Computers & Operations Research 160 (2023) 106393

Contents lists available at ScienceDirect

Computers and Operations Research

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

A logic-based Benders decomposition solution approach for two covering problems that consider the underlying transportation

Vera Fischer [a,*], Sanne Wøhlk [b]

[a] DS&OR - Decision Support and Operations Research, Department of Informatics, University of Fribourg, Bd de Pérolles 90, CH-1700 Fribourg, Switzerland
[b] CORAL - Cluster for Operations Research, Analytics, and Logistics, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark

ARTICLE INFO

Keywords:
Logic-based Benders decomposition
Maximal covering location problem
Transportation
Perfect matching
Waste collection

ABSTRACT

We investigate two maximal covering location problems with capacity restrictions, minimum workload, and transportation. The problems are inspired by a waste collection problem in which large waste containers are scattered throughout the municipality, and the residents bring their waste to these containers. We take the residents' preferences into account when allocating them to locations. When a container is full, a vehicle transports an empty container from the disposal facility (depot) to that location and replaces it. We propose a mixed-integer linear programming formulation for the problems in which vehicles can carry one or two containers, and apply a logic-based Benders decomposition approach for the latter. Here, the sub problem is a multi-period minimum weight perfect matching problem. We show that our logic-based Benders decomposition approach outperforms the direct formulation in terms of solution quality and speed. We further show that transportation of two containers at a time reduces the distance to be driven by 29.5% on average, without compromising the covering level. Furthermore, we analyze the effect of imposing a minimum workload as well as the effect of changing the focus between transportation and covering.

1. Introduction

This article considers two variants of a covering problem that consists in identifying a subset of candidate locations to obtain maximum coverage over a set of demand nodes while minimizing the underlying total transportation cost. We refer to this problem as the maximal covering location problem (MCLP) that considers the underlying transportation (MCLP-T). We assume that a given amount of demand is associated with each demand node. To justify the opening of a candidate location, the total amount of demand assigned to it must lie between a given minimum and maximum threshold. For each demand node, we assume a total preference ordering of the candidate locations according to some convenience measure (e.g., walking distance, proximity to interesting points), and it must be assigned to its most preferred location that is opened. It is considered covered if its demand is allocated to a candidate location that lies within a given maximal distance, and we refer to such a location as acceptable for that demand node. All other candidate locations are considered unacceptable for that demand node. If no acceptable candidate location is available, a demand node is assigned to the most preferred unacceptable location as a last resort.

Given the set of selected candidate locations and their capacities, the total demand assigned to each location and the number of visits needed within a given time horizon can be deduced. Demand is evenly distributed over the whole time horizon, such that each location is visited at regular intervals. In other words, given the day of the first visit to one of the locations, the days of all subsequent visits to that location are known. We consider two variants of the problem. In the first variant, we consider vehicles that can satisfy the demand of exactly one location at a time. The locations are then visited as many days as needed. In the second variant, we consider vehicles that can satisfy up to two locations on a single trip. Since in both problem variants there is a maximum number of two visits in each trip, we do not need to determine a sequence of locations as is typically done in routing problems. Consequently, both problems are defined as transportation problems. Fig. 1 visualizes the idea of the two problems in a simplified way. The green rectangles represent the selected locations, and the green and blue arcs show a solution for vehicles of capacity one and two, respectively. In this example, five of the six demand nodes are assigned to acceptable locations while one is allocated to an unacceptable location.

We have observed this problem in the context of waste collection. In many countries, a curbside system is applied for residential waste

* Corresponding author.
E-mail addresses: vera.fischer@unifr.ch (V. Fischer), sanw@econ.au.dk (S. Wøhlk).

https://doi.org/10.1016/j.cor.2023.106393
Received 19 October 2022; Received in revised form 15 August 2023; Accepted 15 August 2023
Available online 22 August 2023

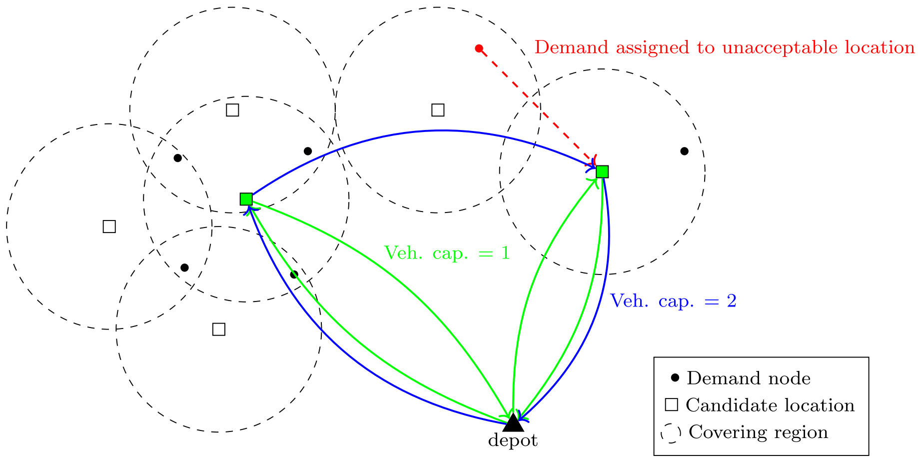


Fig. 1. Selection of candidate locations and solution with vehicles of capacity 1 or 2.

Fig. 2. Side loader vehicle carrying one container (Patent Alpenluft; SAAG).

consisting of heavy trucks stopping at almost each household, which causes high fuel consumption, emissions and noise. We are interested in designing more efficient and sustainable waste collection systems. Therefore, in this article, we investigate an alternative design which has been proposed by the industrial partner of the underlying research project. In this system, large waste containers are scattered throughout the municipality, and residents are asked to take their waste to these locations. At each of these locations exactly one container is established. When a container is full, a special side loader vehicle as shown in Fig. 2 transports an empty container from the disposal facility (depot) to that location and replaces it. The vehicle then transports the full container back to the disposal facility and discharges it. Currently, such vehicles can carry one container at a time, and we investigate the added benefit of using vehicles that can carry up to two containers at a time. To the best of our knowledge, this has not yet been analyzed.

The main contributions of this paper are the following. Based on our observations in practice and following (Farahani et al., 2012)'s recommendations, we formulate a capacitated version of the MCLP that additionally considers a minimum threshold to justify the opening of container locations. Second, we provide mathematical formulations for the problems with vehicle capacities of one and two containers, respectively. Third, we propose a logic-based Benders decomposition approach that decomposes the problem with vehicle capacity two into an MCLP in the master problem and a minimum cost perfect matching problem over a time period in the sub problem. This approach allows to overcome the weakness of the exact formulation with its additional indicator variables and big-M-constraints. Fourth, we derive managerial insights with respect to the different vehicle capacities, the integration of the transportation distance in the objective, the imposition of a minimum workload on container locations, and the allocation of uncovered demand.

The remainder of the paper is organized as follows. In Section 2, we briefly position our work in relation to existing literature. The problem is formally defined in Section 3 where we also discuss our assumptions. The two covering problems are presented in Sections 4 and 5. To be more specific, we define the problem and its two variants in the context of waste collection. The logic-based Benders decomposition approach is described in Section 6. The computational experiments using data from our industrial partner are reported in Section 7, and concluding remarks follow in Section 8.

2. Related work

Our problems fall into the category of so-called covering problems in which a customer is said to be covered by a facility, if the facility lies close enough to the customer, as defined by the problem statement (Kolen and Tamir, 1984). Covering problems represent one of the most popular problems among facility location problems (FLP). Schilling

(1993) and Farahani et al. (2012) present consecutive literature reviews on covering problems in facility location. They classify models which use the concept of covering in two main categories: set covering location problem (SCLP), where coverage of demand is required, and MCLP where coverage of demand is optimized. They also present areas that can be considered for further research such as capacitated facilities and multi-objective covering problems. More recently, García and Marín (2019) provide a review on covering problems and their main models and results. They introduce a general covering model which can be adapted to model specific situations as they can be seen as particular cases of the general model. For a general overview of location-covering problems see Laporte et al. (2019) and Church et al. (2018).

The SCLP aims at finding locations to place facilities at minimum cost while satisfying a specified level of coverage, i.e., service distance (Toregas et al., 1971). In practice, limited resources might, however, not allow to cover all the demand for a given level of coverage. Church and ReVelle (1974) introduce the MCLP that locates a given fixed number of facilities such that the amount of demand covered within the acceptable service distance is maximized. Most of the articles in existing literature on covering problems consider facilities that are uncapacitated. To overcome this limitation, Current and Storbeck (1988) present a capacitated version of the MCLP formulation. The MCLP is most applicable in the public sector (e.g., governmental organizations) where the goal is to maximize service to people with limited resources, such as in waste collection (Schilling, 1993).

Current and Schilling (1994) introduce the maximal covering tour problem (MCTP) in which the tour must visit only p of the n nodes on the network. The MCTP maximizes the total demand covered, as does the MCLP, but it also minimizes the total tour length. In general, these two objectives conflict with each other such that no single solution exists that optimizes all of the objectives simultaneously. A set of efficient solutions can be found, for which an improvement in one objective requires a degradation in the other objective. The authors propose a heuristic to generate an approximation of the set of efficient solutions. To measure the performance of solutions to the MCTP, a weighted sum of the two objectives can be used. The MCTP belongs to the location-routing problems (LRP) that represent an approach to model and solve locational problems while considering vehicle routing aspects to decrease the overall system cost (Prodhon and Prins, 2014; Nagy and Salhi, 2007). We refer the reader to these references for an overview on LRP, as the problems considered in this article (see Section 1) are defined as transportation problems in which no classical routing decisions are made.

