Computers & Industrial Engineering 178 (2023) 109142

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Hierarchical clustering-based algorithms for optimal waste collection point locations in large-scale problems: A framework development and case study

Adam Viktorin [a], Dušan Hrabec [a,*], Vlastimír Nevrlý [b], Radovan Šomplák [b], Roman Šenkeřík [a]

[a] Faculty of Applied Informatics, Tomas Bata University in Zlín, Nad Stráněmi 4511, 760 05 Zlín, Czech Republic
[b] Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic

ARTICLE INFO

Keywords:
Criteria-based clustering
Collection network design planning
Waste management
MILP reduction techniques
Waste container location
Computational complexity

ABSTRACT

The cities face the challenge of optimizing investments in waste management to meet EU standards while maintaining economic affordability. One of the issues is the optimal location for specialized waste collection points. The main target is to find the lowest number of collection points that would still attain waste production, and the average walking distance to the waste container would be kept beneath the tolerable limit for citizens. The population density and waste production vary over city parts; thus, the need for specialized containers in more populated city centers, industrial zones, or household streets differs. This paper develops a new computational approach providing a robust generalized decision-support tool for waste collection bin location and allocation. This task leads to a mixed-integer linear program which is not solvable for larger cities in a reasonable time. Therefore, hierarchical clustering is applied to simplify the model. Two strategies for solving waste bin allocation (for multiple variants of the model formulation) are implemented and compared – sub-problem definition and representative selection approaches. The resulting framework is tested on the artificial instance and a few case studies where the structure and properties of results are discussed. The combination of presented approaches proved to be appropriate for large-scale instances. The representative selection approach leads to a better distribution of containers within the area in the single-objective model formulation.

1. Introduction

Waste Management (WM) has become one of the most challenging issues due to urban development, population growth, and lifestyle changes (Darmian et al., 2020). It presents an important application area where the planning of logistics problems is typically used as the increased consumption levels are causing an exacerbation of the problem (Bing et al., 2016). The waste bin-related network design problems needed either for optimization of existing collection networks or network design of newly separated waste types is, especially computationally, a challenging problem asking for new operational research approaches (Olmez et al. 2022). Clustering recently provides a suitable tool to reduce the computational complexity of decision-making problems and to solve them reasonably with reasonable results (Caramia & Pizzari, 2022).

Logistics and clustering-related WM problems can be seen from two different perspectives. First, issues of regional collection areas from the perspective of waste processing facilities, where a node in the collection network is a municipality or micro-region (Antunes et al., 2008). Second, waste collection of bins and containers from the municipality's perspective. The waste containers are often aggregated into groups. For example, waste collection is organized at the street level, where all households on the same street are serviced by the same vehicle (Zbib & Laporte, 2020). In both cases, clustering reduces the complexity of real-size problems, typically large-scale mixed integer problems. Rarely the network detail is based on individual bins and containers (Liang et al., 2022).

The primary step towards effective waste collection in a municipality and subsequent transport to a waste processing facility is the location of waste containers and bins (Matušinec et al., 2022). Herein, some research has already been done, especially on the socio-economic level dealing with various objectives when locating the waste collection containers (Tralhão et al., 2010). The formulation of a particular mathematical model strongly influences the problem's computational complexity, especially when the model reflects and optimizes multiple criteria (Nevrlý et al., 2021). A typical problem in this type of

* Corresponding author.
E-mail addresses: aviktorin@utb.cz (A. Viktorin), hrabec@utb.cz (D. Hrabec), vlastimir.nevrly@vutbr.cz (V. Nevrlý), radovan.somplak@vutbr.cz (R. Šomplák), senkerik@utb.cz (R. Šenkeřík).

https://doi.org/10.1016/j.cie.2023.109142
Received 13 May 2022; Received in revised form 1 February 2023; Accepted 2 March 2023
Available online 8 March 2023
0360-8352/

application is the high number of integer (often binary) variables and restrictions that go against each other and are so difficult to achieve, which leads to the extremely complex computational task not only in bigger cities (Chavez et al., 2021). A common solution approach is to divide cities into smaller logical parts that can be separately solved to achieve optimality. This approach is used across all application fields, but it has several problematic points:

- Inaccuracy of the solution can arrive on the boundaries of separated parts.
- Some parts can still be too large to be solved to optimality. It can typically happen in the city centers, where the division into smaller parts is not clear
- Another issue arises when solving the outlying localities. It is worth considering placing an option of separate collection for some waste types due to expansive waste collection services.
- Another issue that can arise is the difference between a real infrastructure of sidewalks and a linear distance (by air); this can occur, e. g., in the case of a block of flats with front and back doors. How should the task then be partitioned? Does it influence the clustering algorithm and the solution, respectively?

The location of collection containers is a current topic in the field of WM, not only due to the societal pressure to increase the rate of separation and recycling. Solving complex tasks is thus in the interest of current research.

This paper follows previous studies on the optimal location of waste bins and containers when optimizing the number of bins and the walking distance of waste producers to the assigned waste bin (Matušinec et al., 2022, Nevrlý et al., 2021). This paper develops and suggests a robust generalized decision-support tool for waste collection bin location and allocation. Typically, this task leads to a mixed-integer linear program (MILP) which is not solvable for larger problems (e.g., cities) in a reasonable time. Therefore, hierarchical clustering is applied to simplify the model. Two strategies for solving waste bin allocation are implemented and compared – representative selection (Section 3.4) and sub-problem definition (Section 3.5) approaches.

The rest of the paper is organized as follows. Section 2 provides a literature review on the various bin location problems, clustering algorithms in general and existing clustering approaches applied in the WM. Section 3 describes the model, which is used in two strategies for waste bin network generation. Methods are presented in a detailed step-by-step guide. Section 4 demonstrates new approaches to selected areas and comments on results from the practical point of view. Finally, the paper concludes with Section 5, which also features further applicability of the presented approaches and discusses future research possibilities.

2. Literature review

This section aims to identify articles focusing on related research fields, such as waste network design optimization, operational research approaches, and computational optimization, emphasizing clustering methods.

2.1. Waste management in operations research

Operations research has many applications in solid waste management. One of the most important applications is the design of a waste collection network in the form of a multi-echelon logistics structure, including the collection centers location and allocation of demand areas using mathematical models (Eiselt & Marianov, 2015), optimization algorithms (Rabbani et al., 2018), and multi-criteria decision-making methods (Soltani et al., 2015). Following an operations research perspective, the tasks are divided into strategic, affecting the long-term, and tactical/operational, affecting the medium-short term (Caramia & Pizzari, 2022). Ghiani et al. (2014) provide a survey that studies such a division. The disposal collection and location problem has been studied both from a strategic perspective (Bing et al., 2016) and from a tactical perspective (Eiselt & Marianov, 2014). Another type of waste collection network design-related decision is operational, e.g., the transportation of municipal waste in the urban environment also releases environmental pollutants in case of improper performing waste transportation operations (Expósito-Márquez et al., 2019).

However, most reviewed articles focus only on single-level single-objective models, with just a few delving into multi-objective optimization. On the other hand, most operations research applications in municipal solid waste management involve a location problem to find strategic decisions (Ghiani et al., 2014). Some research papers employing multi-objective optimization have been published. These involve strategic decisions such as Waste-to-Energy facility location (Abdallah et al., 2021), designing a network to minimize transportation and facility costs, land use stress, and impact on public health (Olapiriyakul et al., 2019), or even three joint cost-minimizing objectives being the cost of establishing collection centers and collecting waste and environmental and social impacts (Darmian et al., 2020).

Therefore, the upcoming sections attempt to review the waste bin location problems from the mathematical modeling perspective, and the computational solution approaches perspective.

2.2. Waste management logistics: Bin location problems

In recent years, with the emergence of environmental problems, the importance of considering the environmental dimension in objective functions has been demonstrated. It has become one of the researchers' concerns alongside considering economic factors (Rabbani et al., 2020). Over the past decades, WM tasks have become more complex due to rapid urbanization, which has led to the emergence of different optimization techniques and heuristic methodologies. Logistics-related WM problems belong to the reverse logistics field that deals with the flow of products or goods from the consumer to an earlier stage of the supply chain. Supply chain management usually needs to solve the location problem of network components first (Li, 2019).

Location and location-allocation optimization are essential parts of each waste collection problem (Sbihi & Eglese, 2010). The appropriate collection service and frequency (i.e., the number of bin emptying or visits to individual houses in door-to-door systems) are strongly linked to the capacity allocation to collection sites (i.e., the size and number of bins present). Given the demand for a collection site, by increasing (reducing, respectively) the number of bins located there, one may reduce (increase, respectively) the required service frequency of the site. Varying capacity allocation criteria impact the fixed investment costs associated with service setup (Hemmelmayr et al., 2014). Optimizing waste bin locations and vehicle routings can also be acquired by combining mathematical algorithms with GIS (Erfani et al., 2017). Using GIS, another analysis of the municipal solid waste (MSW) repository location was performed based on street width and population density (Oliaei & Fataei, 2016). The results made it possible to evaluate the best capacity solution for the given localities. The GIS-based approach was also used to re-evaluate the containers' total number and site for separately collected wastes (Zamorano et al., 2009). The excessive number of containers may significantly raise purchase and collection costs. Methods used in GIS systems do not include clustering or size reduction algorithms, so the tailor-made approaches can solve even large-scale tasks and specific constraints.

Problems combining the decisions on location (where to locate) and allocation (size and number of bins to locate) are often called location-allocation problems. Typical location-allocation models for waste containers turn out to be MILPs (Ghiani et al., 2012). Thus, currently, some research has utterly dealt with the only bin location problem (Herrera-Granda et al., 2019). For example, Nevrlý et al. (2021) focus on the optimal location of municipal waste containers while searching for a trade-off between various criteria. Matušinec et al. (2022) developed a

Table 1
Most recent (2018–2022) and significant research on bin sitting.

	Label	GIS	Optimization	Heuristics	Waste	Objective	Max. no. nodes
Letelier et al. (2022)	A	yes	yes	no	MSW, recyclable	single	230
Matušinec et al. (2022)	B	no	yes	no	WCO	single	thousands
Yalcinkaya and Uzer (2022)	C	yes	yes	no	–	single	245
Slavík et al. (2021)	D	yes	no	no	bio	single	21,595
Mahéo et al. (2023)	E	yes	yes	yes	–	single	–
Gläser and Stücken (2021)	F	no	yes	yes	–	single	288
Nevrlý et al. (2021)	G	no	yes	no	Plastic	multi	3,000
Adeleke and Ali (2021)	H	no	yes	yes	–	single	500
Blazquez and Paredes-Belmar (2020)	I	no	yes	yes	–	single	1,345
Rathore et al. (2020)	J	yes	yes	no	–	single	tens
Rossit et al. (2020)	K	no	yes	yes	–	multi	115
Toutouh et al. (2020)	L	no	yes	yes	–	multi	115
Shi et al. (2020)	M	yes	yes	yes	e-waste	multi	100
Nevrlý et al. (2019)	N	no	yes	no	MMW	multi	–
Rossit et al. (2019a)	O	yes	yes	no	MSW (dry and humid)	multi	88
Rossit et al. (2019b)	P	yes	yes	no	MSW	multi	88
Barrena et al. (2019)	Q	no	yes	yes	–	single	39
Herrera-Granda et al. (2019)	R	yes	yes	no	MSW	multi	999
Aka and Akyüz (2018)	S	yes	yes	no	recyclable	multi	29
Yaakoubi et al. (2018)	T	no	yes	yes	MSW	single	75
Vu et al. (2018)	U	yes	no	no	MSW	single	25

Note: WCO = waste cooking oil; MMW = mixed municipal waste.

Table 2
Application of individual objectives and constraints.

