>
</tr>
<tr>
<td>10136</td>
<td>-3600</td>
<td>-1100</td>
<td>-3200</td>
<td>-2000</td>
<td>-2350</td>
<td>-2600</td>
</tr>
<tr>
<td>10137</td>
<td>-3300</td>
<td>-1100</td>
<td>-4500</td>
<td>-2900</td>
<td>-2200</td>
<td>-3700</td>
</tr>
<tr>
<td>10138</td>
<td>-3100</td>
<td>-1400</td>
<td>-800</td>
<td>1000</td>
<td>-2250</td>
<td>100</td>
</tr>
<tr>
<td>10139</td>
<td>-4400</td>
<td>-2200</td>
<td>-900</td>
<td>0</td>
<td>-3300</td>
<td>-450</td>
</tr>
<tr>
<td>10141</td>
<td>-4800</td>
<td>-1400</td>
<td>-2300</td>
<td>-900</td>
<td>-3100</td>
<td>-1600</td>
</tr>
<tr>
<td>10142</td>
<td>-5700</td>
<td>-3200</td>
<td>-2500</td>
<td>-900</td>
<td>-4450</td>
<td>-1700</td>
</tr>
<tr>
<td>10143</td>
<td>-3100</td>
<td>-1400</td>
<td>0</td>
<td>1500</td>
<td>-2250</td>
<td>750</td>
</tr>
<tr>
<td>10144</td>
<td>-2700</td>
<td>-1400</td>
<td>1000</td>
<td>2600</td>
<td>-2050</td>
<td>1800</td>
</tr>
<tr>
<td>10145</td>
<td>-3800</td>
<td>-3100</td>
<td>-200</td>
<td>1200</td>
<td>-3450</td>
<td>500</td>
</tr>
<tr>
<td>10146</td>
<td>-5900</td>
<td>-3700</td>
<td>-1000</td>
<td>400</td>
<td>-4800</td>
<td>-300</td>
</tr>
<tr>
<td>10147</td>
<td>-3100</td>
<td>-1400</td>
<td>2800</td>
<td>4600</td>
<td>-2250</td>
<td>3700</td>
</tr>
<tr>
<td>10148</td>
<td>-4800</td>
<td>-1400</td>
<td>3700</td>
<td>5700</td>
<td>-3100</td>
<td>4700</td>
</tr>
<tr>
<td>10149</td>
<td>-3800</td>
<td>-1400</td>
<td>1800</td>
<td>3600</td>
<td>-2600</td>
<td>2700</td>
</tr>
<tr>
<td>10151</td>
<td>-6200</td>
<td>-3100</td>
<td>1500</td>
<td>4000</td>
<td>-4650</td>
<td>2750</td>
</tr>
<tr>
<td>10152</td>
<td>-1400</td>
<td>600</td>
<td>1500</td>
<td>3000</td>
<td>-400</td>
<td>2250</td>
</tr>
<tr>
<td>10153</td>
<td>0</td>
<td>2600</td>
<td>1500</td>
<td>3000</td>
<td>1300</td>
<td>2250</td>
</tr>
<tr>
<td>10154</td>
<td>-400</td>
<td>1200</td>
<td>2400</td>
<td>4800</td>
<td>400</td>
<td>3600</td>
</tr>
<tr>
<td>10155</td>
<td>-1400</td>
<td>-400</td>
<td>2000</td>
<td>5600</td>
<td>-900</td>
<td>3800</td>
</tr>
<tr>
<td>10156</td>
<td>-400</td>
<td>2500</td>
<td>5000</td>
<td>8000</td>
<td>1050</td>
<td>6500</td>
</tr>
</tbody>
</table>
4.1 Fixed scenario

We simulated the simplified MILP model 1000 times using the FICO® Xpress v8.13.0 optimization library in Python 3.9.12. Table 4 presents the optimization results, which took approximately 1 minute per run. On the other hand, the algorithms were executed using Julia 1.8.5. We set the total number of simulations to 10000 in A1S1-A1S2 and MILP-A1S2, while we set it to 1000. A2 took approximately 2 minutes to execute. Instead, A1S1-A2S2 and MILP-A1S2 took approximately 2 minutes and 20 seconds, respectively. We ran all scripts on an Intel® Core™ i7-10750H CPU.
Table 4. Areas bounds, inhabitants’ percentages, and inhabitants’ cumulative percentages