In the context of waste collection, Adeleke and Olukanni (2020) present an overview on facility location problems in solid waste management. They focus on the decision of locating waste collection facilities such as recycling centers, waste-to-energy facilities, or containers/bins within a waste collection network. To design an effective solid waste management system, locating the waste containers is very crucial. Ghiani et al. (2012) propose an integer programming model for the capacitated location of collection sites while considering the quality of service such that each customer is served by the collection site nearest to him. They minimize the total number of sites to be opened and determine the optimal allocation of waste bins to the collection sites such that all the demands are satisfied. They also propose a construction heuristic for obtaining good solution quality within an extremely reduced computational time. Computational results on real-life data of the city Nardò in Italy show that both exact and heuristic approaches provide consistently better solutions than the approach currently implemented with fewer collection sites and bins used. Cubillos and Wøhlk (2020) address the location-routing problem of recycling drop-off stations by solving an MCTP. They propose a heuristic approach inspired by a variable neighborhood search, which can effectively find good quality solutions. Based on a set of real-life instances, they show the trade-off between covering and collection costs for urban and rural areas.

3. Problem definition

In this section, we formally define our problem. To make it more tangible for the reader, we phrase it in the context of waste collection. The problem is defined over a complete graph $G = (N, A)$, with $N = F \cup V \cup \{\sigma\}$. F is a set of candidate locations, and for each location $j \in F$, we need to decide whether or not to place a waste container there. V is the set of demand nodes representing the residential buildings. σ represents the disposal facility (depot) that holds a sufficiently large fleet of identical vehicles. We use d_{ij} to denote the (potentially asymmetric) distance from $i \in N$ to $j \in N$. It may be that a candidate location is the same physical place as one of the residential nodes. In this case, they are represented by separate nodes with a distance of zero between them.

Each node $i \in V$ can refer to a single residential building or group several together, as buildings are assigned to the nearest node in V. We denote its waste by w_i. Let ρ be the maximal acceptable walking distance for a resident. Then for each residential node $i \in V$, the set of candidate locations F is partitioned into an acceptable set F_i^+ and an unacceptable set F_i^-, with $F = F_i^+ \cup F_i^-$, $F_i^+ \cap F_i^- = \emptyset\ \forall i \in V$ such that $d_{ij} \leq \rho\ \forall j \in F_i^+$ and $d_{ij} > \rho\ \forall j \in F_i^-$. Residents could take their waste to unacceptable locations if no container is placed within an acceptable distance. This is typically the case for distant residential buildings for which the placement of additional containers would be unjustifiable and would have a disproportionate impact on the final solution. Note that residents can walk in both directions, which is why their distance matrix becomes symmetric. More precisely, for any i and j, if $d_{ji} < d_{ij}$, we assume for both directions the shorter one d_{ji}. Furthermore, we assume a total ordering $\pi(i,j)$ of $j \in F$ for every $i \in V$ such that $\pi(i,j) < \pi(i,j')$ if j is preferred over j' by i and $\pi(i,j) < \pi(i,j')\ \forall j \in F_i^+, j' \in F_i^-$. In case two locations, $\{j, j'\}$, are equally preferred by i, their ordering is built arbitrarily.

At most p locations can be established with a container, and a container $j \in F$ has a limited capacity Q. A container can only be placed at a candidate location if it gathers a minimum amount of waste (i.e., minimum workload) $w^{\min}$. Each resident must be assigned to its most preferred container location. Let $S \subseteq F$ be the set of established locations. Then $i \in V$ must be assigned to $j \in S$ such that $\pi(i,j) < \pi(i,j')\ \forall j' \in S \setminus \{j\}$. If $i \in V$ is assigned to a location in F_i^+, then i is said to be covered. Otherwise, it is said to be uncovered.

Let $\mathcal{T} = \{1, \ldots, T\}$ be the set of days in the time period considered for this problem. Note that this set of T days will repeat itself over a longer time horizon. We assume that each location $j \in F$ is visited at most once a day $t \in \mathcal{T}$. To ensure that enough capacity is available, the following equation must hold: $pQT \geq w^{\text{tot}}$, where $w^{\text{tot}} = \sum_{i \in V} w_i$.

To replace the container at an established location $j \in F$, we consider special side loader vehicles (see Fig. 2) and model our problems as transportation problems in which vehicles do not perform classical routes. In the first problem variant, a vehicle with a capacity of one container must perform the round trip σ - j - σ, which induces a transportation cost of $c_j = d_{\sigma j} + d_{j\sigma}$. In the second problem variant, a vehicle with a capacity of two containers may also perform a round trip which induces a transportation cost of $c_{jj'} = \min\{(d_{\sigma j} + d_{jj'} + d_{j'\sigma}), (d_{\sigma j'} + d_{j'j} + d_{j\sigma})\}$. Note that such a vehicle takes the shorter round trip, which is why we use the smaller travel cost.

A subset $S \subseteq F$ represents a feasible solution if the minimum workload and the capacity of the containers installed at locations in S are respected, and the waste of each residential node is allocated to its most preferred location in S. Note that, in contrast to the MCLP, we do satisfy demand of uncovered residential nodes. Indeed, for each residential node $i \in V$, we allocate its waste to its most preferred location in S, even if uncovered. Considering that the total demand at each location must be between a given minimum and maximum threshold, the allocation of uncovered demand might therefore affect the location decision such that more locations are established which lie closer to the uncovered residential nodes. This effect will be discussed in Section 7.6.

The goal of the optimization problem is to determine the subset $S \subseteq F$ that maximizes the total amount of waste of the residential nodes $i \in V$ that have been assigned to acceptable locations, i.e., a location $j \in F_i^+ \cap S$, or equivalently, minimizes the waste that is assigned to unacceptable locations, and that also minimizes the total transportation distance that arises from the number of replacements of containers performed during $\mathcal{T}$. To balance these two objectives, we use a weighting parameter α.

3.1. Model assumptions

To define our problem in a concise manner, we make the following assumptions. Some of these assumptions simplify reality and exclude, for instance, uncertainty in demand. However, we believe that they are reasonable to build a first model of the problem.

(A1): The amount of demand at each location is deterministically known. In our problem, we consider the locations of waste containers, which is a long-term decision. Containers are placed such that they cover a maximum amount of waste from the residents, while keeping the total transportation distances as low as possible. However, we do not provide solutions for any transportation operations that may have to be executed on-demand due to received sensor information on the fill levels of the containers. Furthermore, our observation from real-life operations is that fluctuations in demand are reduced when aggregating over multiple households. As in reality the amount of demand is uncertain, demand estimations can be based on the maximum or 95% quantiles of the demand distributions (instead of the average or median) to ensure service quality in any situation.

(A2): The demand is evenly distributed over time. Within a time horizon of seven days, we assume that each day is equally preferred by all residents to dispose their waste. It can be argued that in reality residents may have different preferences, as they might, for instance, have more waste to dispose on weekend days. Nevertheless, as a starting point, we choose to model our problem using a uniform distribution of aggregated waste over time.

(A3): Residents bring the waste to their most preferred location. Given the locations of the waste containers, it makes sense to assume that residents would choose the location that is most convenient for them. Our model is capable of handling any preference ordering of the container locations with respect to some convenience measure such as, for instance, walking distance or proximity to interesting points (e.g., train station, shopping center). In our experiments, we use the walking distances from the residential buildings to the container locations as a measure, which is the most natural indicator. Moreover, it seems fair to assume that residents would always bring their waste to the same location, rather than changing it from one day to another, since people generally tend to repeat their actions out of habit (e.g., going to work every day).

(A4): A time horizon of seven days is considered. In practice, waste collection is often executed using weekly plans, such that the same collection tours are repeated every week. Assuming a time horizon of seven days, which will repeat itself over a longer time horizon, can therefore be justified.

(A5): Activities can be done in any order. Vehicles with a capacity of two containers can visit up to two locations consecutively. We assume that each of the empty containers on the vehicle can be replaced by a full container without rearranging them. This is also our observation in practice, where vehicles are equipped with cranes that can handle containers individually. Furthermore, as is common in the academic literature, we minimize the travel distances and do not consider operation times in our problems.

(A6): The number of containers installed is restricted by their minimum and maximum capacity. In our problem, we impose a minimum workload and a maximum capacity at each container location and due to these restrictions, it may not be allowed to add a container to a location, considering the other container locations. Hence, these two constraints have the joint effect of limiting the number of container locations.

4. Covering problem with vehicles with a capacity of one container

To model the covering problem with vehicles of capacity one mathematically, we use three sets of variables. Let y_j be a binary variable that is equal to 1 if a container is established at location $j \in F$, and a_{ij} be a binary variable that is equal to 1 if a residential node $i \in V$ (and thereby its waste w_i) is allocated to the location j. Let s_j be an integer variable that represents the number of visits at location $j \in F$. The problem is formulated in (1). We refer to this formulation as the *Cap1*-formulation in the remainder of the paper.

$$\min \ \alpha\Big(\sum_{i\in V}\sum_{j\in F_i^-} w_i a_{ij}\Big) + (1-\alpha)\Big(\sum_{j\in F} c_j s_j\Big) \tag{1a}$$

$$\text{s.t.} \ \sum_{j\in F} a_{ij} = 1 \qquad \forall i \in V \tag{1b}$$

$$\sum_{j'\in F:\pi(i,j')>\pi(i,j)} a_{ij'} \le 1 - y_j \qquad \forall i \in V, j \in F \tag{1c}$$

$$\sum_{i\in V} w_i a_{ij} \ge w^{\min} y_j \qquad \forall j \in F \tag{1d}$$

$$\sum_{i\in V} w_i a_{ij} \le s_j Q \qquad \forall j \in F \tag{1e}$$

$$\sum_{j\in F} y_j \le p \tag{1f}$$

$$s_j \le y_j T \qquad \forall j \in F \tag{1g}$$

$$a_{ij}, y_j \in \{0,1\} \qquad \forall i \in V, j \in F \tag{1h}$$

$$s_j \in \mathbb{Z}_{\ge 0, \le T} \qquad \forall j \in F \tag{1i}$$

The first term of the objective function (1a) defines the goal of minimizing the total waste that is assigned to unacceptable locations, while the second term represents the minimization of the total transportation cost. Constraints (1b) ensure that the waste of each residential node is assigned to exactly one location. Constraints (1c) impose that each residential node is allocated to its most preferred open location. Constraints (1d) enforce the minimum workload $w^{\min}$ on open locations. Constraints (1e) enforce that over the whole planning period, a location j is limited to the capacity Q times the number of visits s_j at location j. Constraints (1f) state that at most p containers can be placed at the candidate locations. Constraints (1g) ensure that a residential node can only be allocated to a location if it is opened. Constraints (1h)–(1i) define the domains of the variables.