Paper label	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Objectives minimizing																					
number of collection points/bins	✓	✓	x	x	✓	–	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	x	x	x
walking distance (different formulations)	x	x	✓	x	x	–	✓	x	x	x	✓	✓	x	✓	x	x	x	✓	x	x	x
number of address points within walking distance of the given value	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
the estimated cost of bin allocation and operation within a certain period (e.g., five years)	x	x	x	x	x	x	✓	x	x	x	✓	x	x	x	✓	✓	x	x	x	✓	x
Objectives maximizing																					
number of address points within a threshold distance (e.g., 80 m)/maximal coverage	x	x	x	✓	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	✓
total utilization of collection points	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Constraints																					
the capacity of collection points/bins	✓	✓	x	✓	✓	✓	✓	✓	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	x	✓	x
maximal number of collection points or bins	x	x	x	✓	x	x	✓	✓	✓	✓	x	x	✓	x	✓	✓	✓	✓	x	x	✓
maximal number of bins	x	x	✓	x	x	x	✓	✓	✓	✓	x	x	✓	x	✓	✓	✓	✓	x	x	x
minimal utilization of collection points	x	x	x	x	x	x	✓	x	x	x	x	x	x	✓	x	x	x	x	x	x	x
maximal average walking distance	x	✓	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
maximal individual walking distance	x	x	x	x	x	✓	x	✓	x	x	✓	✓	x	x	✓	✓	x	✓	x	x	x
maximal volume-weighted walking distance	x	x	x	x	x	x	✓	x	x	x	x	x	x	x	x	x	x	x	x	x	x
restriction on the number of address points within walking distance over a given value	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
relations between bins for different commodities	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

*Note: ✓=yes; x = no; -=no information.

decision-making approach for the optimal location of cooking oils and fat waste bins while optimizing the walking distance from waste producers to the nearest collection site and the number of containers. Problems are formulated and solved with various objective functions and limiting constraints while applying different solutions. The overview and use of specific modeling attributes in recent research are summarized in Table 1.

Articles that did not specifically describe the detailed features of the task or the resources used were not included in the review because their contribution is insignificant for the purposes of this research. The way the problem was solved was divided into three categories – GIS, (mathematical) optimization, and heuristics. For simpler problem formulations, GIS can serve as a quality tool for quick analysis of collection point locations. However, its use is often charged by purchasing licenses or by subscription. Unfortunately, GIS offers only a limited selection of criteria for finding the optimal design. In some papers, this tool is used only as a source of input data or in the pre-processing phase. The second part defines the articles where a mathematical optimization model was formulated and it was subsequently solved by conventional exact algorithms. In some cases, some form of heuristics had to be used. Most often, the objective function is defined as single criteria, but even multi-objective formulation enjoys great interest. However, case studies are usually done only on small to medium-sized tasks, which significantly limits the applicability of the presented approaches. When the task is already tested on a larger instance, it is only a single-criteria problem with limited complexity of constraints, as shown by the detailed analysis in Table 2. Common objectives are the walking distance to the nearest collection point and the number of collection points/bins. The aim is to minimize these objectives, and an appropriate compromise must be found in the case of the multi-objective formulation. This may be based on the requirements and financial capabilities of the contracting authority of the study. The combination of objectives can cause an increase in walking distance for a non-negligible number of waste producers from the outlying localities. Herein, the construction of the individual walking distance criterion seems to be an option.

Table 2 shows the most frequently used criteria for optimization, both in the form of the objective function and in the form of constraints. The review also makes clear what trends further research will follow. The models do not consider the continuity of collection points regarding different waste fractions. The selected collection point should be used for several types of waste, with each waste defining its own maximum average walking distance. Suppose the (average) walking distance condition is constructed via a constraint in the mathematical model. In that case, it can cause the walking distance condition to be satisfied, but the volume of waste in containers is decreased, which is an undesirable effect. Therefore, it is necessary to specify the individual conditions to be mutually consistent. The utilization of collection points should also be considered regarding the current and future household separation rate of individual waste streams. For more complex types of tasks or large instances, there is no universal tool for approaching a problem's reduction and solution.

Planning of waste containers number and locations is a long-term and permanently solved problem to a considerable extent. Due to ongoing housing construction and urban development and growth in the number of waste fractions separated and collected, it is necessary to react and increase the collection capacity and the density of the collection network. Due to the constant maintenance of the infrastructure (repair of roads, sidewalks, hot water, gas, sewerage, parking), adjustment of the location within the city is necessary almost continuously. Regarding prohibited current locations, the key is the capacity of collection points. Both cases should be solvable in a reasonable time with acceptable precision to react and re-plane the collection network, possibly fixing most of the collection infrastructure. Different task formulations have different natures of accuracy and computation time. The aim is to provide a basis for a procedure for solving real case studies.

2.3. Clustering methods and waste bin location problems

Cluster analysis simplifies too complex or too large problems (Ordoñez et al., 2017). Typical logistics applications include location problems, allocation problems, network flow, or collection tasks, where the data are mostly aggregated according to the legislative separation of the area (Moskvichev et al., 2021). Clustering algorithms can be functionally split into a known order and unknown order. While the number of clusters is used as an input parameter in the known-order variant, unknown-order algorithms do not require the modulation order as input, instead of relying on factors such as density to determine the number of clusters (Pla-Sacristán et al., 2019). The most known algorithms that belong to these categories are as follow. For the unknown-order variant, it is the density-based spatial clustering of applications with noise (the so-called DBSCAN), ordering points to identify the clustering structure (OPTICS), or hierarchical clustering. For the known-order variant, it is k-means, k-medoids, or fuzzy c-means (Mouton et al., 2020). The classical clustering algorithms can also be classified into the following several kinds: partitional clustering, hierarchical clustering, density-based clustering, grid clustering, and the model algorithm (Kim et al., 2009; Li, 2019). The partition clustering algorithm, aimed at the database object, calculates the distance from all the samples to the cluster center (Xu et al., 2020).

Some studies proposed heuristic algorithms based on the nearest distance or clustering. The clustering of nodes using the k-medoids method was proposed by Mokhtarzadeh et al. (2021). K-medoids is one of the two most famous algorithms for clustering data; the other is k-means. However, as was mentioned above, these algorithms separate data into predefined k different groups, attempting to minimize the distance between nodes in each group. Hierarchical clustering is a suitable approach for the so-called (multi-criteria) territory partitioning (Lidouh & De Smet, 2016). Hierarchical clustering has also been applied for territory partitioning problems considering a maximum within-cluster distance, i.e., clusters where all distances were less than the chosen maximum transportation distance (Laasasenaho et al., 2019).

Combining clustering and logistics decisions has been widely applied (Wang et al., 2018). Hemmelmayr et al. (2014) developed a model and algorithm for integrated bin allocation and vehicle routing planning. They determine the service frequency and the days of visit associated with this service frequency. The solution approach combines an effective variable neighborhood search metaheuristic for the routing part with a MILP-based exact method for the bin allocation part. They follow a route-first, cluster-second method initially proposed by Beasley (1983). More recently, Jammeli et al. (2021) provided a bi-objective model with randomness to handle the optimization of waste collection. Due to many variables and the high degree of uncertainty, it was impossible to solve the problem by an exact algorithm in a reasonable computational time.

Recently, clustering methods have been applied to reduce the computational complexity of waste bin location problems, typically leading to mixed-integer optimization tasks. Caramia and Pizzari (2022) used fractional programming to solve two strategic objectives related to the costs incurred and utility generated in servicing customers when clustering, location, and allocation are performed on the supply chain associated with the considered WM problem. In their problem formulation, the municipality, as a decision-maker, organizes citizens and collection centers through a clustering method, where a cluster of citizens is assigned to one specific collection center.

The review does not provide any reason to prefer special approaches regarding subsequent optimization. No link to the character of the territorial division was analyzed either, which identifies a gap for further research. Usually, k-means is used where the number of clusters is specified by the input, while the selected number is related to the task solvability.

2.4. Novelty and contribution

This paper follows the previous studies (Matušinec et al., 2022, Nevrlý et al., 2021) on the optimal location of waste bins and containers when optimizing two contradictory objectives, the number of bins and the walking distance of waste producers to the assigned waste bin. Such optimization problem, especially in larger cities (large-scale problems), leads to computationally extremely complex tasks (Matušinec et al., 2022). The calculation time of each task depends on many factors. One of the biggest influences is the selected capacity at individual sites in connection with the total waste production in the network. The waste generation rate is derived from the specified frequency of the collection cycle. By defining these conditions, it is possible to calculate the theoretical value of the minimum number of collection points. The walking distance condition then creates pressure to open more collection points.

The literature review above shows that using other tools like GIS is not possible, especially when defining multiple objectives or complex constraints. Using the discussed modeling ideas in combination with clustering provides a suitable approach to reduce the complexity of the computational problems and so to solve the problem in a reasonable time. The main contribution of this work is to propose a novel approach to reduce the complexity of bin location problems by the following approaches that are compared in this paper:

1. Applying a hierarchical clustering for selecting uniformly located representatives and the number of edges (the edge represents assigning a waste producer to one specific bin) is also significantly reduced. The aim is to create the maximal number of clusters with one representative; the total number of representatives is solvable for the whole area (Section 3.4).
2. Applying a hierarchical clustering for splitting the complex problem into solvable sub-problems such that the sub-problems are similarly complex considering all possible edges. The aim is to create a minimal number of clusters such that the largest cluster (with the highest number of nodes) is still solvable with exact methods (Section 3.5) or to define the clusters' size, which the previous approach can solve.

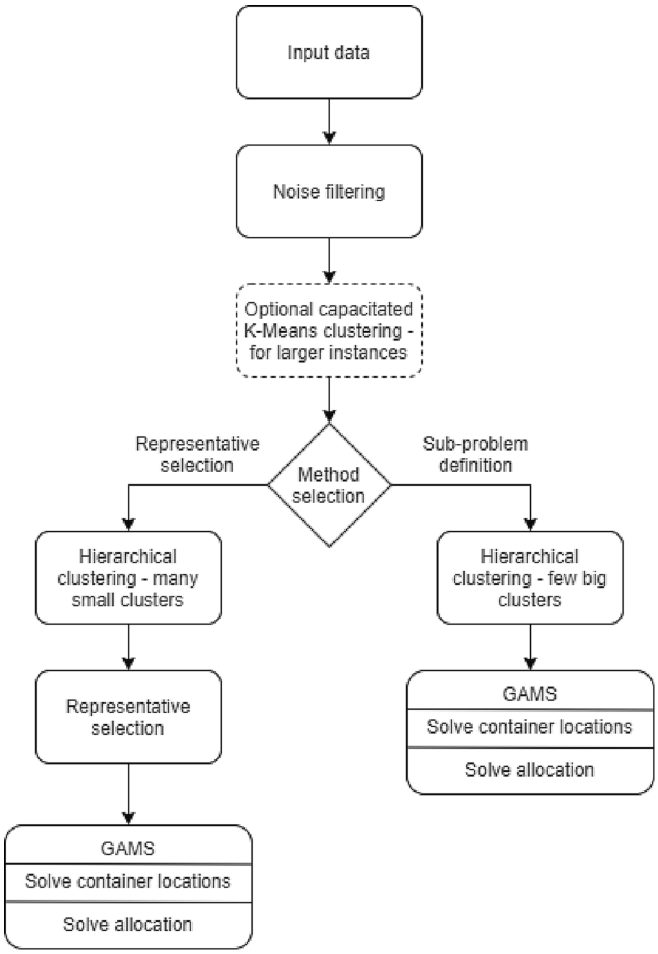

Fig. 1. Scheme of applied steps.

The combination of these approaches was even proved to be the best option for large-scale instances. The approach also considers noise filtering prior to the optimization, which slightly reduces the task size, and especially, it bans address points by providing separate containers in sparsely populated areas. The filtering is performed through the waste quantity evaluation in the predefined radius.

Therefore, the main contribution of this paper is to develop new computational approaches to the waste container location problem with specific constraints. More specifically, the paper provides suitable optimization tools to solve the large-scale problems on waste bin network design when optimizing the number of bins and average walking distance. Especially two different clustering approaches are tested and analyzed in combination with a commercial optimization solver (CPLEX) that is unable to solve real-scale problems without a method as clustering used.

3. Methods and algorithms

The optimization model used in the proposed approaches is adapted from Matušinec et al. (2022) and Nevrlý et al. (2021); see Appendix A. The single-objective formulation was demonstrated for the fat and cooking oil wastes. It considers the total number of containers as the objective function while restricting the average walking distance between addresses (buildings associated with a specific number of occupying inhabitants) and assigned collection points. Each collection point has a predefined capacity, which cannot be exceeded by the cumulative waste generation of assigned address points. The model allows situating of containers directly in front of all address points. Since the model is formulated as a MILP with a high number of binary variables and constraints, it is computationally demanding for larger instances. However, appropriately chosen simplifications can help significantly speed up or even enable the calculation.

3.1. Workflow of approaches

Approaches for simplification will be dealt with in this section by applying certain clustering steps. Fig. 1 defines all steps of tested algorithms. It starts with noise filtering, eliminating isolated points in terms of cumulative waste production in the neighborhood. Then, it is divided into two branches, in which different approaches are applied to simplify the original task. In the end, the problem is always solved by the exact method within the GAMS modeling environment. The solution is always divided into two steps – container location and then the allocation of address points to the collection points so that the walking distance is minimal. The task reduction approaches are described in detail in the following text.