<table>
<thead>
<tr>
<th>APL</th>
<th>(x_a)</th>
<th>(y_a)</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>-725</td>
<td>350</td>
<td>1</td>
<td>5.89</td>
<td>173.00</td>
<td>.946</td>
</tr>
<tr>
<td>28</td>
<td>-725</td>
<td>1135</td>
<td>1</td>
<td>2.55</td>
<td>106.58</td>
<td>.919</td>
</tr>
<tr>
<td>6</td>
<td>575</td>
<td>500</td>
<td>.96</td>
<td>2.41</td>
<td>101.11</td>
<td>.893</td>
</tr>
<tr>
<td>14</td>
<td>800</td>
<td>1290</td>
<td>.95</td>
<td>1.73</td>
<td>87.36</td>
<td>.877</td>
</tr>
<tr>
<td>31</td>
<td>390</td>
<td>-500</td>
<td>.95</td>
<td>5.32</td>
<td>157.79</td>
<td>.921</td>
</tr>
<tr>
<td>30</td>
<td>600</td>
<td>-2500</td>
<td>.83</td>
<td>2.80</td>
<td>101.76</td>
<td>.841</td>
</tr>
<tr>
<td>7</td>
<td>-50</td>
<td>-5000</td>
<td>.79</td>
<td>.40</td>
<td>50.82</td>
<td>.696</td>
</tr>
<tr>
<td>26</td>
<td>-350</td>
<td>-950</td>
<td>.75</td>
<td>1.54</td>
<td>72.48</td>
<td>.757</td>
</tr>
<tr>
<td>5</td>
<td>-1050</td>
<td>-1115</td>
<td>.61</td>
<td>2.58</td>
<td>86.17</td>
<td>.739</td>
</tr>
<tr>
<td>15</td>
<td>3100</td>
<td>50</td>
<td>.60</td>
<td>2.80</td>
<td>89.24</td>
<td>.733</td>
</tr>
<tr>
<td>22</td>
<td>2900</td>
<td>3125</td>
<td>.48</td>
<td>1.67</td>
<td>58.98</td>
<td>.599</td>
</tr>
<tr>
<td>21</td>
<td>3050</td>
<td>-2250</td>
<td>.46</td>
<td>2.14</td>
<td>68.17</td>
<td>.632</td>
</tr>
<tr>
<td>16</td>
<td>-550</td>
<td>-2950</td>
<td>.45</td>
<td>1.12</td>
<td>48.08</td>
<td>.550</td>
</tr>
<tr>
<td>25</td>
<td>-1800</td>
<td>-5450</td>
<td>.44</td>
<td>1.69</td>
<td>58.02</td>
<td>.587</td>
</tr>
<tr>
<td>20</td>
<td>-2350</td>
<td>-2600</td>
<td>.32</td>
<td>.55</td>
<td>28.74</td>
<td>.378</td>
</tr>
<tr>
<td>19</td>
<td>-2200</td>
<td>-3700</td>
<td>.30</td>
<td>1.26</td>
<td>42.22</td>
<td>.468</td>
</tr>
<tr>
<td>3</td>
<td>-2250</td>
<td>100</td>
<td>.25</td>
<td>.81</td>
<td>30.17</td>
<td>.372</td>
</tr>
<tr>
<td>11</td>
<td>-3300</td>
<td>-450</td>
<td>.15</td>
<td>.00</td>
<td>6.31</td>
<td>.097</td>
</tr>
<tr>
<td>9</td>
<td>-3100</td>
<td>-1600</td>
<td>.15</td>
<td>.59</td>
<td>20.20</td>
<td>.263</td>
</tr>
<tr>
<td>17</td>
<td>-4450</td>
<td>-1700</td>
<td>.15</td>
<td>.82</td>
<td>24.43</td>
<td>.300</td>
</tr>
<tr>
<td>1</td>
<td>-2250</td>
<td>750</td>
<td>.14</td>
<td>.62</td>
<td>20.29</td>
<td>.262</td>
</tr>
<tr>
<td>32</td>
<td>-2050</td>
<td>1800</td>
<td>.12</td>
<td>.52</td>
<td>17.38</td>
<td>.230</td>
</tr>
<tr>
<td>8</td>
<td>-3450</td>
<td>500</td>
<td>.08</td>
<td>.36</td>
<td>11.66</td>
<td>.162</td>
</tr>
<tr>
<td>4</td>
<td>-4800</td>
<td>-300</td>
<td>.06</td>
<td>.18</td>
<td>7.05</td>
<td>.103</td>
</tr>
<tr>
<td>18</td>
<td>-2250</td>
<td>3700</td>
<td>.04</td>
<td>.16</td>
<td>5.42</td>
<td>.079</td>
</tr>
<tr>
<td>10</td>
<td>-3100</td>
<td>4700</td>
<td>.04</td>
<td>.00</td>
<td>0.77</td>
<td>.012</td>
</tr>
<tr>
<td>29</td>
<td>-2600</td>
<td>2700</td>
<td>.04</td>
<td>.12</td>
<td>4.76</td>
<td>.071</td>
</tr>
<tr>
<td>13</td>
<td>-4650</td>
<td>2750</td>
<td>.03</td>
<td>.03</td>
<td>2.26</td>
<td>.034</td>
</tr>
<tr>
<td>27</td>
<td>-400</td>
<td>2250</td>
<td>.03</td>
<td>.00</td>
<td>1.64</td>
<td>.025</td>
</tr>
<tr>
<td>23</td>
<td>1300</td>
<td>2250</td>
<td>.03</td>
<td>.12</td>
<td>4.08</td>
<td>.060</td>
</tr>
<tr>
<td>2</td>
<td>400</td>
<td>3600</td>
<td>.03</td>
<td>.08</td>
<td>3.18</td>
<td>.048</td>
</tr>
<tr>
<td>12</td>
<td>-900</td>
<td>3800</td>
<td>.02</td>
<td>.00</td>
<td>0.60</td>
<td>.009</td>
</tr>
<tr>
<td>33</td>
<td>1050</td>
<td>6500</td>
<td>.00</td>
<td>.00</td>
<td>0.00</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 4 presents the APLs coordinates \((x_a, y_a)\) along with their activation frequency \((act_a = 1)\) over the 5 simulations in column (a) of Table 4. Column (b) shows the average number of modules installed at each APL, while column (c) indicates the average daily demand that each facility is expected to serve. Finally, column (d) represents the average saturation of each APL.