5. Covering problem with vehicles with a capacity of two containers

In this section, we present our model for the problem with vehicles of capacity two. As the model is not directly solvable, we present two different approaches for solving it. First, in Section 5.1, we linearize the model. Second, in Section 6, we take a logic-based Benders approach to solve the problem.

Considering our assumptions in Section 3.1, the covering problem with vehicles of a capacity of two containers is formally defined as follows. Let ζ be a dummy location and define $F' = F \cup \{\zeta\}$. This dummy location enables a feasible matching of locations on days where an odd number of locations needs collection. Note that due to the triangle equation, at most one location will be visited alone in one day. More precisely, the cost of visiting locations $j \in F'$ and $j' \in F'$ in a single trip is less or equal to the cost of visiting each location in a separate trip, i.e., $c_{jj'} \leq c_j + c_{j'}$, where the matching cost to the dummy location $c_{j\zeta}$ is equal to the transportation cost c_j. The goal thus becomes to find a subset $S \subseteq F$ with maximum coverage and to define for each day $t \in \mathcal{T}$ pairs of locations $\{j,j'\} \in S', S' = S \cup \{\zeta\}$ that are visited in the same trip in order to minimize the total transportation distance.

Let $x^t_{jj'}$ be a binary variable that takes value 1 if locations $j,j' \in F' : j \neq j'$ are visited in the same trip in period $t \in \mathcal{T}$. For each location $j \in F$, we introduce the integer variables l_j and u_j that define a lower and upper bound on the number of consecutive days without a visit, such that the capacity Q is never exceeded. The problem is formulated in (2) keeping the variables defined in formulation (1).

$$\min \ \alpha\Big(\sum_{i\in V}\sum_{j\in F_i^-} w_i a_{ij}\Big) + (1-\alpha)\Big(\sum_{t\in\mathcal{T}}\sum_{j\in F'}\sum_{j'\in F'} c_{jj'} x^t_{jj'}\Big) \tag{2a}$$

$$\text{s.t. (1b)} - \text{(1g)} \tag{2b}$$

$$\sum_{j'\in F'} x^t_{jj'} + x^t_{j'j} \leq y_j \qquad \forall j \in F, t \in \mathcal{T} \tag{2c}$$

$$\sum_{t\in\mathcal{T}}\sum_{j'\in F'} x^t_{jj'} + x^t_{j'j} = s_j \qquad \forall j \in F \tag{2d}$$

$$\sum_{t'\in\{t,\dots,t+l_j-1\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \leq 1 \qquad \forall j \in F, t \in \mathcal{T} : t \leq T - l_j + 1 \tag{2e}$$

$$\sum_{t'\in\{t,\dots,T\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} + \sum_{t'\in\{1,\dots,t+l_j-T-2\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \leq 1 \qquad \forall j \in F, t \in \mathcal{T} : t \geq T - l_j + 2 \tag{2f}$$

$$\sum_{t'\in\{t,\dots,t+u_j-1\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \geq 1 \qquad \forall j \in F, t \in \mathcal{T} : t \leq T - u_j + 1 \tag{2g}$$

$$\sum_{t'\in\{t,\dots,T\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} + \sum_{t'\in\{1,\dots,t+u_j-T-2\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \geq 1 \qquad \forall j \in F, t \in \mathcal{T} : t \geq T - u_j + 2 \tag{2h}$$

$$x^t_{jj'} = 0 \qquad \forall j,j' \in F', j \geq j', t \in \mathcal{T} \tag{2i}$$

$$\sum_{j\in F}\sum_{j'\in F'} x^1_{jj'} + x^1_{j'j} \geq \sum_{j\in F} s_j / T \tag{2j}$$

$$l_j \leq u_j \leq l_j + 1 \qquad \forall j \in F \tag{2k}$$

$$s_j, l_j, u_j \in \mathbb{Z}_{\geq 0, \leq T} \qquad \forall j \in F \tag{2l}$$

$$a_{ij}, y_j \in \{0,1\} \qquad \forall i \in V, j \in F \tag{2m}$$

$$x^t_{jj'} \in \{0,1\} \qquad \forall j,j' \in F', t \in \mathcal{T} \tag{2n}$$

The Eq. (2a) defines the two goals of minimizing the total waste that is assigned to unacceptable locations and minimizing the total matching cost. Constraints (2b) define the covering problem with vehicles of capacity one in (1), and they remain unaltered for this extended problem. Constraints (2c)–(2d) ensure that each location j is visited at most once a day t and over the whole time period exactly s_j times. Constraints (2e)–(2h) impose that the number of consecutive days without a visit must be between the two bounds l_j and u_j. Constraints (2k) specify the relationship between the lower and upper bounds. Constraints (2i) forbid a matching of two locations $\{j,j'\}$ such that $j \geq j'$, since a matching $i-j$ is equivalent to $j-i$ and thus has the effect of removing symmetry. To mitigate the negative effects of the inherent symmetry obtained with the time period index t on the x-variables, we introduce symmetry breaking constraints as follows. Let $s^{\text{tot}} = \sum_{j\in F} s_j$ be the total number of visits to all locations in the time period $\mathcal{T}$. Now we assume that at least the average number of visits, namely s^{tot}/T, are made in the first period $t = 1$. This leads us to constraints (2j). Finally, constraints (2l)–(2n) define the domains of the variables.

5.1. Linearizing the problem

Clearly, this model is not solvable with the l_j and u_j variables in sum and set statements (see constraints (2e)–(2h)). To overcome this issue, we introduce additional variables d_{jk} and δ_{jk} that indicate whether a visit is needed at location $j \in F$ within a period of $k \in \mathcal{T}$ consecutive days according to l_j and u_j. More precisely, if $k \leq l_j$, then $d_{jk} = 1$, and if $k \geq u_j$, then $\delta_{jk} = 1$. Such variables are called *indicator variables* and are linked to other variables in the problem to indicate certain states (Williams, 2013). Using these additional variables, constraints (2e)–(2h) can be replaced by (3c)–(3i), which results in formulation (3). We refer to this formulation as the *BigM*-formulation in the remainder of the paper.

$$\min \ \text{(2a)} \tag{3a}$$

$$\text{s.t. (2b)–(2d), (2i)–(2j)} \tag{3b}$$

$$l_j \leq M d_{jk} + k - 1 \qquad \forall j \in F, k \in \mathcal{T} \tag{3c}$$

$$u_j \geq k + 1 - M\delta_{jk} \qquad \forall j \in F, k \in \mathcal{T} \tag{3d}$$

$$\sum_{t'\in\{t,\dots,t+k-1\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \leq 1 + M - M d_{jk} \qquad \forall j \in F, k \in \mathcal{T}, t \in \mathcal{T} : t \leq T - k + 1 \tag{3e}$$

$$\sum_{t'\in\{t,\dots,T\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} + \sum_{t'\in\{1,\dots,t+k-T-2\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \leq 1 + M - M d_{jk} \qquad \forall j \in F, k \in \mathcal{T}, t \in \mathcal{T} : t \geq T - k + 2 \tag{3f}$$

$$\sum_{t'\in\{t,\dots,t+k-1\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \geq 1 + M\delta_{jk} - M \qquad \forall j \in F, k \in \mathcal{T}, t \in \mathcal{T} : t \leq T - k + 1 \tag{3g}$$

$$\sum_{t'\in\{t,\dots,T\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} + \sum_{t'\in\{1,\dots,t+k-T-2\}}\sum_{j'\in F'} x^{t'}_{jj'} + x^{t'}_{j'j} \geq 1 + M\delta_{jk} - M \qquad \forall j \in F, k \in \mathcal{T}, t \in \mathcal{T} : t \geq T - k + 2 \tag{3h}$$

$$d_{jk}, \delta_{jk} \in \{0,1\} \qquad \forall j \in F, k \in \mathcal{T} \tag{3i}$$

$$\text{(2k)–(2n)} \tag{3j}$$

Constraints (3c)–(3d) link the indicator variables d_{jk} and δ_{jk} with the bound variables l_j and u_j, where M is a constant representing an upper bound for l_j and u_j. Constraints (3e)–(3h) define for each period of k consecutive days starting at day t whether location j has to be visited at most/least one time. Constraints (3i) define the domains of the indicator variables.

6. Logic-based Benders decomposition approach

The formulation developed in Section 5.1 is quite weak due to the extra variables and *BigM*-constraints. In this section, we propose a logic-based Benders decomposition approach, originally introduced by Hooker and Ottosson (2003), to decompose the problem into the following two sub problems:

The *master problem* comprises the selection of a subset of candidate locations that cover a maximum amount of waste from the residential nodes. We call such a subset a *max cover*. At each integer node of the branch and bound tree of the master problem, we provide the solution (which is a *max cover*) to the sub problem.

The *sub problem* involves a matching of locations in the *max cover* (i.e., pairs of locations visited in the same trip) over the time horizon $\mathcal{T}$ at minimum total cost. If the master problem solution satisfies all the bounds produced by the sub problems, convergence has been achieved. To get an integer solution at the end of the solution process, we simply return the best combined solution found over all iterations.