3.2. Noise filtering

Noise in this task is defined as address points that are too distant from the rest of the network. These are detached dwellings, often also cottage areas or smaller clusters of family houses. It is not desirable to allocate bins in such areas due to insufficient cumulative waste production and the high costs of collecting these isolated and almost empty bins. Therefore, the noise was defined by the criterion for each address point as follows:

- If the sum of waste in the radius (given by mean walking distance) is lower than, e.g., 10% of the container capacity, the address point is identified as noise. A different walking distance can be targeted for each type of waste, as well as the minimum utilization of the allocated capacity of the collection point. In this regard, consultation with WM experts is always appropriate.

This condition will reduce the total number of points considered and ensure that bins are not distributed in sparsely populated areas.

3.3. Optional capacitated K-means clustering

This step is optional for larger problem instances. If the time complexity of the whole task is too high, the instance can be divided into smaller instances with a size constraint. The Capacitated K-Means algorithm (Geetha et al., 2009) was selected thanks to its ability to produce constraint-sized clusters and lower computational complexity. This approach was used for the second case study presented in this paper (see Section 4.5).

3.4. Representative selection

The first approach tries to reduce the total number of edges within the network. Since the graph is defined as bipartite, it connects all address points with all candidate locations for the container. A similar variant of reducing the size of the task is to limit it to a certain percentage of the nearest candidates for collection points for each address point. Usually, each address point can be at the same time also the candidate location for the container. The number of edges thus increases with the square of the number of address points. This method will thus seek some representatives, which will reduce the number of candidate locations. It will still be allowed to place the container in front of each address point if it reaches certain self-sufficiency (no other edges will be assigned). The selection of these representative sites will be made using a modified hierarchical method that considers the details of waste production and other aspects. It can be defined as updated agglomerative hierarchical clustering with Ward's method (Ward, 1963). Ward's method has been chosen because it can generate compact clusters.

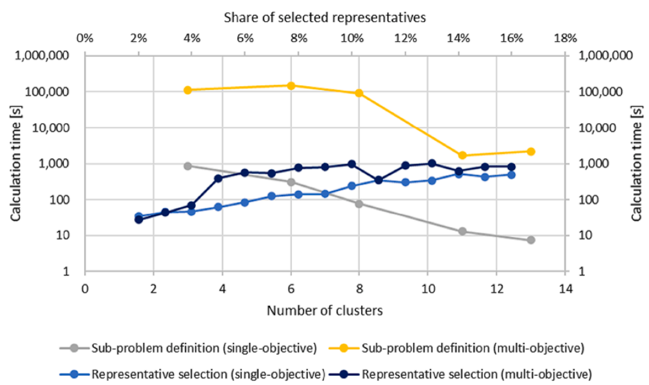

Fig. 2. Calculation times for artificial instance – different approaches and formulations.

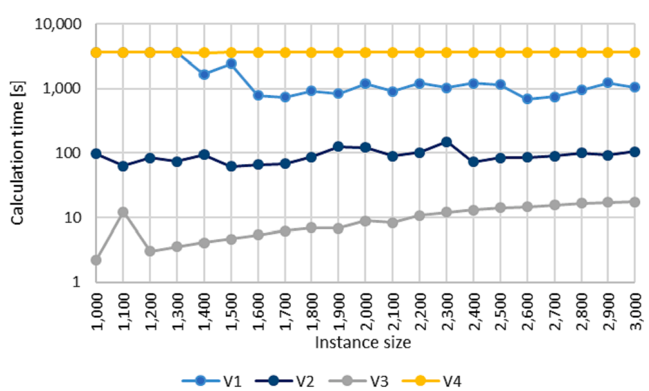

Fig. 3. Calculation times for artificial instances of varying sizes – different single-objective variants.

Moreover, it performs well in separating clusters if there is noise between them. Updates in this procedure represent a repeated application of the method until the goal is reached. Each iteration follows these rules:

- Stops at the given number of clusters – equal to the number of representatives.
- The maximum size of the cluster is limited to 100% of the container capacity. Merging of two clusters is possible only if the sum of their waste production is lower than the limit. This condition is integrated by distance penalization (close clusters that should be merged but overcome the capacity are considered far apart).

The final step of the method corresponds with the calculation of representatives. Address points that are closest to the cluster center in terms of distance are selected as representatives.

3.5. Sub-problem definition

Another way to simplify the size of a task is to divide it into several more solvable subtasks. In this method, address points will be distributed into groups with relatively similar sizes (with low computational complexity). Each of these groups defines a separate problem that will be solved by ordinary methods to achieve optimality. However, such a simplification may disrupt the solution on the boundaries of each group (cluster). The inaccuracies will be discussed in the case study. The method is also the updated agglomerative hierarchical clustering with Ward's method, however, with different rules and criteria:

- Maximum cluster size is limited (e.g., 2,000 address points with regard to the model formulation and common housing development in the Czech Republic).
- If there are no more clusters to merge without breaking the size limit – stop.

By applying this method, several clusters are defined, where for each is the problem solved to optimality. This procedure is limited mainly by computational complexity according to the predefined size of the cluster.

3.6. Artificial problem testing

An artificial problem instance was generated in order to test various approaches to the optimization problem at hand. The instance is comprised of 2,000 address points in a radius of 500 m generated by a uniform pseudo-random number generator. For each address point, the number of inhabitants was generated randomly from the range of 2 to 17.

Both approaches, i.e., representative selection (Section 3.4) and sub-problem definition (Section 3.5) were tested on this artificial problem instance to determine which approach is more suitable and what is the influence of the number of representatives and maximum cluster size on the problem-solving. Moreover, the problem specification with multi-objective minimization of the average walking distance and number of container locations was tested alongside the original problem definition with maximum average walking distance constraint and minimization of the number of container locations (see Appendix A for model specifications). The maximum container capacity at each container location was set to 1,100 l; the maximum allowed walking distance for the multi-objective approach was set to 200 m and the maximum allowed average walking distance was set to 100 m.

For the representative selection approach, the number of representatives was set from 2% up to 16% with the step of 2% and a special case of 100% representatives – complete problem instance. For the sub-problem definition, the maximum sizes of clusters were set to 200, 250, 334, 500, and 1,000 address points. The results for the multi-objective problem definition are depicted in Table B.1 and Table B.2 in Appendix B (representative selection and sub-problem definition, respectively), and the results for the single-objective case are in Table C.1 and Table C.2 in Appendix C. The results of calculation time requirements are visualized in Fig. 2, using a logarithmic scale. For the small testing instance of 2,000 address points (corresponding to a small municipality), the calculation time rises when the number of clusters decreases and the share of selected representatives increase. The higher the number of clusters, the fewer address points are in each cluster, and thus each individual task has fewer variables resulting in a shorter calculation time. Especially the multi-objective task with a 100% share of representatives presents an extremely complex task (its solution was found within several days of computations).

The empirical results do not show monotonicity, which is given by the concrete computational problem; however, the increasing trend in computational trend is evident. It can also be seen in the result tables that the multi-objective approach might not be suitable for larger problem instances since the time complexity is high, and therefore the result would not be produced in a reasonable time. The need for simplification techniques is therefore necessary. It is also apparent that the solution in the representative case is highly dependent on the fraction of addresses that are considered as representants, which is not so pronounced in the single-objective case.

A similar procedure was applied to different formulations of the single-objective model. The individual problem formulations are listed in Appendix A. A total of four variants of single-objective problems were compiled from the equations. Variant 1 (V1) distributes a predetermined number of collection points with unlimited capacity while minimizing the average walking distance. Variant 2 (V2) limits the average walking

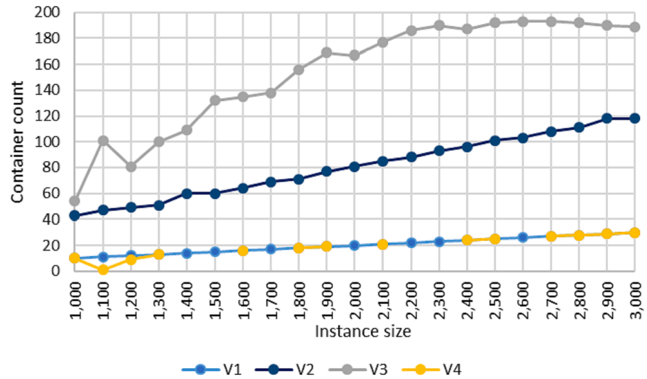

Fig. 4. Container counts for artificial instances of varying sizes – different single-objective variants.

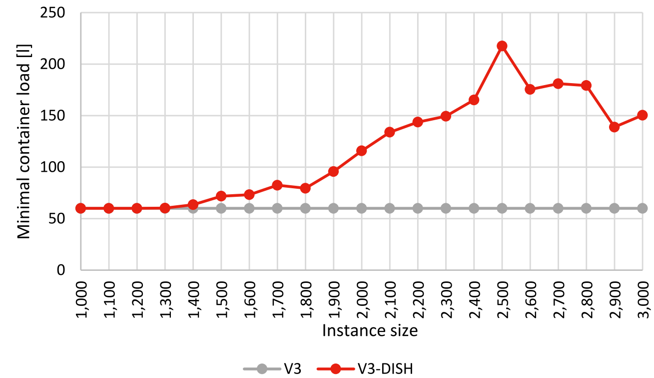

Fig. 7. Maximized minimal container load on artificial instances of different sizes (problem variant V3) – comparison between CPLEX solver (V3) and DISH (V3-DISH).

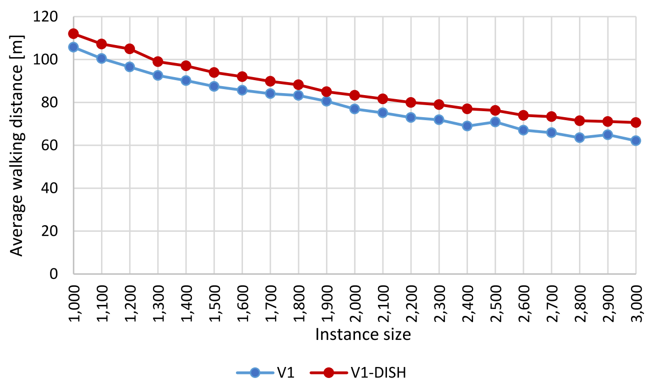

Fig. 5. Minimized average walking distance on artificial instances of different sizes (problem variant V1) – comparison between CPLEX solver (V1) and DISH (V1-DISH).

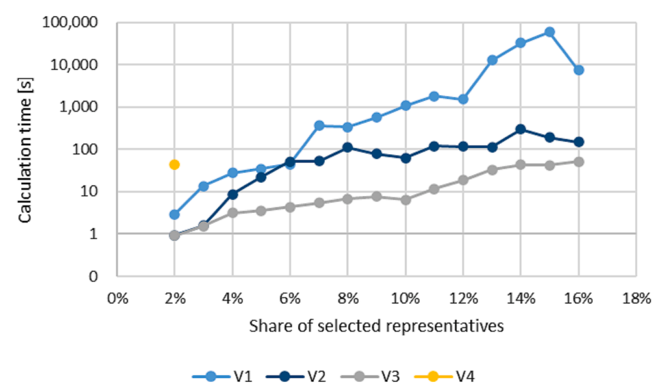

Fig. 8. Calculation times for artificial instances with varying ratios of representatives – different single-objective variants.

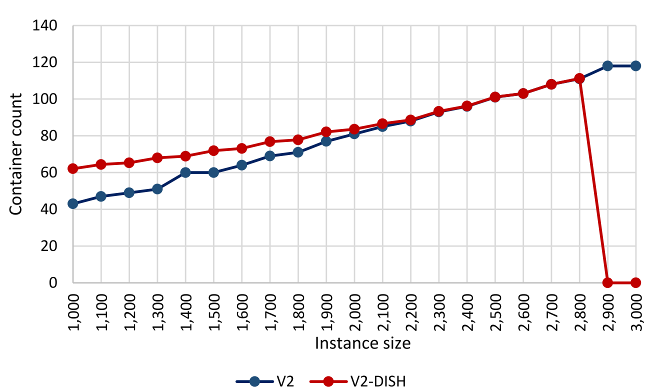

Fig. 6. Minimized container count on artificial instances of different sizes (problem variant V2) – comparison between CPLEX solver (V2) and DISH (V2-DISH).

distance to the selected threshold (target), the capacity of collection points is limited, and the number of collection points is minimized. In variant 3 (V3), the aim is to maximize the minimum utilization of all collection points while limiting the average walking distance and the capacity of collection points. Variant 4 (V4) limits the number of collection points with unlimited capacity and minimizes the maximum individual walking distance. These models were tested on 21 artificial instances with varying instance sizes (from 1,000 to 3,000) with a fixed number of selected representatives – 200 (see Fig. 3). The computational budget for the CPLEX optimizer was limited to one hour.