Column (a) of Table 4 reveals that at least ten APLs have an activation frequency of 50%, with an average saturation ranging from approximately 69.6% to 95% and an average number of modules of around three. To cover at least 95% of the estimated potential users in each simulation, between 11 and 13 APLs are required. The same results were obtained by executing Algorithm 1 in its entirety and, separately, by replacing its step one output with the MILP APL sequencing. Algorithm 2 also led to the same conclusions. The results of the various algorithms used are presented in Table 5.
The sequence A1S1.APL represents the order of APLs evaluated using Algorithm 1 Step 1, while A1S1-A1S2.APL is the result of Algorithm 1 Step 2 with A1S1.APL as the fixed input sequence. The differences between these sequences imply that activating two or more facilities may prevent the placement of the other ones and vice versa. A2.APL is the output of Algorithm 2, while MILP.APL refers to the APL column in Table 4, which is used as input to Algorithm 1 Step 2 to obtain the MILP-A1S2.APL facilities sequence.

The A1S1-A1S2.APL, A2.APL, and MILP-A1S2.APL sequences share both similarities and differences. Although the first APLs in A1S1-A1S2 and A2 are mostly the same, with minor differences in their order, the MILP-A1S2 sequence is completely different. The only exception is APL 31, which appears in the seventh, second, and first positions in the three sequences, respectively. To cover around 90% of the estimated demand, a minimum of ten or eleven APLs is
required. In all three cases, the first APLs cover approximately 15.5% of the total estimated demand (the intersections with the y-axis), and each additional facility provides similar increments in coverage (curvature). This is due to the combined effect of users’ homogeneous demand towards APL usage (Ottaviani et al., 2020) and the volumetric capacity of the facilities.