Due to the integer variables in the sub problem, standard linear duality cannot be used to develop classical Benders cuts. Therefore, we generate a Benders cut based on the bound on the total cost provided by the sub problem implied by the given *max cover*, which is then added to the master problem. This decomposition allows us to avoid the variables l_j and u_j, since their values can be derived from the solution of the master problem. Hence, l_j and u_j are parameters of the sub problem so that the initial constraints (2e)–(2h) can be used in the sub problem. Thereby, the indicator variables d_{jk} and δ_{jk} and all the linearization constraints (3c)–(3i) are no longer needed. We refer to this approach as the *Benders* in the remainder of the paper.

6.1. Master problem

The master problem is formulated in (4). To tighten the bounds, we add a relaxation of the sub problem to the master problem. More precisely, we keep the linear constraints on the matching ((2c), (2d), (2i)–(2j)) in the formulation, but relax the variables $x^t_{jj'}$. The associated cost is then added to the objective function (4b).

$$\min \ Z \tag{4a}$$

$$\text{s.t. } Z = \alpha\Big(\sum_{i\in V}\sum_{j\in F_i^-} w_i a_{ij}\Big) + (1-\alpha)\Big(\sum_{t\in\mathcal{T}}\sum_{j\in F'}\sum_{j'\in F'} c_{jj'}x^t_{jj'}\Big) \tag{4b}$$

$$\text{(2b)–(2d), (2i)–(2j)} \tag{4c}$$

$$a_{ij}, y_j \in \{0,1\} \qquad \forall i\in V, j\in F \tag{4d}$$

$$s_j \in \mathbb{Z}_{\geq 0,\leq T} \qquad \forall j \in F \tag{4e}$$

$$x^t_{jj'} \geq 0 \qquad \forall j,j' \in F', \forall t\in\mathcal{T} \tag{4f}$$

$$Z \geq 0 \tag{4g}$$

6.2. Sub problem

In the sub problem, we solve a matching problem for a given solution to the master problem (4). Let $S \subseteq F$ be a *max cover* that represents a feasible solution to the master problem. More precisely, S consists of the locations for which $y_j = 1$ in the current master problem solution. Given the locations of the containers, let $q_j = \sum_{i\in V} w_i a_{ij}$ be the total waste that is gathered at each location $j \in S$. The number of visits needed at each location $j \in S$ can thus be deduced as $s_j = \lceil q_j/Q \rceil$, and the lower and upper bounds on the number of consecutive days without a visit, l_j and u_j, are calculated as $l_j = \lfloor T/s_j \rfloor$ and $u_j = \lceil T/s_j \rceil$, respectively. Note that s_j, l_j and u_j are constants in the sub problem. Furthermore, let $S' = S \cup \{\zeta\}$, with ζ being a dummy location to always guarantee a feasible matching of locations. The objective function (5a) defines the goal of minimizing the total transportation cost of the vehicles. Constraints (5b) specify the matching problem. Constraints (5c) define the domains of the $x^t_{jj'}$-variables, which are binary in the sub problem.

$$\min \ Z_{sub} = \sum_{t\in\mathcal{T}}\sum_{j\in S'}\sum_{j'\in S'} c_{jj'}x^t_{jj'} \tag{5a}$$

$$\text{s.t. (2c)–(2j)} \tag{5b}$$

$$x^t_{jj'} \in \{0,1\} \qquad \forall j,j' \in S', t\in\mathcal{T} \tag{5c}$$

6.3. Benders cuts

Given an optimal solution to the sub problem, let $val(Z) = \alpha Z_{unco} + (1-\alpha)Z_{sub}$ be the combined value of the objective function (5a) of the sub problem and the uncovered cost $Z_{unco} = \sum_{i\in V}\sum_{j\in F_i^-:j\in S} w_i a_{ij}$ of the subset S. The Benders cut (6) states that the objective function value of future master problems Z must be greater or equal to $val(Z)$, if all the candidate locations of the subset $S \subseteq F$ are opened and the candidate locations $j \in F\setminus S$ are closed. This Benders cut is then added to the master problem.

$$Z \geq val(Z) - val(Z)\Big(\sum_{j\in S}(1-y_j) + \sum_{j\in F\setminus S} y_j\Big) \tag{6}$$

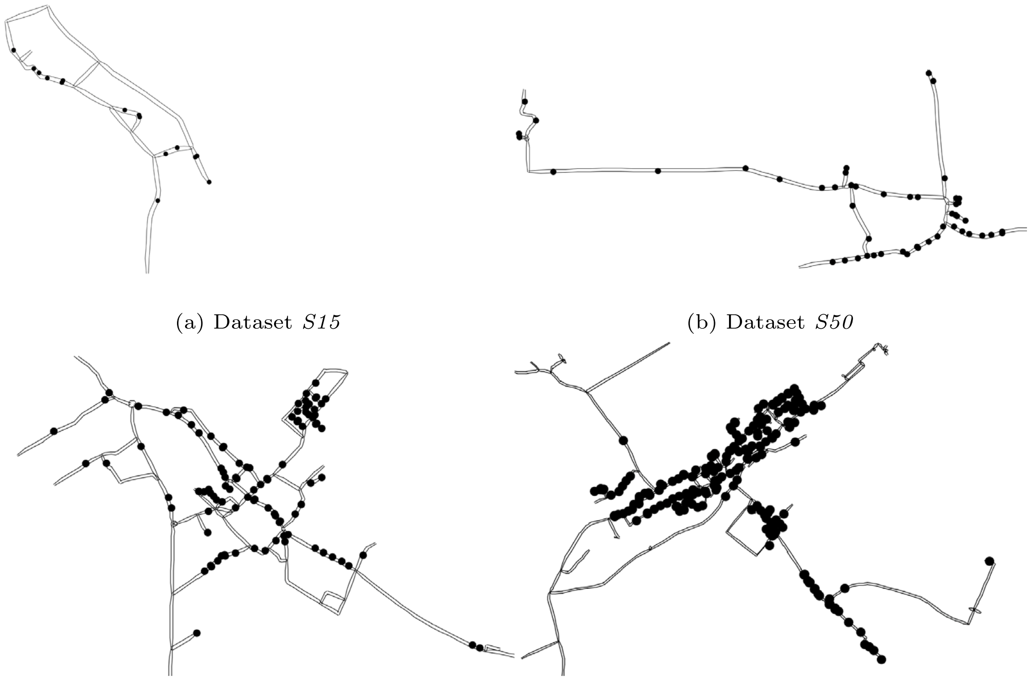
7. Computational experiments

In this section, we present the results of our computational experiments. The MILP formulations and the logic-based Benders decomposition approach have been implemented in Java. To solve the formulations, we use the Gurobi 9.1.2 MIP solver via its Java API. In the Benders master problem, we make use of the Gurobi callback function and solve the Benders sub problem for each integer master solution. Given the sub problem solution, we then generate a Benders cut which is added as a lazy constraint to the master problem. The instances were tested on a computer with a 2.7 GHz Intel Core i5 processor, 32 GB of RAM, operating under Windows 10. A time limit (TL) of 3 hours was imposed on each instance and formulation.

In Section 7.1, we describe the various problem instances considered for the tests. Section 7.2 compares the *BigM*-formulation (Section 5.1) against the *Benders* approach (Section 6) to obtain algorithmic insight. Section 7.3 assesses the benefit of having vehicles with a capacity of two containers for transportation opposed to capacity one vehicles, and Section 7.4 analyzes the impact of including the transportation distances in the objective. Finally, Sections 7.5–7.6 discuss the impact of imposing a minimum workload on container locations and the effect of satisfying demand of uncovered residential nodes on the location decision, respectively. In all tables, we mark the best results in bold.

7.1. Problem instances

To derive the set of instances, we consider four datasets based on real data from Switzerland. These datasets represent directed road-network graphs. The nodes of each graph G either represent residential nodes (V), candidate locations for the containers (F), or they belong to the so-called "intersection" nodes that constitute the road network. The four datasets are called $S|V|$ with $|V| \in \{15, 50, 100, 200\}$ being the number of residential nodes. Fig. 3 shows the street networks and residential nodes of the four datasets. The smallest dataset serves to derive toy instances for development purposes. For each residential node $i \in V$, we partition the set of candidate locations F into an acceptable set F_i^+ and an unacceptable set F_i^- based on the maximal walking distance (ρ). To determine the transportation distances, we construct a distance matrix based on the shortest paths in G between each pair of candidate locations $\{j, j'\} \in F$ and the depot σ.



(a) Dataset *S15* (b) Dataset *S50*

(c) Dataset *S100* (d) Dataset *S200*

Fig. 3. Street networks and residential nodes of the four datasets.

We specify a capacity of $Q = \lceil 0.1w^{\text{tot}} \rceil$ kilograms (kg) for each candidate location and $T = 7$ days for the time horizon. Following our assumption (*A6*, Section 3), we define an infinite upper bound (p) on the number of containers placed. To derive multiple instances from each graph, we define three different values for the maximal walking distance $\rho \in \{0, 150, 300\}$ meters (m) and the minimum workload $w^{\min} \in \{1, \lfloor 0.5Q \rfloor, \lfloor 0.75Q \rfloor\}$ kg. Note that to examine the impact of imposing a minimum workload on locations (Section 7.5), we also include instances where the *constraints* (1d) are ignored, i.e., $w^{\min} = 1$. To evaluate the trade-off between the two conflicting objectives, we define 11 values for the weighting parameter $\alpha \in \{\epsilon, 0.1, \dots, 0.9, 1-\epsilon\}$, with $\epsilon = 0.001$ so that both objectives are weighted in all instances. This results in a total of $3 \cdot 3 \cdot 11 = 99$ instances for each of the four datasets. To support future research in this area, we prepared a subset of those instances for benchmark tests, which is available under www.optimization.dk/MCTP. In the following sections, detailed results for those benchmark instances are provided along with analyses based on aggregated results. The set of benchmark instances was selected to be as representative as possible over all the datasets by keeping all parameters but one fixed while also aiming at keeping the number of instances low. More precisely, for each dataset we set the maximal walking distance and the minimum workload to their middle values as a starting point, namely $\rho = 150$ and $w^{\min} = \lfloor 0.5Q \rfloor$, and selected three representable values for $\alpha \in \{0.2, 0.5, 0.8\}$. To account for other values of ρ and $w^{\min}$, we added some instances for the two medium-sized datasets by varying only one of the two parameter values at a time. Upon request, further instances can be shared.