As can be seen in Fig. 3, the computational time for V1 and V2 is constant, while for V3, it increases with the instance size. For V4, the CPLEX optimizer was not able to provide a feasible solution within the specified time budget. A case of an outlier can be seen, for instance, with 1,100 address points, where the computational time and also the number of collection points (Fig. 4) are not as expected. In the authors' opinion, this anomaly can be accounted to the pseudo-random generation of the artificial instance.

In order to verify the CPLEX solver result, the optimization of these four problem variants (V1, V2, V3, and V4) was also performed by the metaheuristic algorithm DISH (Viktorin et al. 2019) specifically tuned for these problems. The DISH algorithm was providing locations of the containers, and for V2, V3, and V4, the CPLEX was solving the waste logistic and providing the objective function value as feedback for the metaheuristic. The stopping criterion for each optimization run was set to one hour of computational time, and due to the stochasticity of the method, ten runs for each problem variant and problem instance were performed to get a reasonable statistic. The average result from the metaheuristic over the ten runs is compared to the CPLEX solver solution in the following figures.

As can be seen from Fig. 5, the CPLEX solver is able to provide a solution with a better average walking distance for V1 than the metaheuristic approach for all problem instances. A similar result is displayed in Fig. 6, which shows the performance of both approaches on problem variant V2. The metaheuristic approach is able to provide a competitive solution for larger instances (2,100 to 2,800 address points) but was unable to provide a feasible solution within the given computational time for the two largest instances (2,900 and 3,000 address points). A different result can be seen for the V3, where the DISH metaheuristic is able to provide a solution for larger instances of the problem with higher minimal utilization of containers while the CPLEX solver solution has at

Table 3
The city of Zlín – address count and population (source: Czech Statistical Office, the year 2018).

Characteristics	Units	Number
Address points' count	[-]	6,571
Address points' count after noise filtering	[-]	6,212
Population	[cap]	51,721
Population after noise filtering	[cap]	50,903
Waste container capacity	[l]	240
Waste production	[l/cap/collection cycle]	0.25
Maximum avg. walking distance	[m]	500

least one container on the minimal allowed utilization (60 l). The DISH solution also usually uses fewer containers while preserving the condition of a maximal average walking distance of 50 m. Thus, for this type of problem, the metaheuristic is able to provide a better solution but for the price of longer computation. On the other hand, the metaheuristic was not able to provide a feasible solution within one hour of computational time for any of the problem instances of V4. The result tables are available in Appendix D. Fig. 7

Further calculation time testing was performed for the representative selection method. The results are shown in Fig. 8; they point out the significant differences in computational complexity between the tested variants. The values on the y-axis are in a logarithmic scale to better represent the differences. V3 achieves the best results in terms of calculation speed, and even for the growing proportion of considered representatives, the calculation time is in the order of units up to tens of seconds. V2 calculation time is an order of magnitude higher value. The calculation time of V1 reaches the level of thousands of seconds. V4 is the most demanding in calculation time. It was not calculated for the chosen simplifications even after a few days, except for the scenario with 2% representatives. Due to the small size of the test task, the computational time is relatively enormous even when using simplification techniques, which underlines the importance of their application. The results for the sub-problem definition method have a similar character.

For the solution of real instances, it is recommended to use a combination of both approaches or individual approaches based on the targets of the calculations. If a fast solution with as few containers as possible is desired, then it is advisable to use a representative selection with a few percent share. On the other hand, if a slightly larger number of collection points is not important, the sub-problem definition takes place. The case study will demonstrate the approaches on a single-objective variant of the model.

4. Case studies

The first case study selected in this work considers the two approaches mentioned above, representative selection (Section 3.4) and sub-problem definition (Section 3.5), for the city of Zlín (Czech Republic) and the allocation of cooking oils and fat waste containers since the presented approach has especially contributed to the waste fractions with a less density (of the collection network) and production (Matušinec et al., 2022). This city was selected because it features a densely populated city center, large housing estates typical for Czech cities, industrial zone and old household streets, and remote populated satellite towns. However, it is still small enough to demonstrate the benefits and drawbacks of both approaches visibly. The main characteristics (number of address points and population) of the city of Zlín and the waste collection details are summarized in Table 3. All collection points were assumed to have only a single container with a predefined capacity. Waste production amount is adopted for cooking oil and fat waste. The waste production is given per collection cycle – regular collection every

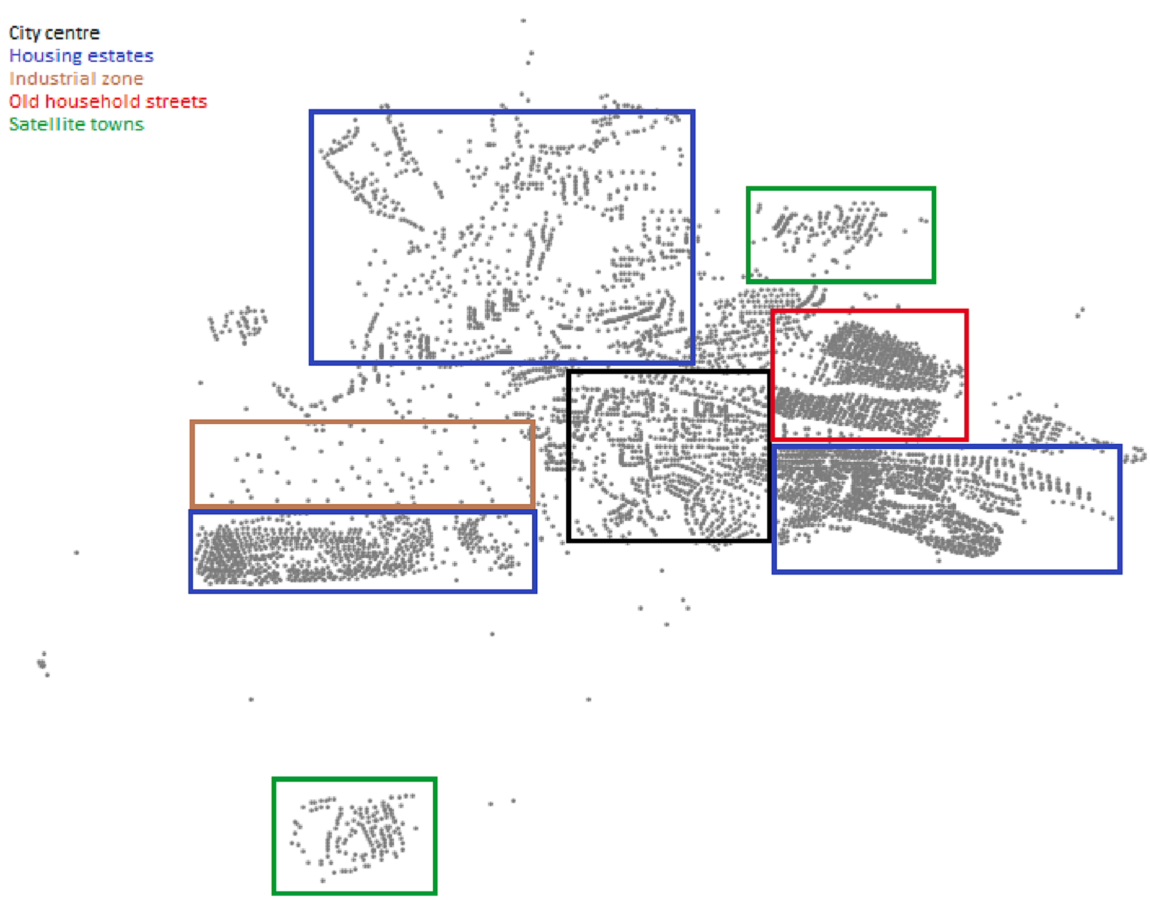

Fig. 9. The city of Zlín – approximate structure map.

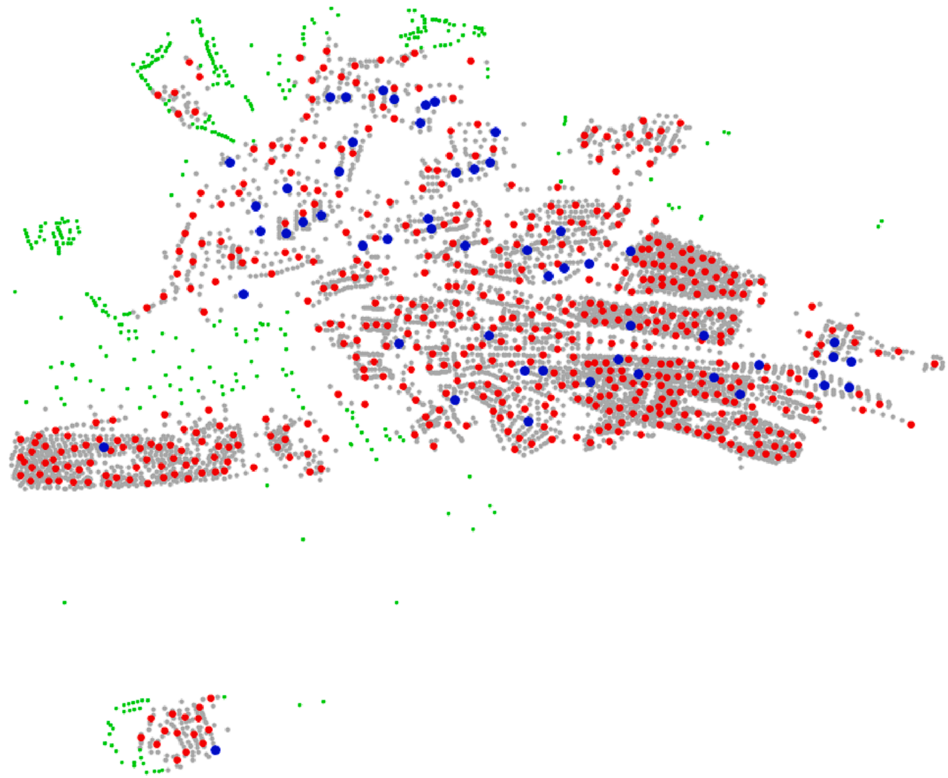

Fig. 10. The city of Zlín – representative selection (gray points = address points, red points = representatives, green points = noise, blue points = waste container locations).

four weeks – while the waste production quantity was determined according to the steps described in Matušinec et al. (2022). The overall structure of the city addresses is visible in Fig. 9.

As can be derived from Table 3, the number of people capable of filling the waste container in one collection cycle is 240/0.25 = 960. Thus, the whole city's minimal number of waste containers would be 51,721/960 = 53.88 ≈ 54. However, this number does not attain the average walking distance condition and does not reflect the undesirable collection areas – sparsely populated areas (see next Section 4.1 with noise filtering). The resulting number of containers covering a given area will thus start at the smallest possible number, and given these limitations, their number will increase, but it will happen that the resulting value will remain at this number.

4.1. Zlín – Noise filtering

The first step in evaluating this case study is filtering the noise points, which are not desirable for waste bin allocation. The result after applying the noise filtering method described in Section 3.1 is depicted in Fig. 10. It can be seen that the industrial area was successfully filtered out since there is a low number of address points with a low population. The same goes for less populated and remote areas. The numbers of addresses and inhabitants filtered out are in Table 3. The noise ratio was 5.8% for the address points' count and 1.6% for the population.

The noise filtering leads to a minimal number of waste containers, 50,903 / 960 = 53.02 ≈ 54, which is still the same minimal number as it was for the unfiltered dataset. Although the minimal number of waste containers still does not attain the average walking distance condition, the representative selection approach is able to provide a solution with exactly 54 waste containers, as is described in the next section.

4.2. Zlín – Representative selection

In this approach, approximately 10% (600) of the data address points are considered to allocate waste containers by the method described in Section 3.3. This approach generates evenly distributed waste collection representatives, which are depicted in Fig. 10. The map also depicts the final step of the representative selection approach, which is the selection of waste container locations by the optimization model solved in GAMS software by CPLEX optimizer (Gams Development Corportaion, 2013).

This approach yields 54 collection points, with an average walking distance of exactly 500 m and an average container utilization of 98.19%. The computation time needed to evaluate this approach was approximately 6.5 h for clustering and model preparation and 4.5 h for GAMS solving. The approach works very well in determining the minimum number of collection points, but it does not place the containers appropriately within the space. The approach to fulfill this purpose is a multi-objective formulation.