Figure 4 visually represents the A1S1-A1S2, A2, and MILP-A1S2 facility networks to clarify the differences in the three approaches.
4.2 Sensitivity analysis

We conducted a sensitivity analysis to assess the impact of average tolerated travel distance ($\overline{tol}$) on the performance of the three algorithms. The sensitivity analysis involved testing a range of values (i.e., from 0 meters to 6600 meters, in increments of 100), and evaluating the resulting
maximum covered demand \((\text{max}\%\text{dem}_c)\) as well as the minimum number of activated APLs required to achieve 95% coverage (when possible). Given that the capacity of APLs can be increased through expansion, whereas the distance between facilities is fixed, the sensitivity analysis focused solely on varying the average tolerated travel distance while keeping the average demand constant.

Figure 5 shows the results of the sensitivity analysis. The number of APLs activated is plotted as continuous lines on the left \(y\) axis, while the dashed lines on the right \(y\) axis indicate the maximum percentage of covered cumulative demand \((\text{max}\%\text{dem}_c)\). Based on the results, the number of APLs decreases by half when \(\overline{\text{tol}} > 1600\). For \(\overline{\text{tol}} = 1800\), the same results as in Table 5 apply. We observe that A1S1-A1S2 is not monotone decreasing in the region \(\overline{\text{tol}} \geq 3100\), as the algorithm follows a greedy approach and does not optimize globally. In contrast, both A2 and MILP-A1S2 exhibit monotone decreasing behavior.

![Figure 5. Sensitivity analysis (with \(\overline{\text{dem}} = 0.027\)). Continuous lines represent the number of APLs required to cover 95% of the demand. Dashed lines represent the maximum \%\text{dem}_c.](image)

### 5 Discussions and conclusions

This work generates both theoretical and practical contributions.

The theoretical contribution of the study is related to the two approaches to solving the covering location problem of an APL network and their application to the city of Turin. The first approach consists of the robust optimization of a MILP model. The model is described in terms of the objective function, constraints, decision variables, and parameters involved. A simplified version is suggested whereby the possible locations of the facilities to install are fixed. This model is thus optimized \(S\) times, changing the user-related variables during each simulation. We averaged the results from the \(S\) runs, obtaining a custom ranking of the APLs based on their activation frequency and time. The second approach is represented by the execution of an ad-hoc APLs selection algorithm. In this regard, two criteria for selecting the next APL to install are tested: for the first criterion, the same input sequence of order is simulated; for the second criterion, the selection of the next APL depends on the simulated outcomes of the following possible option.

The practical implications of the study are the following. Practitioners are provided two approaches for solving the covering location problem related to the APL network installation. The MILP approach is focused on the simultaneous optimization of the APLs and the number of extra modules per facility. Nevertheless, it does not allow identification of the incremental percentage of
Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal

satisfied demand provided by the individual APL activation. On the other hand, the algorithmic approach considers the APLs and their capacity one at a time, allowing for the evaluation of the total percentage of the satisfied demand curve according to the \( a \)th APL. In this perspective, all approaches provide similar results (i.e., reach 95% of total satisfying demand with 10 to 11 APLs). The algorithms provide an equivalent level of performance as the MILP, and thus they can easily be replicated in different geographical contexts. This allows an understanding of how many facilities are required per square meter, given the population-tolerated travel distance to pick up an order. Therefore, LSPs can use the proposed solution to fine-tune their APL strategy according to the level of service proposed (i.e., the percentage of demand reached by the APL) and exogenous variables such as the customers’ propensity and flexibility towards the usage of the parcel locker.

This study has some minor limitations. Firstly, only distances from customers’ households are considered, whereas it has been proven that some customers prefer to pick their online purchases during their commuting. This would have required knowing all the workplace locations for all customers involved in the survey, which would have significantly hindered the generalization of the results. Secondly, users coordinates are randomized uniformly within the eight areas. In any case, we believe that the differences between a uniform distribution and the actual distribution of the population within the areas are not likely to result in variation in the output variables of the methods tested.

Future research can improve the study in several ways. On the one hand, one can analyze the default locations to install facilities based on several factors evinced from the literature (e.g., pre-existing facilities, logistical hubs, public spaces, etc.). On the other hand, one can relax the assumptions on which the methods are based, e.g., that the entire demand of a user is assigned to the nearest active APL, or include emission-related costs in the objective function and constraints.