7.2. Comparison of the Benders approach and the BigM-formulation

In this section, we compare the *Benders* approach with the *BigM*-formulation for the covering problem with vehicles of a capacity of two containers. For each of the two approaches and each graph size, Table 1 presents the total number of instances, the number of instances solved to optimality, the number of instances solved to near optimality, i.e., Gurobi gap $\leq 1\%$ (excluding the ones solved to optimality), and the sum of the two. Furthermore, the table reports the average and the worst Gurobi gaps for those instances that were not solved to optimality. For a fair comparison of the average computation times, we only consider the subset of instances that were solved to optimality within the TL by both approaches, namely the *Benders* approach and the *BigM*-formulation. This results in a total of 99 instances of dataset *S15*, 62 instances of dataset *S50*, 12 instances of dataset *S100*, and no instance of dataset *S200*. These are the instances which are included in the calculation of the average computation times in the last line of Table 1.

We observe that both approaches could solve all toy instances to optimality but only a few instances of the biggest dataset. The main differences are observed for instances of sizes 50 and 100. For these instances, the *Benders* approach was able to solve more instances to optimality or near optimality. Over all datasets, the *Benders* approach was faster in finding optimal solutions and reported lower average and worst gaps for the instances that were not solved to optimality. This clearly shows the superiority of our *Benders* approach. In general, with increasing size of the datasets, both approaches return fewer optimal solutions, report higher computation times for those instances solved to optimality, and have higher gaps for the remaining instances.

Table 2 shows detailed results for the benchmark instances introduced in Section 7.1, which is a selected subset of all instances that is available online to support future research, and reports computation times, best obtained feasible solutions, lower bounds, and Gurobi gaps for each of the two approaches. The same pattern as in the aggregated table can be observed here. Both approaches reach the TL for the largest instances, while the smallest instances could be solved to optimality, whereas for the medium sized instances, the *Benders* approach generally shows better performance.

Fig. 4 visualizes the evolution of the upper and lower bound values over time for the *Benders* and the *BigM* approach for four of the benchmark instances. Each point in the plots represents a new value reported by Gurobi. We observe that for the first two instances of datasets *S15* and *S50* the *Benders* method converges quicker than the *BigM* approach. For the other two instances neither of the approaches is able to prove optimality within the given TL, however, the *Benders* method is able to find a bound and a feasible solution faster and thus provides better bounds early in the solution process.

Table 1
Aggregated results of the *Benders* approach and the *BigM*-formulation for each of the four datasets.

Dataset	*S15*		*S50*		*S100*		*S200*	
	BigM	*Benders*	*BigM*	*Benders*	*BigM*	*Benders*	*BigM*	*Benders*
# instances	99	99	99	99	99	99	99	99
# optimal	99	99	72	**81**	16	**17**	3	3
# nearly optimal	0	0	0	**6**	1	**7**	2	**8**
# opt. + nearly opt.	99	99	72	**87**	17	**24**	5	**11**
average gap (%)	–	–	11.7	**3.8**	28.4	**14.9**	33.6	**20.2**
worst gap (%)	–	–	18.9	**15.7**	94.0	**50.1**	100.0	**84.0**
average comp. time (s)	214.0	**51.8**	1013.1	**444.4**	5363.6	**1421.3**	–	–

Table 2
Results of the *Benders* approach and the *BigM*-formulation for the benchmark instances.

Instance	Alpha	Comp. time (s)		Upper bound		Lower bound		Gap (%)	
$(\|V\| - \rho - w^{\min})$		*BigM*	*Benders*	*BigM*	*Benders*	*BigM*	*Benders*	*BigM*	*Benders*
15-150-0.5Q	0.2	107.5	**15.1**	27.1	27.1	27.1	27.1	0.0	0.0
15-150-0.5Q	0.5	226.1	**57.1**	21.9	21.9	21.9	21.9	0.0	0.0
15-150-0.5Q	0.8	257.6	**58.1**	8.8	8.8	8.8	8.8	0.0	0.0
50-150-0.5Q	0.2	9374.2	**2542.8**	106.5	106.5	106.5	106.5	0.0	0.0
50-150-0.5Q	0.5	9270.2	**527.9**	87.9	87.9	87.9	87.9	0.0	0.0
50-150-0.5Q	0.8	1145.9	**178.5**	58.6	58.6	58.6	58.6	0.0	0.0
100-150-0.5Q	0.2	TL	TL	**163.5**	164.0	**135.1**	134.0	**17.3**	18.3
100-150-0.5Q	0.5	TL	TL	146.1	**141.5**	115.7	**117.0**	20.8	**17.3**
100-150-0.5Q	0.8	TL	TL	97.4	97.4	80.8	**81.6**	17.1	**16.2**
200-150-0.5Q	0.2	TL	TL	**726.8**	735.6	572.4	**583.2**	21.2	**20.7**
200-150-0.5Q	0.5	TL	TL	661.6	**641.2**	**470.0**	441.6	**29.0**	31.1
200-150-0.5Q	0.8	TL	TL	483.2	**440.4**	260.7	**284.8**	46.1	**35.3**
50-150-1	0.5	TL	**1397.9**	87.1	87.1	81.6	**87.1**	6.3	**0.0**
50-150-0.75Q	0.5	1645.1	**289.1**	87.9	87.9	87.9	87.9	0.0	0.0
50-0-0.5Q	0.5	116.1	**28.2**	182.0	182.0	182.0	182.0	0.0	0.0
50-300-0.5Q	0.5	TL	**3646.6**	78.4	78.4	68.9	**78.4**	12.0	**0.0**
100-150-1	0.5	TL	TL	129.1	**126.7**	112.2	**115.4**	13.1	**8.9**
100-150-0.75Q	0.5	TL	TL	147.1	**143.7**	122.2	**123.4**	16.9	**14.1**
100-0-0.5Q	0.5	4369	**1933.5**	337.2	337.2	337.2	337.2	0.0	0.0
100-300-0.5Q	0.5	TL	TL	203.7	**100.6**	83.4	**84.4**	59.1	**16.1**

These experiments show the superiority of the *Benders* approach both from a computational speed and solution quality point of view. Fig. 5 presents boxplots of the number of Benders cuts introduced for the instances that were solved to optimality or near optimality by this approach. Each of the four boxplots represents one dataset, where the number of instances included is indicated below. We observe a large variance of the number of cuts for the considered instances going from zero to almost 6000 cuts and with an average of 283 cuts. For readability reasons, the plots only show outlier values up to 2500 cuts.

7.3. Benefit of vehicles with a capacity of two containers

To assess the benefit of vehicles with a capacity of two containers rather than one, we first compare the two formulations (*Cap1* and *Benders*) with respect to optimality and computation times. Due to the better performance of the *Benders* approach (as discussed in the previous Section 7.2), we use this approach (instead of the *BigM*-formulation) to compare the two different covering problems in this section. Table 3 presents the same key information as Table 1, but for the *Cap1*-formulation and the *Benders* approach. Note that the values for *Benders* are the same as in Table 1, except for the average computation times for which only instances that were solved to optimality by both approaches are considered.

Both approaches could solve all toy instances to optimality. For the instances of the other three datasets, the *Cap1*-formulation found more optimal solutions. In total, however, the *Benders* approach provides more optimal and near optimal solutions for the largest dataset. The average gap of those six near optimal solutions of the *Benders* approach lies below 0.4%. Over all datasets, the *Cap1*-formulation was faster in proving optimality for the instances that were solved to optimality within the TL by both approaches (*Cap1* and *Benders*, resulting in 99 of *S15*, 81 for *S50*, 11 for *S100* and 0 for *S200* instances) and reported lower average and worst gaps for the remaining instances. This indicates that the problem in which the vehicle can only carry one container is easier to solve. Similar to the previous section, the table shows that with increasing size of the datasets, both approaches solve less instances to optimality, report higher computation times to prove optimality and higher average and worst gaps. The benchmark results are shown in Table 4. We observe again that both approaches reach TL for the large instances, except that *Benders* was able to solve the instance *S100-0-0.5Q* to optimality while *Cap1* was not. The small instances could all be solved to optimality by both approaches, and *Cap1* was faster in doing so, except for the instance *S50-150-0.75Q* for which *Benders* returned an optimal solution in almost half of the time. In addition, this table provides a first managerial insight, namely, that the upper bound values of the *Benders* approach are smaller for all instances. These values represent the combined objectives of uncovered waste units and transportation distances.