4.3. Zlín – Sub-problem definition

It was experimentally determined on the same computer (AMD Ryzen 9 3900X 12cores 3.8 GHz, 32 GB RAM, Nvidia GeForce GT 710) that the GAMS CPLEX solver is able to provide a solution for a fully connected network of approximately two thousand clustered address points in a reasonable time (2.5 h). It is assumed that larger clusters for this case study significantly increase the computational complexity since the average walking distance increases over 500 m for multiple waste container allocation points, and the solution space is much more constrained. Thus, the data address points were according to the method described in Section 3.4. hierarchically clustered into four clusters (1,084, 1,739, 1,455, and 1,934 address points in each), and those four clusters were solved individually to optimality by GAMS software. The combined solution is provided in Fig. 11.

The result of this approach is 55 selected locations for waste containers with an average walking distance of exactly 500 m and an average container utilization of 96.41%. The computation time needed

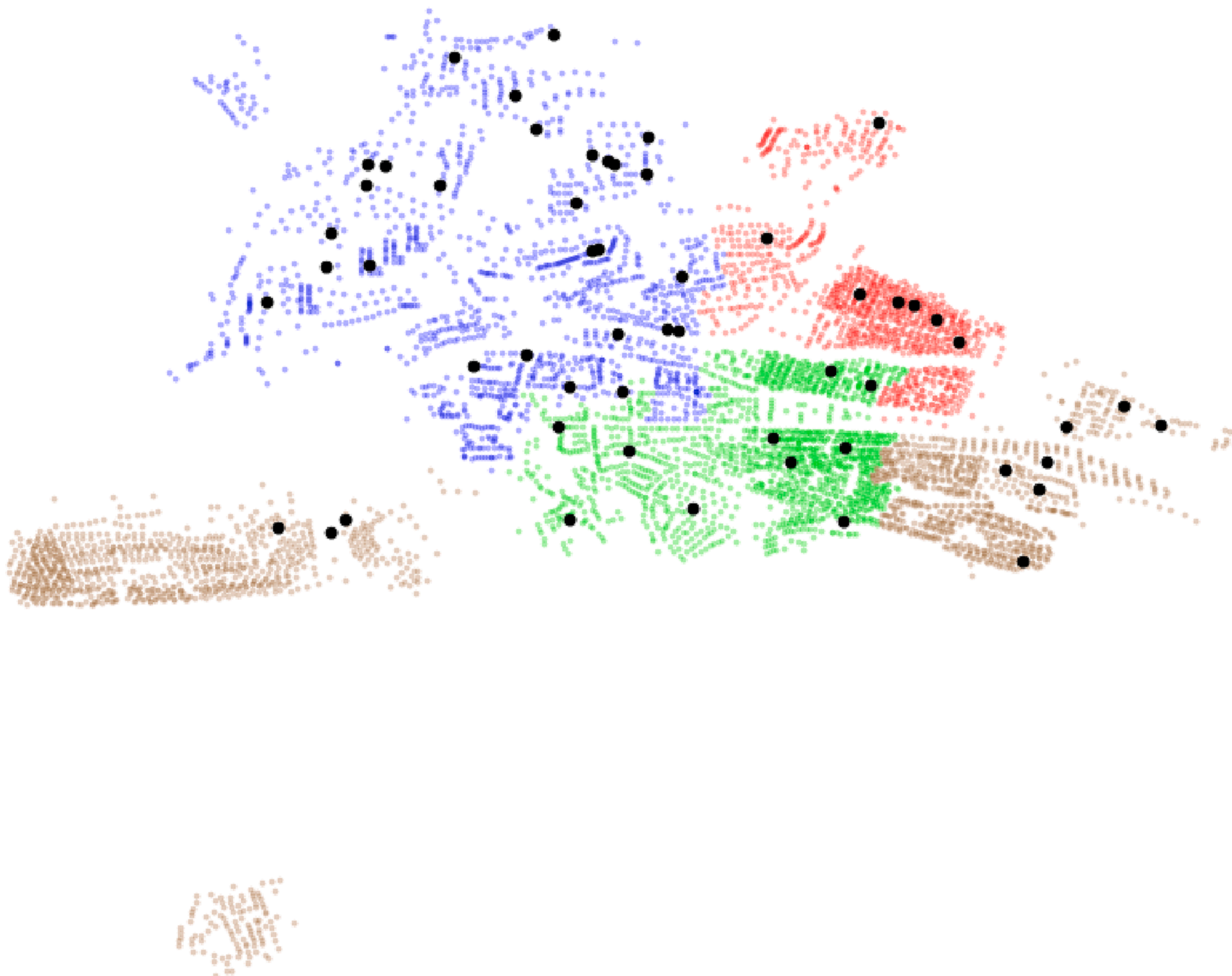

Fig. 11. The city of Zlín – waste container locations (black points) selected by sub-problem definition approach.

Table 4
The city of Prague – address count and population (source: Czech Statistical Office, the year 2018).

Characteristics	Units	Number
Address points' count	[-]	107,398
Address points' count after noise filtering	[-]	102,927
Population	[cap]	1,246,774
Population after noise filtering	[cap]	1,235,352
Waste container capacity	[l]	240
Waste production	[l/cap/collection cycle]	0.25
Maximum avg. walking distance	[m]	500

for the evaluation of this approach was approximately 6.5 h for clustering and model preparation and 0.6 h for GAMS solving. From the point of view of the distribution of collection points in space, this approach looks much worse than the representative selection method.

4.4. Zlín – Approach comparison

While the resulting number of waste containers is 54 for representative selection and 55 for the sub-problem definition approach, respectively, and the average walking distance is similarly below the targeted limit (the sub-problem definition performs a little bit better because of a higher number of allocated collection points). The main difference lies in the global location of waste containers. It can be seen in Fig. 8 that the satellite town in the south-west part of the map gets allocated to one of the collection points in a north-bound housing estate, which would be inconvenient for residents from that satellite town since they would have to travel to a distant location and probably would rather choose to not recycle. Moreover, it can be seen that the sub-problem definition approach sometimes tends to select waste container locations close to one another. This is due to the used model formulation (single-objective in Appendix A), optimizing the number of waste collection points w. r. t. the maximal average walking distance condition. The representative selection approach provides relatively evenly distributed representants after the clustering step, and thus it is able to provide a more evenly distributed collection infrastructure. This leads to easily implementable solutions for real-world scenarios when using the representative selection approach. The optimization of the average walking distance at the same time in the complex multi-criteria model would require distribution into more clusters or a reduced share of selected representatives (depending on the number of address points in the instance).

4.5. Prague – Case study

For the second case study, the capital and biggest city of the Czech Republic was selected. It is several times bigger than Zlín, which was selected in the previous case study. Its population is ca. 1.2 million and involves more than 100 thousand address points (see Table 4). Prague involves all the various area characteristics discussed above; that is a rational reason for testing the developed approach. The distribution of address points is depicted in Fig. 9, and as can be seen, a small number of address points are on the outskirts of the city. These points are filtered

Table 5
The city of Prague – clusters' characteristics and statistics with time complexity.

Cluster	1	2	3	4	5	6	7
Address count	7,500	6,991	7,500	7,500	7,500	7,500	7,500
Population	44,363	24,530	73,106	98,453	94,440	179,391	84,388
Minimum waste container count (rounded up)	47	26	77	103	99	187	88
Waste containers (difference from minimum)	47 (0)	27 (+1)	77 (0)	112 (+9)	99 (0)	193 (+6)	92 (+4)
Avg. container utilization [%]	98.32	94.64	98.90	91.57	99.37	96.82	95.55
Model preparation time [h]	5.13	9.72	6.14	5.58	5.12	5.25	5.13
GAMS solution [h]	13.52	31.09	14.20	6.90	6.99	5.12	4.39
Cluster	8	9	10	11	12	13	14
Address count	7,500	5,936	7,500	7,500	7,500	7,500	7,500
Population	48,497	76,123	71,193	99,492	113,884	108,188	119,304
Minimum waste container count (rounded up)	51	80	75	104	119	113	125
Waste containers (difference from minimum)	51 (0)	80 (0)	80 (+5)	104 (0)	121 (+2)	113 (0)	132 (+7)
Avg. container utilization [%]	99.05	99.12	92.70	99.65	98.04	99.73	94.15
Model preparation time [h]	2.55	1.14	2.50	2.42	2.43	2.42	1.92
GAMS solution [h]	14.74	1.14	13.85	5.44	11.75	5.19	11.79

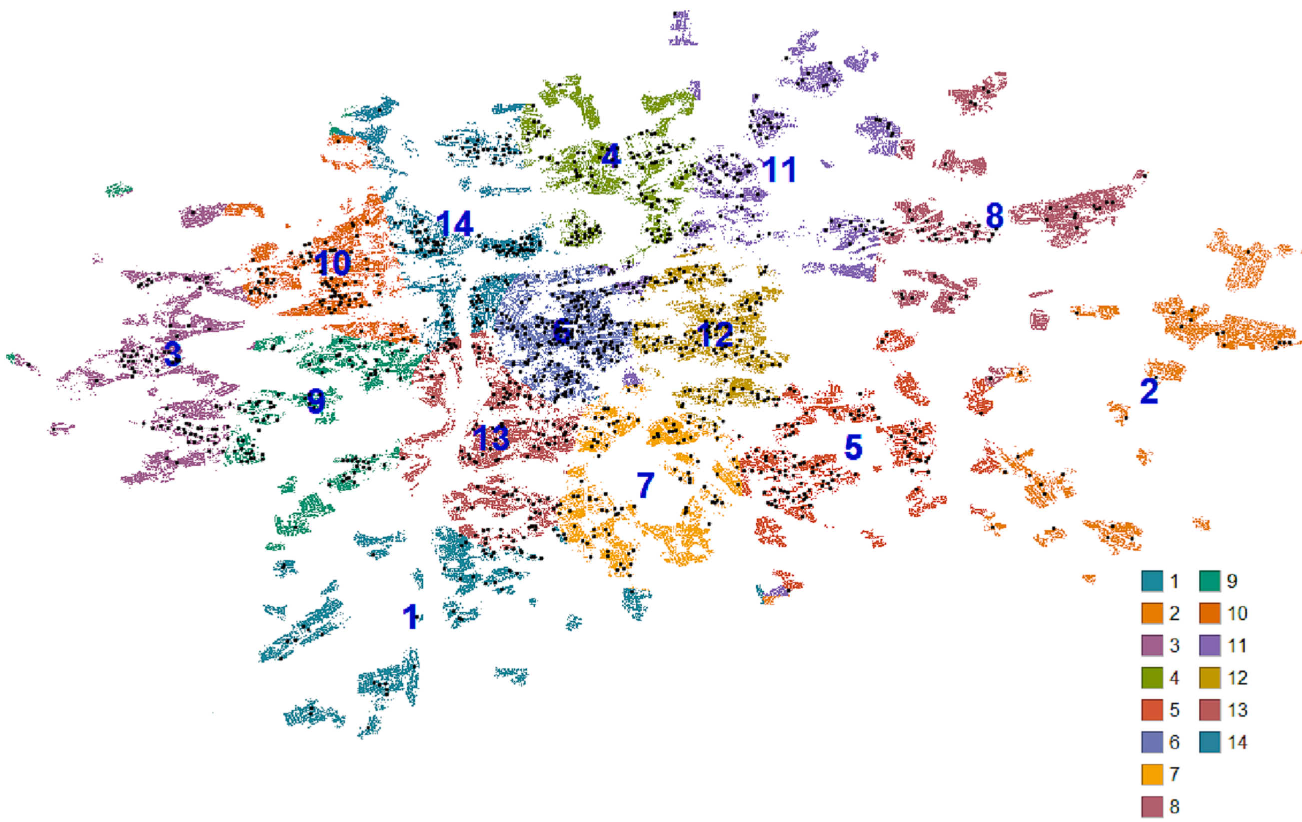

Fig. 12. The city of Prague – clustering (without noise) and waste container locations by representative approach.

out in the noise filtering step, but for the purpose of better visual representation and space-saving, upcoming figures depict the only area with no filtered points. Thanks to the benefits of the representative selection approach described in Section 4.4, it was selected for the waste collection bins allocation in Prague as well.

The first step in this case study is to remove noise address points, which would be undesirable for waste collection, and it is the same process as described in Section 3.1. The results of this step are summarized in Table 5. This leads to a minimum number of 1,287 waste bins (1,235,352 / 960 = 1,286.83 ≈ 1,287) without the maximum walking distance constraint. The noise ratio was 4.3% for the address points' count and 0.9% for the population.