Notation List

This manuscript utilizes the following abbreviations: unformatted abbreviations represent parameters, italicized abbreviations denote variables, bolded abbreviations indicate vectors, and bolded and capitalized abbreviations signify matrices.

Table 6. Acronyms with definitions

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u )</td>
<td>User index</td>
</tr>
<tr>
<td>( a )</td>
<td>APL index</td>
</tr>
<tr>
<td>( \text{nUsers} )</td>
<td>Total number of users</td>
</tr>
<tr>
<td>( \text{nAPLs} )</td>
<td>Total number of APLs</td>
</tr>
<tr>
<td>( U )</td>
<td>Users set</td>
</tr>
<tr>
<td>( A )</td>
<td>APLs set</td>
</tr>
<tr>
<td>( z )</td>
<td>Uniformly distributed random variable</td>
</tr>
<tr>
<td>( x )</td>
<td>Cartesian coordinate</td>
</tr>
<tr>
<td>( y )</td>
<td>Cartesian coordinate</td>
</tr>
<tr>
<td>( \text{de}m )</td>
<td>User daily demand</td>
</tr>
<tr>
<td>( \text{tol} )</td>
<td>User tolerated travel distance</td>
</tr>
<tr>
<td>( \text{sat} )</td>
<td>User Boolean coverage variable</td>
</tr>
<tr>
<td>( D_{ua} )</td>
<td>Euclidian distance between ( u )th user and ( a )th APL</td>
</tr>
<tr>
<td>( \text{pct} )</td>
<td>Users’ frequency for given area</td>
</tr>
<tr>
<td>( \text{pct}_{\text{cum}} )</td>
<td>Users’ cumulative frequency for given area</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>Exponential distribution shape parameter</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>Beta distribution shape parameter</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Beta distribution shape parameter</td>
</tr>
<tr>
<td>( c_{\text{APL}} )</td>
<td>APL installation cost</td>
</tr>
<tr>
<td>( c_{\text{Module}} )</td>
<td>APL module installation cost</td>
</tr>
<tr>
<td>( \text{act}_a )</td>
<td>APL Boolean activation variable</td>
</tr>
<tr>
<td>( \text{SL} )</td>
<td>Service level</td>
</tr>
</tbody>
</table>
Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal

\[ M \] \quad \text{Linear programming big M parameter}

\[ \epsilon \] \quad \text{Linear programming small \( \epsilon \) parameter}

\[ mD_u \] \quad \text{Distance between \( u \)th user and closest APL}

\[ \text{ass}_{APL}^u \] \quad \text{\( u \)th APL assignment to \( u \)th user Boolean variable}

\[ \text{cap}^{APL} \] \quad \text{APL orders capacity}

\[ \text{cap}^{Module} \] \quad \text{APL module orders capacity}

\[ mod_a \] \quad \text{\( a \)th APL number of modules}

\[ \text{mod}^{\text{max}} \] \quad \text{Maximum number of modules per APL}

\[ SF \] \quad \text{User population scaling factor}

\[ APL^* \] \quad \text{Candidate APL}

\[ \text{max}_\text{dem} \] \quad \text{Maximum demand covered by candidate APL}

\[ \text{pot}_\text{dem} \] \quad \text{Potential demand covered by candidate APL}

\[ \text{pot}_\text{users} \] \quad \text{Potential number of users covered by candidate APL}

\[ nSats \] \quad \text{Total number of users covered by selected APL}

\[ pcts \] \quad \text{Matrix of percentage covered demand values by each APL}

\[ pcts_s \] \quad \text{Matrix of percentage covered demand in each simulation}

\[ APL_{\text{maxs}} \] \quad \text{Array of candidate APLs}

\[ APL_{\text{order}} \] \quad \text{Sequenced array of APLs}

Data Availability Statement

All data and models used during the study are available in a repository online in accordance with funder data retention policies. All models or code that support the findings of this study are available from the corresponding author upon reasonable request.

References


Locating Automated Parcel Lockers (APL) with known customers’ demand: a mixed approach proposal