To analyze the savings gained by carrying two containers for each of the objectives separately, Fig. 6 presents for each of the datasets a boxplot of the improvement in distance (%) and the coverage ratio of solutions that were solved to optimality or near optimality by both approaches (99 of *S15*, 88 for *S50*, 20 for *S100*, and 5 for *S200*). More precisely, the distance improvement was computed as $(Dist_{Cap1} - Dist_{Benders})/Dist_{Cap1} * 100$ and the coverage ratio as $Cov_{Benders}/Cov_{Cap1}$, with distance values in meters and coverage values in %. Fig. 6(a) clearly shows an improvement in distance with vehicles of a capacity of two containers as opposed to vehicles of a capacity of one container. On average, transportation distances were 29.5% shorter, with 21.3% for dataset *S15*, 40.4% for dataset *S50* and 30.2% for dataset *S100*. For dataset *S200*, the average distances increased by 3.5%. In the five instances included for this dataset, the respective

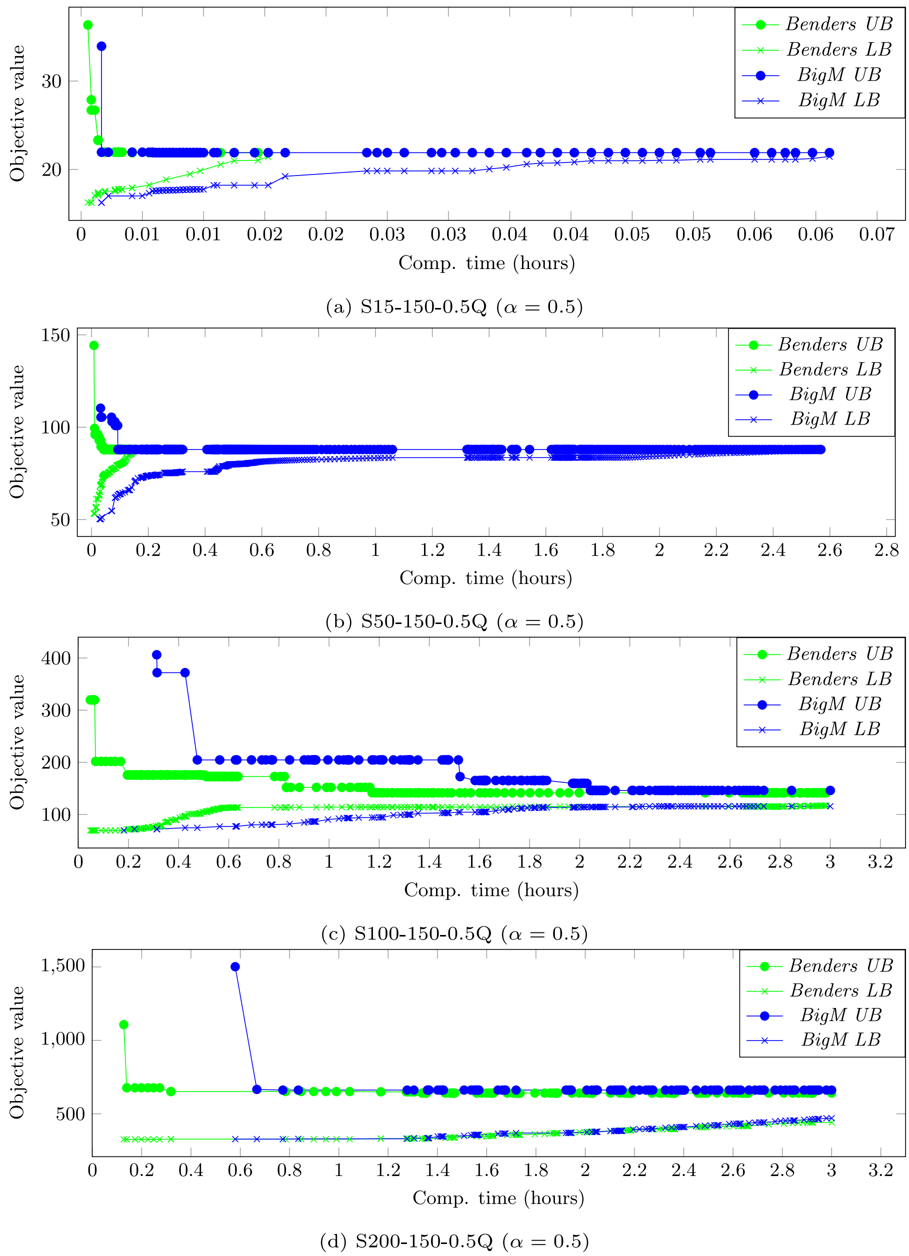


Fig. 4. Convergence graph of the *Benders* and the *BigM* approach for four benchmark instances.

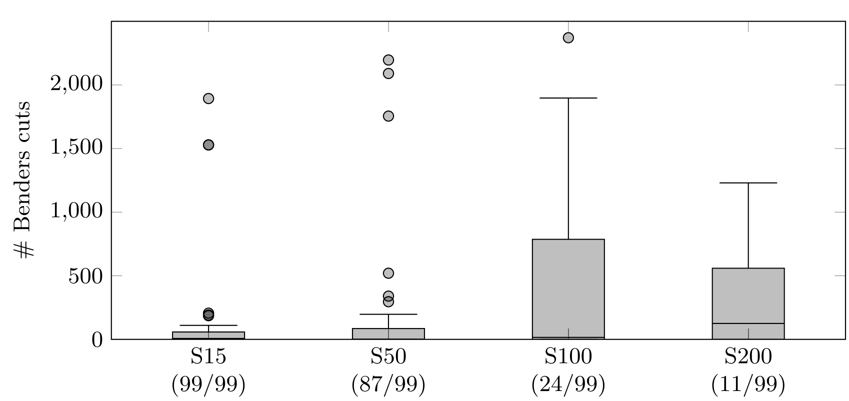


Fig. 5. Number of Benders cuts added to the master problem for instances solved to optimality or near optimality (gap $\leq 1\%$).

Table 3
Aggregated results of the *Cap1*-formulation and the *Benders* approach for each of the four datasets.

Dataset	*S15*		*S50*		*S100*		*S200*	
	Cap1	*Benders*	*Cap1*	*Benders*	*Cap1*	*Benders*	*Cap1*	*Benders*
# instances	99	99	99	99	99	99	99	99
# optimal	99	99	**99**	81	**32**	17	**5**	3
# nearly optimal	0	0	0	6	2	7	2	8
# opt. + nearly opt.	99	99	**99**	87	**34**	24	7	**11**
average gap (%)	–	–	–	3.8	**11.4**	14.9	**13.0**	20.2
worst gap (%)	–	–	–	15.7	**38.1**	50.1	**78.3**	84.0
average comp. time (s)	**1.1**	51.8	**295.0**	1408.9	**3114.1**	3616.8	–	–

Table 4
Results of the *Cap1*-formulation and the *Benders* approach for the benchmark instances.

Instance	Alpha	Comp. time (s)		Upper bound		Gap (%)	
($\|V\| - \rho - w^{\min}$)		*Cap1*	*Benders*	*Cap1*	*Benders*	*Cap1*	*Benders*
15-150-0.5Q	0.2	**0.8**	15.1	32.8	**27.1**	0.0	0.0
15-150-0.5Q	0.5	**1.1**	57.1	34.4	**21.9**	0.0	0.0
15-150-0.5Q	0.8	**1.0**	58.1	13.8	**8.8**	0.0	0.0
50-150-0.5Q	0.2	**1491.9**	2542.8	168.8	**106.5**	0.0	0.0
50-150-0.5Q	0.5	**502.0**	527.9	136.6	**87.9**	0.0	0.0
50-150-0.5Q	0.8	**87.0**	178.5	78.8	**58.6**	0.0	0.0
100-150-0.5Q	0.2	TL	TL	272.2	**164.0**	**15.5**	18.3
100-150-0.5Q	0.5	TL	TL	212.9	**141.5**	**9.8**	17.3
100-150-0.5Q	0.8	TL	TL	131.1	**97.4**	**1.8**	16.2
200-150-0.5Q	0.2	TL	TL	1269.0	**735.6**	**12.5**	20.7
200-150-0.5Q	0.5	TL	TL	960.8	**641.2**	**13.2**	31.1
200-150-0.5Q	0.8	TL	TL	579.8	**440.4**	**12.4**	35.3
50-150-1	0.5	**127.7**	1397.9	136.1	**87.1**	0.0	0.0
50-150-0.75Q	0.5	543.7	**289.1**	136.6	**87.9**	0.0	0.0
50-0-0.5Q	0.5	**12.5**	28.2	228.6	**182.0**	0.0	0.0
50-300-0.5Q	0.5	**1300.9**	3646.6	122.1	**78.4**	0.0	0.0
100-150-1	0.5	TL	TL	201.0	**126.7**	**4.0**	8.9
100-150-0.75Q	0.5	TL	TL	212.9	**143.7**	**8.7**	14.1
100-0-0.5Q	0.5	TL	**1933.5**	417.1	**337.2**	2.9	**0.0**
100-300-0.5Q	0.5	TL	TL	167.3	**100.6**	**7.7**	16.1

α-values were either 0.9 or 1, which means that more focus was given to the coverage objective than to the transportation distances. Furthermore, for three out of the five instances, only covering was considered with $\alpha = 1$. Looking at Fig. 6(b), the boxplot of dataset *S200* supports this observation, such that in these five instances (where focus was on coverage) both approaches were able to cover the same amount of waste, with the exception of one instance where *Benders* reached higher coverage. Coming back to dataset *S200* in Fig. 6(a), we conclude that the increase in distance for *Benders* is due to the focus on coverage. Considering the few other instances with negative distance improvement values (datasets *S15* and *S100*), we observe that coverage is always higher for *Benders*, such that the increase in distance was compensated by an increase in coverage. Finally, when ignoring temporarily the smallest dataset, we observe a general increase in the variance of the values for the bigger datasets that include less instances.