Since the resulting data address points' count is still over 100,000, the representative selection approach with the suggested 10% of address points considered as representants would lead to a complex task unsolvable in a reasonable time. Thus, an additional clustering step (see Section 3.2) is used for the division of the data set into smaller blocks that would be solvable but would still provide a complex solution. It was shown in the first case study that a city of approximately 6,200 data address points is solvable in a fairly reasonable time, and therefore, the proposed clustering characteristics were selected as follows. The maximum number of address points in a cluster was set to 7,500, and the desired number of clusters k was set to 14 (102,927 / 7,500 = 13.72 ≈ 14). Afterward, the representative selection was applied in each cluster. The number of representatives was kept the same as in the first case study – 600 for each cluster. The final step is a selection of waste container locations in each cluster by the optimization model solved in GAMS software by CPLEX optimizer (Gams Development Corportaion, 2013). The result of clustering and waste container locations are presented in Fig. 12. The characteristics of clusters and results are

Fig. 13. The city of Prague – cluster #6 (most populated) – waste container locations (Black points) selected by representative selection approach.

summarized in Table 5, together with the time needed for model preparation and solution. An example of a detailed solution of the most populated cluster #6 is presented in Fig. 13.

The representative selection approach solving the waste collection bins locations in Prague yields a solution with 1,328 bins and average utilization of 96.90% w.r.t. the maximum walking distance condition of 500 m. This is only 34 bins more than the minimum needed number in clustered variant (1,294) and 41 bins more than in the not clustered variant (1,287). The results confirm that when the CPLEX solver deploys the minimum possible number of containers for a cluster and such a solution exists, it is much faster compared to the instance, where some extra must be added (e.g., see clusters #2, #10, #12, #14). At the same time, it can be seen that this is not always the case, which shows that it is very difficult or impossible to find a general description of the task's nature and suggest general recommendations.

5. Discussion and conclusions

This paper has aimed to suggest a general clustering-based computational approach to solve large waste collection network design problems. A social-managerial problem can arise when a collection point must be ensured in compliance with the environment, i.e., finding a specific location for a bay with containers may not be possible in some environments. The authors follow up on their previous research, especially Matušinec et al. (2022), where the authors provide a mathematical model to optimize waste bin location from a city perspective. The optimal model reflects a walking distance constraint that was developed by Nevrlý et al. (2021). However, large-scale problems could not be solved with traditional optimization models and solvers, so this paper deals with hierarchical and other clustering-based algorithms to provide suitable simplification approaches to solve such problems. Since the paper provides a new model and solution approach, no alternative methodology leading to some comparable results could be used and tested.

Based on testing computations for various formulations of waste bin location problems, a new approach to large-scale problems simplification, allowing solving the problem in a reasonable time, was developed. However, regarding each particular problem, it is necessary to perform some calibration computations exposing the computational limits of the particular problem and especially a solvable problem size. Based on such computational analysis, the input parameters for the developed approach are determined, and so the large-scale problems are simplified to solvable clusters. A new view of such a problem can be to seek maximal clusters that are computationally solvable and have a high number of clusters while respecting a suitable combination of mentioned conditions and criteria. One central node (medoid) can be used as a representative for a potential waste container(s) location. The solution to such a problem is the main novelty of the presented research, motivated by the WM's real problems. However, the approach is defined in a general way and so is transferable to other application areas. The optimal location of public transport stations and the optimal infrastructure layout of local doctors, pharmacies, shops, etc., can show an example.

When comparing the results of the two presented methods (representative selection and sub-problem definition), the resulting number of allocated containers is very similar. The difference can be seen in the containers' location, where the sub-problem definition sometimes tends to select containers close to each other. It is also caused by the selected model (see Appendix A), where the walking distance is included only in the form of constraint. The representative selection method should be used when the minimum number of collection points is strictly required.

The solution to multi-objective formulation also performed better for the representative selection method (artificial instance). Unfortunately, multi-criteria models are even worse at getting solutions fast, so the maximal size of clusters and the percentage of representatives have to be adapted accordingly. The representative selection method is better for the single-objective formulation because it provides evenly distributed representatives within the area. In each municipality, the number of separately collected waste is constantly increasing, while each waste has different characteristics and requirements for the density of the collection network. The density also has an economic impact in the form of collection costs. The amendment of the legislation forces mandatory textile collection from 2025. Mayors and people responsible for changes to the collection system (densification of the bin collection network) will welcome the proposed approaches and different formulations of the models, as they will be able to compare their results and subsequently make managerial decisions.

The approach adopted for the city of Prague resulted in 14 clusters, which were afterward solved by the presented representative selection approach. The combination of these methods seems the best option for solving large-scale instances, as in the proposed case study. However, some clusters are distributed in space with isolated, separate units. These might be identified prior to the optimization and possibly solved separately. This approach could lead to too many separate clusters, which is not desirable. Further research within this area is needed. The authors are inclined to the possibility of not solving the process of creating clusters with a deterministic number of clusters but dividing the problem according to the density of areas and further dividing the clusters that do not meet the size limits. Choosing a combination of the proposed procedures with another logical criterion may be ideal. The proposed approaches cannot generally find optimal values; still, at a small cost of deviation from the optimum, they enable computability in a reasonable time and subsequent implementation of the outputs to actual bin deployment.

The suggested approach can be applied in various other case studies. Besides the applications and possible algorithms improvements, there are several potential directions for future research. One of them is reflecting the difference between a real infrastructure of sidewalks and a linear distance (by air) which can typically occur when considering a block of flats with front and back doors. This can lead to several research questions: Does it have significant benefits? How should the task be partitioned? Does it influence the clustering algorithm and the solution, respectively? The other potential future research directions can involve the application of other waste-generation fractions and points (Letelier et al., 2022), stochastic characteristics (Adeleke and Ali, 2021), or objective functions with a sustainable approach (Rathore et al., 2020).

CRediT authorship contribution statement

Adam Viktorin: Methodology, Software, Data curation, Writing – original draft. **Dušan Hrabec:** Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Vlastimír Nevrlý:** Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Radovan Šomplák:** Conceptualization, Supervision. **Roman Šenkeřík:** Supervision.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by Grant No. GA 20-00091Y „Development of Sustainable Waste Management: Methods and Operations Research Perspectives" of the Czech Science Foundation. The authors also gratefully acknowledge the financial support provided by the Technology Agency of the Czech Republic, grant No. SS02030008 "Center of Environmental Research: Waste management, circular economy and environmental security". This work was also supported by the Tomas Bata University under project no. FSR FORD 5-6/2022-23/FLKŘ/001 Sustainability in Transport: Modern Trends and their Impact on the Environment.

Appendix A

Appendix A involves a table describing the mathematical notation used in the mathematical model, which was presented in Matušinec et al. (2022) and Nevrlý et al. (2021) – parts that were used in the paper are below. Further, some variants of model formulation were used for testing purposes of developed approaches (see Table A1).

Single-objective formulation.

$$\min z_1 \; or \; \min z_2 \tag{1}$$

$$z_1 = \sum_{a \in A} \delta_a \tag{2}$$

$$z_2 = \frac{\sum_{a \in A} p_a \sum_{j \in J} x_j \, d_j \, M_{out(j,\ a)}}{\sum_{a \in A} p_a} \tag{3}$$

$$z_2 \leq w_{tar} \tag{4}$$

$$q_a \leq \delta_a \, c_a \forall a \in A \tag{5}$$

$$y_a + \sum_{j \in J} x_j \, M_{out(j,\ a)} = 1$$

$$\forall a \in A \tag{6}$$

$$q_a = p_a y_a + \sum_{j \in J} \sum_{b \in A} p_b \, x_j \, M_{out(j,\ b)} \left(-M_{in(j,\ a)}\right)$$

Table A1
Models notation.

Sets	
A	Set of address points, $a, b \in A$
J	Set of edges, $j \in J$
Parameters	
p_a	Estimated production of fat waste by address point a [l]
d_j	Edge distance [m]
$M_{out(j,a)}$	Matrix of outflow edges j from a [-]
$M_{in(j,a)}$	Matrix of inflow edges j to a [-]
c_a	Container capacity [l]
c_{min}	Minimum container load [l]
w_{tar}	Target average walking distance [m]
w_{max}	Maximal walking distance [m]
w_{total}	Total number of collection points [-]
u_a	The utilization of the collection point in node a [-]
Variables	
z_1	Value of objective function – number of collection points [-]
z_2	Average walking distance [m]
z_4	Maximal utilization of collection points [%]
x_j	The proportion of waste production that flows along the edge j [-]
y_a	The proportion of waste production from address point a that is assigned to the collection point a [-]
w_a	Walking distance of address point a[m]
q_a	Load at collection point [l]
Binary variables	
δ_a	Existence of a collection point at the address point a [-]

$$\forall a \in A \tag{7}$$

$$x_j \geq 0 \forall j \in J, \tag{8}$$

$$y_a, q_a \geq 0 \forall a \in A \tag{9}$$

$$\delta_a \in \{0, 1\} \forall a \in A. \tag{10}$$

The second model includes these additional constraints, and the objective function is replaced by multi-objective, which uses weights resulting from single objective solutions z_1^* and z_2^*. The model was used for testing artificial instances.

Multi-objective formulation.

$$\min \frac{z_1}{z_1^*} + \frac{z_2}{z_2^*} \tag{11}$$

$$w_a = \sum_{j \in J} x_j d_j M_{out(j,a)} \forall a \in A \tag{12}$$

$$w_a \leq w_{max} \forall a \in A \tag{13}$$

$$w_a \geq 0 \forall a \in A \tag{14}$$

$$c_{min} \delta_a \leq q_a \forall a \in A. \tag{15}$$

Single-objective variant 1 (V1).

$$\min z_2$$

s.t. (2),(3),(5),(6),(7),(8),(9),(10),(15) and.

$$z_1 = w_{total}$$

Single-objective variant 2 (V2).

$$\min z_1$$

$$s.t. \tag{15}$$

Single-objective variant 3 (V3).

$$\min z_4$$

s.t. (2),(3),(4),(5),(6),(7),(8),(9),(10),(15) and.

$$z_4 \leq \frac{q_a}{c_a}$$

Single-objective variant 4 (V4).

$$\min z_4$$

s.t. (2),(5),(6),(7),(8),(9),(10),(12),(13),(15) and.

$$z_1 \leq w_{total}$$

Appendix B

Multi-objective problem definition (Tables B1 and B2).

Table B1
Representative selection testing on artificial problem instance.

Test case	GAMS time [s]	Number of container locations [-]	Utilization [%]	Avg. walking distance [m]
2% – 40 reps	27.24	17	99.91	86.64
3% – 60 reps	42.67	17	99.91	88.30
4% – 80 reps	70.77	21	81.74	75.02
5% – 100 reps	394.50	20	84.93	74.70
6% – 120 reps	563.78	22	77.21	70.83
7% – 140 reps	536.99	25	67.95	65.83
8% – 160 reps	759.24	23	73.85	68.67
9% – 180 reps	810.36	26	65.33	64.37
10% – 200 reps	969.08	26	65.33	65.89
11% – 220 reps	351.19	28	60.67	64.78
12% – 240 reps	892.39	30	56.62	59.98
13% – 260 reps	1,008.06	29	58.57	60.71
14% – 280 reps	619.69	27	62.91	65.07
15% – 300 reps	831.94	27	62.91	66.63
16% – 320 reps	817.89	32	53.08	57.76
100% – 2,000 reps	Unable to compute	–	–	–

Table B2
Sub-problem definition testing on artificial problem instance.

Test case	Number of clusters [-]	Init time – cluster creation [hh:mm:ss]	GAMS time [s]	Number of container locations [-]	Utilization [%]	Avg. walking distance [m]
200 addresses	13	0:00:20	2,194.91	49	34.67	47.09
250 addresses	11	0:00:29	1,685.37	48	35.39	47.15
334 addresses	8	0:01:28	91,652.90	41	41.43	50.95
500 addresses	6	0:02:22	150,502.00	37	45.91	53.55
1,000 addresses	3	0:07:41	113,828.20	36	47.18	54.15

Appendix C

Single-objective problem definition (Tables C1 and C2).

Table C1
Representative selection testing on artificial problem instance.

Test case	GAMS time [s]	Number of container locations [-]	Utilization [%]	Avg. walking distance [m]
2% – 40 reps	33.89	17	99.92	85.32
3% – 60 reps	44.56	17	99.92	94.53
4% – 80 reps	45.83	17	99.92	87.20
5% – 100 reps	62.22	18	94.37	90.33
6% – 120 reps	83.53	17	99.92	92.11
7% – 140 reps	126.16	18	94.37	90.13
8% – 160 reps	140.05	18	94.37	89.24
9% – 180 reps	144.94	17	99.92	91.44
10% – 200 reps	239.56	18	94.37	91.46
11% – 220 reps	346.72	17	99.92	98.48
12% – 240 reps	301.69	18	94.37	93.64
13% – 260 reps	341.86	18	94.37	86.93
14% – 280 reps	518.31	17	99.92	97.09
15% – 300 reps	433.84	18	94.37	95.58
16% – 320 reps	490.20	17	99.92	96.25
100% – 2,000 reps	14,123.02	17	99.92	98.90

Table C2
Sub-problem definition testing on artificial problem instance.

Test case	Number of clusters [-]	Init time – cluster creation [hh:mm:ss]	GAMS time [s]	Number of container locations [-]	Utilization [%]	Avg. walking distance [m]
200 addresses	13	0:00:20	7.41	25	67.95	87.55
250 addresses	11	0:00:29	12.75	24	70.78	88.32
334 addresses	8	0:01:28	77.01	21	80.89	91.71
500 addresses	6	0:02:22	313.36	19	89.40	94.08
1,000 addresses	3	0:07:41	858.66	18	94.37	95.81

Appendix D

Artificial problem results – metaheuristic (see Tables D1–D7).

Table D1
CPLEX solver solution for V1.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count
1000	3600.27	105.71	949.80	10
1100	3600.27	100.43	937.36	11
1200	3600.31	96.53	921.83	12
1300	3600.34	92.54	927.69	13
1400	1673.42	90.17	947.07	14
1500	2426.89	87.46	959.67	15
1600	785.49	85.67	955.00	16
1700	741.13	84.09	972.88	17
1800	918.20	83.25	943.61	18
1900	840.88	80.56	963.89	19
2000	1195.08	76.91	964.95	20
2100	908.50	75.13	964.14	21
2200	1217.23	72.93	951.50	22
2300	1035.63	71.89	962.17	23
2400	1209.83	68.97	953.21	24
2500	1150.33	70.87	968.28	25
2600	688.80	67.05	945.58	26
2700	752.63	65.89	956.74	27
2800	950.17	63.54	948.50	28
2900	1224.02	64.93	971.59	29
3000	1052.59	62.11	940.87	30

Table D2
CPLEX solver solution for V2.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count
1000	96.91	50.00	220.88	43
1100	63.02	50.00	219.38	47
1200	83.75	50.00	225.76	49
1300	74.34	50.00	236.47	51
1400	94.11	50.00	220.98	60
1500	62.19	50.00	239.92	60
1600	66.00	50.00	238.75	64
1700	68.23	50.00	239.70	69
1800	87.20	50.00	239.23	71
1900	125.89	50.00	237.84	77
2000	121.89	50.00	238.26	81
2100	89.97	50.00	238.20	85
2200	101.53	50.00	237.88	88
2300	151.06	50.00	237.96	93
2400	73.30	50.00	238.30	96
2500	84.70	50.00	239.67	101
2600	85.91	50.00	238.69	103
2700	89.20	50.00	239.19	108
2800	99.95	50.00	239.26	111
2900	92.72	50.00	238.78	118
3000	106.24	50.00	239.20	118

Table D3
CPLEX solver solution for V3.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count	Minimal container load [l]
1000	2.25	50.00	175.89	54	60
1100	12.27	50.00	102.09	101	60
1200	3.02	50.00	136.57	81	60
1300	3.59	50.00	120.60	100	60
1400	4.14	50.00	121.64	109	60
1500	4.72	50.00	109.05	132	60
1600	5.45	50.00	113.19	135	60
1700	6.23	50.00	119.85	138	60
1800	7.03	50.00	108.88	156	60
1900	6.84	50.00	108.37	169	60
2000	8.84	50.00	115.56	167	60
2100	8.41	50.00	114.39	177	60
2200	10.78	50.00	112.54	186	60
2300	12.06	50.00	116.47	190	60
2400	13.13	50.00	122.34	187	60
2500	14.19	50.00	126.08	192	60
2600	14.64	50.00	127.38	193	60
2700	15.55	50.00	133.84	193	60
2800	16.55	50.00	138.32	192	60
2900	17.11	50.00	148.29	190	60
3000	17.41	50.00	149.34	189	60

Table D4
CPLEX solver solution for V4.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count
1000	3600.38	269.10	949.80	10
1100	3600.31	499.72	10311.00	1
1200	3600.41	342.32	1229.11	9
1300	3600.42	266.55	927.69	13
1400	3600.19	–	–	–
1500	3600.20	–	–	–
1600	3600.59	194.64	955.00	16
1700	3604.19	–	–	–
1800	3600.67	224.82	943.61	18
1900	3600.81	191.41	963.89	19
2000	3600.31	–	–	–
2100	3600.69	264.33	964.14	21
2200	3600.33	–	–	–
2300	3600.34	–	–	–
2400	3601.02	198.11	953.21	24
2500	3601.06	175.96	968.28	25
2600	3600.41	–	–	–
2700	3601.30	165.68	956.74	27
2800	3601.14	222.18	948.50	28
2900	3601.44	174.43	971.59	29
3000	3601.13	232.56	940.87	30

Table D5
DISH metaheuristic average solution for V1.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count
1000	3600.00	111.99	949.80	10
1100	3600.00	107.21	937.36	11
1200	3600.00	104.94	921.83	12
1300	3600.00	98.99	927.69	13
1400	3600.00	97.04	947.07	14
1500	3600.00	93.94	959.67	15
1600	3600.00	91.99	955.00	16
1700	3600.00	89.83	972.88	17
1800	3600.00	88.21	943.61	18
1900	3600.00	84.97	963.90	19
2000	3600.00	83.33	964.95	20
2100	3600.00	81.65	964.14	21
2200	3600.00	79.98	951.50	22
2300	3600.00	78.95	962.17	23
2400	3600.00	76.99	953.21	24
2500	3600.00	76.25	968.28	25
2600	3600.00	73.96	945.58	26
2700	3600.00	73.37	956.74	27
2800	3600.00	71.45	948.50	28
2900	3600.00	71.08	971.59	29
3000	3600.00	70.63	940.87	30

Table D6
DISH metaheuristic average solution for V2.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count
1000	3600.00	50.00	153.15	62.10
1100	3600.00	50.00	160.34	64.40
1200	3600.00	50.00	169.63	65.30
1300	3600.00	50.00	177.51	68.00
1400	3600.00	50.00	192.54	68.90
1500	3600.00	50.00	200.45	71.90
1600	3600.00	50.00	209.07	73.10
1700	3600.00	50.00	215.51	76.80
1800	3600.00	50.00	218.42	77.80
1900	3600.00	50.00	223.14	82.10
2000	3600.00	50.00	230.92	83.60
2100	3600.00	50.00	233.84	86.60
2200	3600.00	50.00	236.28	88.60
2300	3600.00	50.00	237.20	93.30
2400	3600.00	50.00	237.81	96.20
2500	3600.00	50.00	239.44	101.10
2600	3600.00	50.00	238.69	103.00
2700	3600.00	50.00	239.19	108.00
2800	3600.00	50.00	238.83	111.20
2900	3600.00	50.00	237.98	–
3000	3600.00	50.00	237.22	–

Table D7
DISH metaheuristic average solution for V3.

Size	Time [s]	Avg. walk distance [m]	Avg. container load [l]	Container count	Minimal container load [l]
1000	3600.00	50.00	95.69	99.70	60.00
1100	3600.00	50.00	101.89	101.60	60.00
1200	3600.00	50.00	111.34	99.80	60.00
1300	3600.00	50.00	121.89	99.70	60.20
1400	3600.00	50.00	144.23	93.00	63.50
1500	3600.00	50.00	163.58	88.50	71.80
1600	3600.00	50.00	173.70	88.70	73.20
1700	3600.00	50.00	189.57	87.70	82.40
1800	3600.00	50.00	185.90	91.80	79.50
1900	3600.00	50.00	208.52	87.90	95.70
2000	3600.00	50.00	224.52	86.00	115.82
2100	3600.00	50.00	229.36	88.40	133.90
2200	3600.00	50.00	233.71	89.60	143.70
2300	3600.00	50.00	235.00	94.20	149.40
2400	3600.00	50.00	238.06	96.10	165.21
2500	3600.00	50.00	239.67	101.00	217.55
2600	3600.00	50.00	238.69	103.00	175.39
2700	3600.00	50.00	238.75	108.20	181.00
2800	3600.00	50.00	238.62	111.30	179.20
2900	3600.00	50.00	237.38	118.70	138.83
3000	3600.00	50.00	237.81	118.70	150.40

Appendix F. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cie.2023.109142.

References

Abdallah, M., Hamdan, S., & Shabib, A. (2021). A multi-objective optimization model for strategic waste management master plans. *Journal of Cleaner Production, 284*, Article 124714. https://doi.org/10.1016/j.jclepro.2020.124714

Adeleke, O. J., & Ali, M. M. (2021). An efficient model for locating solid waste collection sites in urban residential areas. *International Journal of Production Research, 59*(3), 798–812. https://doi.org/10.1080/00207543.2019.1709670

Aka, S., & Akyüz, G. (2018). Fuzzy goal programming approach on location-routing model for waste containers. *International Journal of Industrial and Systems Engineering, 29*(4), 413–427. https://doi.org/10.1504/IJISE.2018.094265

Antunes, A. P., Teixeira, J. C., & Coutinho, M. S. (2008). Managing solid waste through discrete location analysis: A case study in central Portugal. *Journal of the Operational Research Society, 59*(8), 1038–1046. https://doi.org/10.1057/palgrave.jors.2602422

Barrena, E., Canca, D., Ortega, F. A., & Piedera-de-la-Cuadra, R. (2019). Optimizing container location for selective collection of urban solid waste. *WIT Transactions on Ecology and the Environment, 231*, 1–9. https://doi.org/10.2495/WM180011

Beasley, J. E. (1983). Route-first cluster-second methods for vehicle routing. *Omega, 11* (4), 403–408. https://doi.org/10.1016/0305-0483(83)90033-6

Bing, X., Bloemhof, J. M., Ramos, T. R. P., Barbosa-Povoa, A. P., Wong, C. Y., & van der Vorst, J. G. A. J. (2016). Research challenges in municipal solid waste logistics management. *Waste Management, 48*, 584–692. https://doi.org/10.1016/j.wasman.2015.11.025

Blazquez, C., & Paredes-Belmar, G. (2020). Network design of a household waste collection system: A case study of the commune of Renca in Santiago, Chile. *Waste Management, 116*, 179–189. https://doi.org/10.1016/j.wasman.2020.07.027

Caramia, M., & Pizzari, E. (2022). Clustering, location, and allocation in two stage supply chain for waste management: A fractional programming approach. *Computers & Industrial Engineering, 169*, Article 108297. https://doi.org/10.1016/j.cie.2022.108297

Chavez, M. M. M., Costa, Y., & Sarache, W. (2021). A three-objective stochastic location-inventory-routing model for agricultural waste-based biofuel supply chain. *Computers & Industrial Engineering, 162*, Article 107769. https://doi.org/10.1016/j.cie.2021.107759

Darmian, S. M., Moazzeni, S., & Hvattum, L. M. (2020). Multi-objective sustainable location-districting for the collection of municipal solid waste: Two case studies. *Computers & Industrial Engineering, 150*, Article 106965. https://doi.org/10.1016/j.cie.2020.106965

Eiselt, H. A., & Marianov, V. (2014). A bi-objective model for the location of landfills for municipal solid waste. *European Journal of Operational Research, 235*(1), 187–194. https://doi.org/10.1016/j.ejor.2013.10.005

Eiselt, H. A., & Marianov, V. (2015). Location modeling for municipal solid waste facilities. *Computers & Operations Research, 62*, 305–315. https://doi.org/10.1016/j.cor.2014.05.003

Erfani, S. M. H., Danesh, S., Karrabi, S. M., & Shad, R. (2017). A novel approach to find and optimize bin locations and collection routes using a geographic information system. *Waste Management and Research, 35*(7), 776–785. https://doi.org/10.1177/0734242X17706753

Expósito-Márquez, A., Expósito-Izquierdo, C., Brito-Santana, J., & Moreno-Pérez, J. A. (2019). Greedy randomized adaptive search procedure to design waste collection routes in La Palma. *Computers & Industrial Engineering, 137*, Article 106047. https://doi.org/10.1016/j.cie.2019.106047

Ghiani, G., Laganà, D., Manni, E., Musmanno, R., & Vigo, D. (2014). Operations research in solid waste management: A survey of strategic and tactical issues. *Computers & Operations Research, 44*, 22–32. https://doi.org/10.1016/j.cor.2013.10.006

Ghiani, G., Laganà, D., Manni, E., & Toraldo, M. (2012). Capacitated location of collection sites in an urban waste management system. *Waste Management, 32*(7), 1291–1296. https://doi.org/10.1016/j.wasman.2012.02.009

Gams Development Corporation. (2013). *GAMS: The Solver Manuals*. Washington, DC., USA: CreateSpace Independent Publishing.