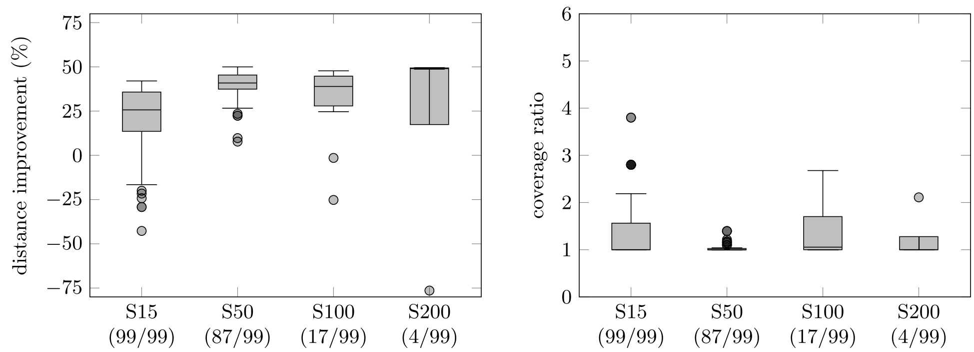
Fig. 6(b) shows that for almost all instances, the same amount of waste or more could be covered with vehicles of a capacity of two containers. On average, the coverage of carrying two containers is 1.2 times the coverage of carrying only one (corresponding to an increase in coverage of 20%), with the highest value observed being almost 6 times as high coverage for instance *S50-150-0.5Q* and $\alpha = 0$. For each of the four datasets, we report average values of 1.3, 1.1, 1.4, and 1.2, respectively. These average values above 1 clearly show that the savings in transportation distances were not achieved at expense of coverage. On the contrary, we even observe higher coverage values at shorter distances. Considering the few instances with coverage ratio below 1 (dataset *S50*), we observe that the respective α-values were all below or equal to 0.5, which means that more focus was given to the distance objective. Furthermore, for the instance with the most extreme coverage ratio below 1 (*S50-150-0.75Q*), only distance was considered ($\alpha = 0$).

7.4. Impact of including transportation distances in the objective

To analyze the impact of including transportation distances in the objective, Fig. 7 visualizes the optimal and nearly optimal solutions (gap $\leq 1\%$) provided by the *Cap1*-formulation for one example instance (*S100-0-1*). The curve shows a clear trade-off between the coverage (%) and the transportation distances (km), as expected in bi-objective problems. We observe the same trend in Table 5, which presents average distance and coverage values over instances solved to optimality or near optimality for each of the four datasets and each value of α. Please note that we aggregate the values for each line separately to get a more complete view with a high number of instances, but consequently the number of instances may differ from line to line. With an increase of α (more focus on coverage), the average coverage values increase at the cost of increased distance. For a few cases, the average coverage decreases compared to the previous value, which is due to the different number of instances considered in each line.

7.5. Impact of imposing minimum workload on container locations

As pointed out in Section 7.1, in this section, we analyze the impact of imposing a minimum workload ($w^{\min}$) on locations. Fig. 8 presents the upper bound values of the *Cap1*-formulation and the *Benders* approach for different values of $w^{\min}$ and α for dataset *S50* with $\rho = 150$. Note that all these instances were solved to optimality or near optimality by both approaches, except for the instance *S50-150-1* at $\alpha = 0.9$ for which the *Benders* approach reported a gap of 3.1%, which justifies using these instances for illustration. For each value of α, the first bar shows the upper bound obtained without imposing a minimum workload. Looking at each α value individually, we observe that the upper bound values generally increase when imposing a minimum



(a) Distance improvement (b) Coverage ratio

Fig. 6. Distance improvement and coverage ratio of vehicles with a capacity of two containers compared to vehicles with a capacity of one container for instances solved to optimality or near optimality (gap $\leq 1\%$).

Table 5
Average distance (km) and coverage (%) values over instances that were solved to optimality or near optimality (gap $\leq 1\%$) by the *Cap1*-formulation for each dataset and each value of α.

Alpha	*S15*		*S50*		*S100*		*S200*	
	Dist. (km)	Cov. (%)	Dist. (km)	Cov. (%)	Dist. (km)	Cov. (%)	Dist. (km)	Cov. (%)
ϵ	3.3	38.7	15.3	46.9	22.2	3.1	–	–
0.1	3.3	38.7	15.7	62.4	22.7	2.5	–	–
0.2	3.3	38.7	16.3	66.5	23.0	4.2	–	–
0.3	3.4	42.9	18.9	78.9	23.0	4.2	–	–
0.4	3.9	53.5	19.5	80.6	24.9	9.2	–	–
0.5	4.1	57.1	20.3	82.1	–	–	–	–
0.6	6.0	80.5	22.0	84.2	35.2	48.0	–	–
0.7	7.5	93.5	24.6	86.1	39.0	45.5	–	–
0.8	7.8	94.6	28.0	87.8	55.2	52.7	210.1	92.9
0.9	8.4	96.4	33.5	89.3	100.8	71.3	235.3	44.6
$1-\epsilon$	8.4	96.4	39.1	89.8	102.4	80.1	1274.7	100.0

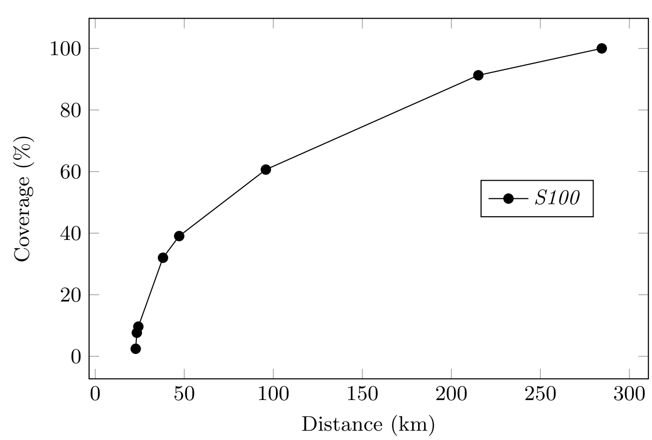


Fig. 7. Trade-off between coverage (%) and distance (km) of solutions provided by the *Cap1*-formulation for one instance of dataset *S100* with $\rho = 0$ and without *constraints* (1d).

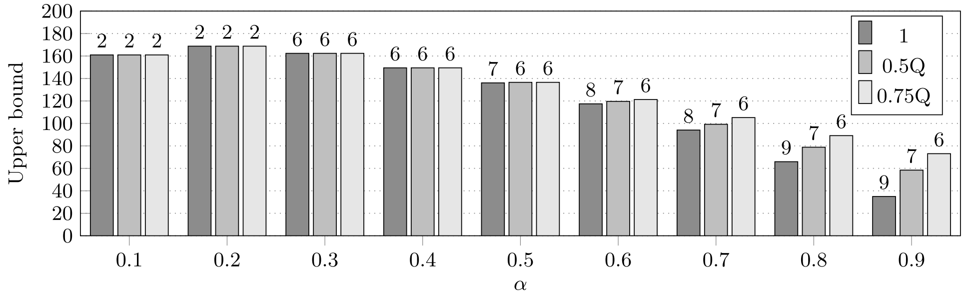
workload greater than 1. The labels on top of the bars indicate the number of containers located in each solution. They clearly show that a high required minimum workload leads to fewer containers (as stated in assumption *A6*). When we consider the whole range of the chart, we note that when the values of α are increased (more focus on coverage), the number of containers increases, which is most evident when not imposing a minimum workload.

7.6. Impact of satisfying demand of uncovered demand nodes on the location decision

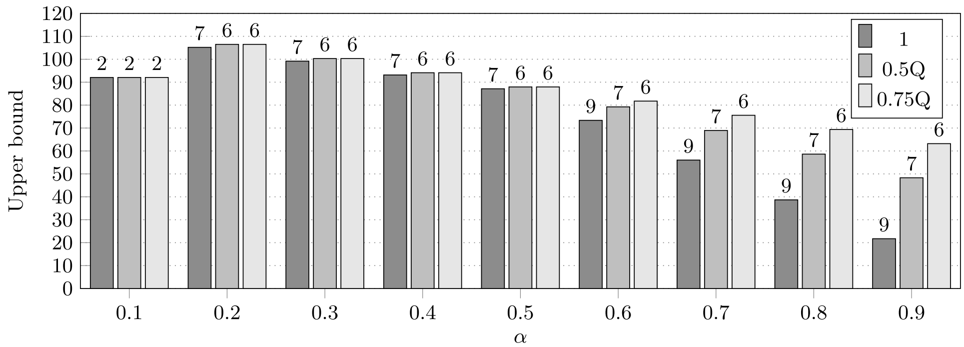
In this section, we verify our claim from Section 3, saying that the allocation of uncovered demand can have an effect on the location decision such that locations lie closer to the uncovered residential nodes. Fig. 9 visualizes the solutions of two example instances of dataset *S100* with $w^{\min} = 0.5$ and $\rho = 150$ (top) and $\rho = 300$ (bottom), respectively, resulting from two different formulations. Figs. 9(a) and 9(c) present the solutions given by our *Cap1*-formulation with $\alpha = 1$ (focus only on covering) in which uncovered demand is assigned and satisfied at container locations. Figs. 9(b) and 9(d) show the solutions given by an adapted *MCLP*-formulation, in which preferences, minimum workloads and capacities are considered, but demand of uncovered demand nodes is left unsatisfied. We refer the reader to Appendix A.1 for a complete formulation of this problem. We point out that the *Cap1*-formulation is the more relevant formulation in practice for waste collection, because even residents who are not formally covered by having a container location close to them still need to be assigned to a container. These figures clearly illustrate that with the *Cap1*-formulation, the container locations are dragged towards the uncovered residential nodes, while with the *MCLP*-formulation, containers are placed more centrally. Furthermore, we observe in Figs. 9(a) and 9(b) that the same number of locations are selected by both formulations whereas in Figs. 9(c) and 9(d) that number is lower in the *MCLP* solution compared to the *Cap1* solution. These effects can be explained by the different treatment of uncovered demand in the two formulations and the imposed capacity limits on container locations.