Geetha, S., Poonthalir, G., & Vanathi, P. T. (2009). Improved K-Means Algorithm for Capacitated Clustering Problem. *INFOCOMP Journal of Computer Science, 8*(4), 52–59.

Gläser, S., & Stücken, M. (2021). Introduction of an underground waste container system–model and solution approaches. *European Journal of Operational Research, 295*(2), 675–689. https://doi.org/10.1016/j.ejor.2021.02.060

Hemmelmayr, V. C., Doerner, K. F., Hartl, R. F., & Vigo, D. (2014). Models and algorithms for the integrated planning of bin allocation and vehicle routing in solid waste management. *Transportation Science, 48*(1), 103–120. https://doi.org/10.1287/trsc.2013.0459

Herrera-Granda, I. D., Imbaquingo-Usiña, W. G., Lorente-Leyva, L. L., Herrera-Granda, E. P., Peluffo-Ordóñez, D. H., & Rossit, D. G. (2019). Optimization of the network of urban solid waste containers: A case study. In M. Botto-Tobar, G. Pizarro, M. Zúñiga-Prieto, M. D'Armas, M. Zúñiga Sánchez (Eds.), *Technology Trends (pp.* 578–589). Communications in Computer and Information Science, vol. 895. https://doi.org/10.1007/978-3-030-05532-5_44.

Jammeli, H., Argoubi, M., & Masri, H. (2021). A bi-objective stochastic programming model for the household waste collection and transportation problem: Case of the city of Sousse. *Operational Research, 21*, 1613–1639. https://doi.org/10.1007/s12351-019-00538-5

Kim, J., Yang, J., & Ólafson, S. (2009). An optimization approach to partitional data clustering. *Journal of the Operational Research Society, 60*(8), 1069–1084. https://doi.org/10.1057/jors.2008.195

Laasasenaho, K., Lensu, A., Lauhanen, R., & Rintala, J. (2019). GIS-data related route optimization, hierarchical clustering, location optimization, and kernel density methods are useful for promoting distributed bioenergy plant planning in rural areas. *Sustainable Energy Technologies and Assessments, 32*, 47–57. https://doi.org/10.1016/j.seta.2019.01.006

Letelier, C., Blazquez, C., & Paredes-Belmat, G. (2022). Solving the bin location–allocation problem for household and recycle waste generated in the commune of Renca in Santiago, Chile. *Waste Management & Research, 40*(2), 154–164. https://doi.org/10.1177/0734242X20986610

Li, J. (2019). Optimal design of transportation distance in logistics supply chain model based on data mining algorithm. *Cluster Computing, 22*, 3943–3952. https://doi.org/10.1007/s10586-018-2544-x

Liang, Y.-C., Minanda, V., & Gunawan, A. (2022). Waste collection routing problem: A mini-review of recent heuristic approaches and applications. *Waste Management & Research, 40*(5), 519–537. https://doi.org/10.1177/0734242X211003975

Lidouh, K., & De Smet, Y. (2016). A model for spatial multi-criteria hierarchical clustering. *International Journal of Multicriteria Decision Making, 6*(1), 1–13. https://doi.org/10.1504/IJMCDM.2016.075617

Mahéo, A., Rossit, D. G., & Kilby, P. (2023). Solving the integrated bin allocation and collection routing problem for municipal solid waste: a Benders decomposition approach. *Annals of Operations Research, 322*, 441–465. https://doi.org/10.1007/s10479-022-04918-7

Matušinec, J., Hrabec, D., Šomplák, R., Nevrlý, V., & Redutskiy, Y. (2022). Cooking oils and fat waste collection infrastructure planning: A regional level outline. *Clean Technologies and Environmental Policy, 24*, 109–123. https://doi.org/10.1007/s10098-021-02087-y

Mokhtarzadeh, M., Tavakkoli-Moghaddam, R., Triki, C., & Rahimi, Y. (2021). A hybrid of clustering and meta-heuristic algorithms to solve a p-mobile hub location-allocation problem with the depreciation cost of hub facilities. *Engineering Applications of Artificial Intelligence, 98*, Article 104121. https://doi.org/10.1016/j.engappai.2020.104121

Moskvichev, O., Moskvicheva, E., & Bulatov, A. (2021). Clustering methods for determination of optimal locations of container storage and distribution centers. *Transportation Research Procedia, 54*, 461–469. https://doi.org/10.1016/j.trpro.2021.02.096

Mouton, J. P., Ferreira, M., & Helberg, S. J. (2020). A comparison of clustering algorithms for automatic modulation classification. *Expert Systems with Applications, 151*, Article 113317. https://doi.org/10.1016/j.eswa.2020.113317

Nevrlý, V., Šomplák, R., Khýr, L., Smejkalová, V., & Jadrný, J. (2019). Municipal solid waste container location based on walking distance and distribution of population. *Chemical Engineering Transactions, 76*, 553–558. https://doi.org/10.3303/CET1976093

Nevrlý, V., Šomplák, R., Smejkalová, V., Lipovský, T., & Jadrný, J. (2021). Location of municipal waste containers: Trade-off between criteria. *Journal of Cleaner Production, 278*, Article 123445. https://doi.org/10.1016/j.jclepro.2020.123445

Olapiriyakul, S., Pannakkong, W., Kachapanya, W., & Starita, S. (2019). Multi-objective optimization model for sustainable waste management network design. *Journal of Advanced Transportation, 2019*, 3612809. https://doi.org/10.1155/2019/3612809

Oliaei, A. S. B., & Fataei, E. (2016). Breakdown of urban waste repository location using GIS (Case study District 3 Region 1 Tabriz), Iran. *Ecology, Environment and Conservation, 22*(4), 2115–2120.

Olmez, B. O., Gultekin, C., Balcik, B., Ekici, A., & Özener, O.Ö. (2022). A variable neighborhood search based matheuristic for a waste cooking oil collection network design problem. *European Journal of Operational Research, 302*(1), 187–202. https://doi.org/10.1016/j.ejor.2021.12.018

Ordoñez, A., Ordoñez, H., Corrales, J. C., Cobos, C., Wives, L. K., & Thom, L. H. (2017). Grouping of business processes models based on an incremental clustering algorithm using fuzzy similarity and multimodal search. *Expert Systems with Applications, 37*, 163–177. https://doi.org/10.1016/j.eswa.2016.08.061

Pla-Sacristán, E., González-Díaz, I., Martínez-Cortés, T., & Díaz-de-María, F. (2019). Finding landmarks within settled areas using hierarchical density-based clustering and meta-data from publicly available images. *Expert Systems with Applications, 123*, 315–327. https://doi.org/10.1016/j.eswa.2019.01.046

Rabbani, M., Heidari, R., Farrokhi-Asl, H., & Rahimi, N. (2018). Using metaheuristic algorithms to solve a multi-objective industrial hazardous waste location-routing problem considering incompatible waste types. *Journal of Cleaner Production, 170*, 227–241. https://doi.org/10.1016/j.jclepro.2017.09.029

Rabbani, M., Sadati, S. A., & Farrokhi-Asl, H. (2020). Incorporating location routing model and decision making techniques in industrial waste management: Application in the automotive industry. *Computers & Industrial Engineering, 148*, Article 106692. https://doi.org/10.1016/j.cie.2020.106692

Rathore, P., Sarmah, S. P., & Singh, A. (2020). Location–allocation of bins in urban solid waste management: A case study of Bilaspur city, India. *Environment, Development and Sustainability, 22*, 3309–3331. https://doi.org/10.1007/s10668-019-00347-y

Rossit, D. G., Nesmachnow, S., & Toutouh, J. (2019a). A bi-objective integer programming model for locating garbage accumulation points: A case study. *Revista Facultad de Ingeniería, 93*. https://doi.org/10.17533/udea.redin.20190509

Rossit, D. G., Nesmachnow, S., & Toutouh, J. (2019b). Municipal Solid Waste Management in Smart Cities: Facility Location of Community Bins. In S. Nesmachnow, & L. Hernández Callejo (Eds.), *Smart Cities* (vol 978, pp. 102–115). Cham: Springer. https://doi.org/10.1007/978-3-030-12804-3_9.

Rossit, D. G., Toutouh, J., & Nesmachnow, S. (2020). Exact and heuristic approaches for multi-objective garbage accumulation points location in real scenarios. *Waste Management, 105*, 467–481. https://doi.org/10.1016/j.wasman.2020.02.016

Sbihi, A., & Eglese, R. W. (2010). Combinatorial optimization and Green Logistics. *Annals of Operations Research, 175*, 159–175. https://doi.org/10.1007/s10479-009-0651-z

Shi, J., Wang, R., Chen, W., Xing, L., & Jin, M. (2020). Bi-objective design of household E-waste collection with public advertising and competition from informal sectors. *Waste Management, 102*, 65–75. https://doi.org/10.1016/j.wasman.2019.10.018

Slavík, J., Dolejš, M., & Rybová, K. (2021). Mixed-method approach incorporating Geographic information system (GIS) tools for optimizing collection costs and convenience of the biowaste separate collection. *Waste Management, 134*, 177–186. https://doi.org/10.1016/j.wasman.2021.07.018

Soltani, A., Hewage, K., Reza, B., & Sadiq, R. (2015). *Multiple stakeholders in multi-criteria decision-making in the context of Municipal Solid Waste Management: A review. Waste management, 35*, 318–328. https://doi.org/10.1016/j.wasman.2014.09.010

Toutouh, J., Rossit, D., & Nesmachnow, S. (2020). Soft computing methods for multi-objective location of garbage accumulation points in smart cities. *Annals of Mathematics and Artificial Intelligence, 88*, 105–131. https://doi.org/10.1007/s10472-019-09647-5

Tralhão, L., Coutinho-Rodrigues, J., & Alçada-Almeida, L. (2010). A multi-objective modeling approach to locate multi-compartment containers for urban-sorted waste. *Waste Management, 30*(12), 2418–2429. https://doi.org/10.1016/j.wasman.2010.06.017

Yaakoubi, E., Benabdouallah, M., & Bojji, C. (2018). Heuristic approaches for waste containers location problem and waste collection routes optimisation in an urban area. *International Journal of Environment and Waste Management, 21*(4), 269–286. https://doi.org/10.1504/IJEWM.2018.093436

Yalcinkaya, S., & Uzer, S. (2022). A GIS-based multi-criteria decision support model for planning municipal solid waste collection points: A case study of Çağdaş Neighbourhood, Çiğli District, Izmir, Turkey. *Waste Management and Research, 40*(8), 1297–1310. https://doi.org/10.1177/0734242X211063733

Viktorin, A., Senkerik, R., Pluhacek, M., Kadavy, T., & Zamuda, A. (2019). Distance based parameter adaptation for success-history based differential evolution. *Swarm and Evolutionary Computation, 50*, Article 100462. https://doi.org/10.1016/j.swevo.2018.10.013

Vu, H. L., Ng, K. T. W., & Bolingbroke, D. (2018). Parameter interrelationships in a dual phase GIS-based municipal solid waste collection model. *Waste Management, 78*, 257–270. https://doi.org/10.1016/j.wasman.2018.05.050

Wang, Y., Assogba, K., Liu, Y., Ma, X., Xu, M., & Wang, Y. (2018). Two-echelon location-routing optimization with time windows based on customer clustering. *Expert Systems with Applications, 104*, 244–260. https://doi.org/10.1016/j.eswa.2018.03.018

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Xu, Q., Zhang, Q., Liu, J., & Luo, B. (2020). Efficient synthetical clustering validity indexes for hierarchical clustering. *Expert Systems with Applications, 151*, Article 113367. https://doi.org/10.1016/j.eswa.2020.113367

Zamorano, M., Molero, E., Grindlay, A., Rodríguez, M. L., Hurtado, A., & Calvo, F. J. (2009). A planning scenario for the application of geographical information systems in municipal waste collection: A case of Churriana de la Vega (Granada, Spain). *Resources, Conservation and Recycling, 54*(2), 123–133. https://doi.org/10.1016/j.resconrec.2009.07.001

Zbib, H., & Laporte, G. (2020). The commodity-split multi-compartment capacitated arc routing problem. *Computers and Operations Research, 122*, Article 104994. https://doi.org/10.1016/j.cor.2020.104994