8. Conclusion

In this article, we formulated a capacitated MCLP and provided mathematical formulations for two covering problems inspired from a waste collection problem in which residents bring their waste to container locations. One problem considers vehicles with a capacity of one container, and the other considers vehicles with a capacity of two containers. We proposed a logic-based Benders decomposition approach for the latter problem and derived several managerial insights

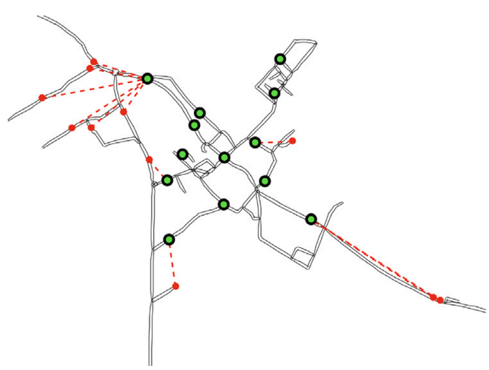


(a) Upper bound of the *Cap1*-formulation for each value of α and for each instance of dataset *S50* with fixed $\rho = 150$ and varying values for $w^{\min}$.



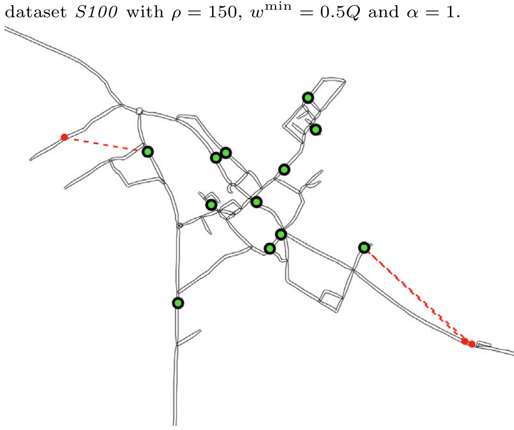
(b) Upper bound of the *Benders* approach for each value of α and for each instance of dataset *S50* with fixed $\rho = 150$ and varying values for $w^{\min}$.

Fig. 8. Impact of imposing different values of minimum workload $w^{\min}$ on the upper bound values for the *Cap1*-formulation and the *Benders* approach illustrated on dataset *S50* with $\rho = 150$ (the labels on top of the bars show the number of containers located).

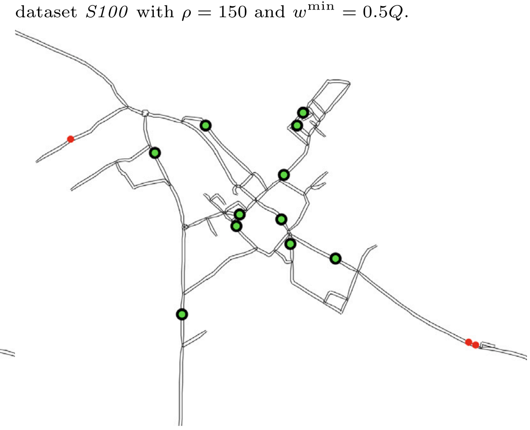


(a) Optimal solution of the *Cap1*-formulation for the dataset *S100* with $\rho = 150$, $w^{\min} = 0.5Q$ and $\alpha = 1$.

(b) Optimal solution of the *MCLP*-formulation for the dataset *S100* with $\rho = 150$ and $w^{\min} = 0.5Q$.



(c) Optimal solution of the *Cap1*-formulation for the dataset *S100* with $\rho = 300$, $w^{\min} = 0.5Q$ and $\alpha = 1$.



(d) Optimal solution of the *MCLP*-formulation for the dataset *S100* with $\rho = 300$ and $w^{\min} = 0.5Q$.

Fig. 9. Optimal solutions of the *Cap1*- and *MCLP*-formulations for two example instances to illustrate the impact of satisfying demand of uncovered demand nodes.

from the computational experiments conducted on a set of real-life instances. The results show that the *Benders* approach clearly outperforms the *BigM*-formulation with respect to the number of optimal and near optimal solutions (225 and 198 in total, respectively), the average and worst gaps (17.6% and 27.3%, and 92.4% and 100%, respectively), and the average computation times (232.6 s and 833.8 s, respectively). The benefit of vehicles with a capacity of two containers was demonstrated by the decrease in transportation distances (29.5% on average) and the increase in coverage (20% on average). Furthermore, we observed a clear trade-off between the coverage (%) and the transportation distances (m), as expected in bi-objective problems. Finally, imposing a minimum workload on locations has the effect of placing fewer containers, resulting in lower coverage, and satisfying uncovered demand leads to more selected container locations which are dragged towards the uncovered residential nodes.

This paper opens several interesting paths for future research. To better represent reality, uncertain demand could be modeled by a set of discrete scenarios with the goal to minimize the worst or average cost (i.e., uncovered demand and transportation cost) over all of them. Furthermore, other distributions of waste over time could be considered, so that days are associated with different preferences as residents might prefer working days over weekends to dispose their waste. To define the preference ordering of the container locations for residents, an interesting extension could be the characterization of a convenience measure that includes indicators related to candidate locations other than walking distance, such as proximity to points of interest (e.g., train station, schools, playgrounds) and walkability (i.e., a measure of how pedestrian-friendly an area is).

CRediT authorship contribution statement

Vera Fischer: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization, Funding acquisition. **Sanne Wøhlk:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization, Supervision, Project administration.

Data availability

We have shared the link to the data in Section 7.1 of the manuscript.

Acknowledgments

The authors gratefully acknowledge the support of Innosuisse, Switzerland under grant 36157.1 IP-EE. They also highly value the involvement of Schwendimann AG in conducting practical experiments and providing the associated data.

Appendix

A.1. MCLP formulation

This formulation is based on the MCLP formulation presented in Farahani et al. (2012) and on the notation defined in Section 3. The binary variable y_j is equal to 1 if a container is established at location $j \in F$, and the binary variable a_{ij} is equal to 1 if residential node $i \in V$ is allocated to the location j. The objective function (A.1a) defines the goal of maximizing the total waste that has been assigned to acceptable locations. Constraints (A.1b) ensure that the waste of each residential node is only assigned to a location that has been opened, and constraints (A.1c) impose that each residential node is allocated to its most preferred location. Constraints (A.1d) enforce the minimum workload $w^{\min}$ on open locations. Constraints (A.1e) define that over the whole planning period, a location j is limited to the capacity Q times the number of days T. Constraints (A.1f) state that at most p containers can be placed at the candidate locations. Constraints (A.1g) define the domains of the variables. We refer to this formulation as the *MCLP*-formulation.

$$\max \sum_{i \in V} \sum_{j \in F_i^+} w_i a_{ij} \tag{A.1a}$$

$$\text{s.t. } a_{ij} \le y_j \qquad \forall i \in V, j \in F_i^+ \tag{A.1b}$$

$$\sum_{j' \in F_i^+ : \pi(i,j') > \pi(i,j)} a_{ij'} \le 1 - y_j \qquad \forall i \in V, j \in F_i^+ \tag{A.1c}$$

$$\sum_{i \in V : j \in F_i^+} w_i a_{ij} \ge w^{\min} y_j \qquad \forall j \in F \tag{A.1d}$$

$$\sum_{i \in V : j \in F_i^+} w_i a_{ij} \le TQ \qquad \forall j \in F \tag{A.1e}$$

$$\sum_{j \in F} y_j \le p \tag{A.1f}$$

$$a_{ij}, y_j \in \{0, 1\} \qquad \forall i \in V, j \in F \tag{A.1g}$$

References

Adeleke, O.J., Olukanni, D.O., 2020. Facility location problems: models, techniques, and applications in waste management. Recycling 5 (2), 10.
Church, R.L., Murray, A., et al., 2018. Location Covering Models. Springer.
Church, R., ReVelle, C., 1974. The maximal covering location problem. In: Papers of the Regional Science Association, Vol. 32, No. 1. Springer-Verlag, pp. 101–118.
Cubillos, M., Wøhlk, S., 2020. Solution of the maximal covering tour problem for locating recycling drop-off stations. J. Oper. Res. Soc. 1–16.
Current, J.R., Schilling, D.A., 1994. The median tour and maximal covering tour problems: Formulations and heuristics. European J. Oper. Res. 73 (1), 114–126.
Current, J.R., Storbeck, J.E., 1988. Capacitated covering models. Environ. Plan. B Plan. Des. 15 (2), 153–163.
Farahani, R.Z., Asgari, N., Heidari, N., Hosseininia, M., Goh, M., 2012. Covering problems in facility location: A review. Comput. Ind. Eng. 62 (1), 368–407.
García, S., Marín, A., 2019. Covering location problems. In: Laporte, G., Nickel, S., Saldanha-da Gama, F. (Eds.), Location Science. Springer, pp. 99–119.
Ghiani, G., Laganà, D., Manni, E., Triki, C., 2012. Capacitated location of collection sites in an urban waste management system. Waste Manage. 32 (7), 1291–1296.
Hooker, J.N., Ottosson, G., 2003. Logic-based benders decomposition. Math. Program. 96 (1), 33–60.
Kolen, A.W., Tamir, A., 1984. Covering Problems. Econometric Institute.
Laporte, G., Nickel, S., Saldanha-da Gama, F., 2019. Location Science. Springer.
Nagy, G., Salhi, S., 2007. Location-routing: Issues, models and methods. European J. Oper. Res. 177 (2), 649–672.
Prodhon, C., Prins, C., 2014. A survey of recent research on location-routing problems. European J. Oper. Res. 238 (1), 1–17.
SAAG, S.A., 2023. Patent alpenluft. URL: https://www.system-alpenluft.ch.
Schilling, D.A., 1993. A review of covering problems in facility location. Locat. Sci. 1, 25–55.
Toregas, C., Swain, R., ReVelle, C., Bergman, L., 1971. The location of emergency service facilities. Oper. Res. 19 (6), 1363–1373.
Williams, H.P., 2013. Model Building in Mathematical Programming. John Wiley & Sons.