Computers & Industrial Engineering 217 (2026) 112057

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Joint optimisation of bin location, allocation, type selection, and collection frequency under stochastic waste generation

Massimo Bertolini, Davide Castellano *, Davide Mezzogori

"Enzo Ferrari" Department of Engineering, Università Degli Studi di Modena e Reggio Emilia, Via P. Vivarelli, 10–41125 Modena, Italy

A R T I C L E I N F O

Keywords:
Waste management
Optimisation
Uncertainty
Heuristics
Coordination

A B S T R A C T

This paper addresses the joint optimisation of bin location, type selection, allocation, and collection frequency for urban waste management, explicitly considering stochastic waste generation. Waste originates from multiple nodes, also potential bin sites, with bins having specific capacities and installation/operating costs. Bin emptying is coordinated via a novel mechanism inspired by the periodic-review Joint Replenishment Problem (JRP). The comprehensive cost function covers bin placement, waste delivery, JRP-based collection ordering (major/minor), penalties for bin over/underutilisation and truck overcapacity, and disposal facility operations. The objective is to find optimal bin locations, types, allocations, and collection frequencies that minimise the long-run expected total system cost rate. Given the problem's inherent complexity, a robust solution framework is employed, embedding an efficient heuristic algorithm (optimising collection frequency) within a standard metaheuristic to navigate the combinatorial search space. Numerical experiments evaluate the solution method's performance and the model's behaviour under diverse operating conditions. Key results demonstrate that the covariance structure of waste generation significantly dictates optimal system configuration and overall cost-effectiveness; high inter-node correlation, for instance, necessitates more flexible capacity provision and frequency planning to effectively mitigate amplified penalty risks. The study also reveals strong system sensitivity to variations in truck and bin capacities, as well as key cost parameters. Crucially, allowing the system to adapt its configuration consistently outperforms static approaches, highlighting the significant value of dynamic resource reallocation. These findings offer valuable, data-driven managerial insights for designing robust and cost-effective waste management systems, particularly when operating under inherent uncertainty.

1. Introduction

Municipal solid waste (MSW) management represents one of the most persistent and resource-intensive challenges facing modern urban centres. The confluence of rapid urbanisation, population growth, and evolving consumption patterns has led to an unprecedented increase in both the volume and complexity of waste streams, including challenging new categories like e-waste (Ravindra & Mor, 2019). For local authorities, the task of efficiently collecting, transporting, treating, and disposing of this waste is a critical public service, essential for safeguarding public health, preserving environmental quality, and maintaining urban aesthetics. The stakes are substantial; MSW management can consume up to 50% of municipal budgets in some developing nations, and mismanagement poses severe risks, including soil and water contamination from leachate, air pollution, and the spread of disease (Kumar & Agrawal, 2020; Singh, 2019). Achieving an effective MSW system therefore requires solutions that thoughtfully balance economic sustainability, environmental responsibility, and social equity.

Within the broader scope of waste management, the collection phase is widely recognised as the most operationally complex and costly component, often accounting for 50–70% of the total system expenditure (Tavares et al., 2009). Operational inefficiencies at this stage have cascading negative effects, leading not only to inflated costs but also to increased vehicle emissions, traffic congestion, and significant public health concerns such as vector proliferation and unpleasant odours (Blazquez & Paredes-Belmar, 2020). Consequently, the design of the waste collection infrastructure–the interface between waste generators and the collection system–is a tactical decision of paramount importance. This design process involves a set of deeply interconnected decisions that collectively define the system's efficiency and resilience (Hemmelmayr et al., 2014).

Effectively tackling this design challenge requires a systematic and

* Corresponding author.
E-mail address: davide.castellano@unimore.it (D. Castellano).

https://doi.org/10.1016/j.cie.2026.112057
Received 4 September 2025; Received in revised form 27 February 2026; Accepted 22 April 2026
Available online 26 April 2026

quantitative approach that addresses four fundamental and interdependent pillars. First, bin location determines where collection points are physically placed, a decision that directly impacts citizen convenience and subsequent routing efficiency (Rossit & Bard, 2025). Second, bin type selection involves choosing the appropriate capacity and design for each bin, creating a crucial trade-off between upfront capital investment and the risk of service failures like overflows (Nevrlý et al., 2021). Third, waste allocation dictates which households or commercial entities are assigned to each collection point, thereby determining the waste accumulation profile at each bin (Letelier et al., 2022). Finally, collection frequency establishes the schedule for emptying the bins, a decision that embodies the core operational trade-off between minimising costly collection tours and maintaining an adequate level of service to prevent overflows (Ghiani et al., 2012). The inherent interdependencies among these pillars mean that they cannot be optimised in isolation; a decision to use smaller bins, for instance, necessitates more frequent collections or a denser network of bin locations.

The primary challenge complicating these interdependent decisions is the inherent uncertainty in waste generation. Waste production is not a deterministic process but is fundamentally stochastic, exhibiting significant variability influenced by daily and weekly cycles, seasonality, public events, socioeconomic factors, and unpredictable human behaviour (Dyson & Chang, 2005; Singh, 2019). Traditional deterministic optimisation models, which rely on average or forecasted waste generation rates, are ill-equipped to handle these real-world fluctuations. This limitation is a key motivator for the development of stochastic models, whose value has been repeatedly demonstrated in the literature (Spinelli et al., 2025). Ignoring uncertainty inevitably leads to system designs that are either inefficiently over-provisioned with oversized bins and overly frequent collections, or dangerously under-provisioned, resulting in frequent bin overflows, sanitation problems, and widespread citizen dissatisfaction (Markov et al., 2020). To design a system that is both cost-effective and robust, it is essential to move beyond deterministic assumptions and explicitly incorporate the stochastic nature of waste generation.

This paper addresses this critical need by proposing a comprehensive, integrated stochastic optimisation framework to co-optimise the four core tactical decisions: bin location, type selection, allocation, and collection frequency. Our objective is to design a system that minimises the long-run expected total cost, explicitly accounting for the operational risks and penalties arising from stochastic variability. The framework makes several novel contributions. First, it models waste generation not only through its mean and variance but also through its covariance structure, capturing the critical, yet often overlooked, statistical interdependence of waste generation rates across different urban zones. Second, it coordinates the complex collection schedules across potentially thousands of bins via a novel mechanism inspired by the periodic-review joint replenishment problem (JRP) from inventory theory, providing an economically grounded basis for frequency planning. Given the inherent complexity of the resulting mixed-integer nonlinear program (MINLP), we implement a robust solution framework to solve it. Specifically, we demonstrate that combining a standard metaheuristic to navigate the combinatorial search space with an embedded, efficient heuristic algorithm for frequency optimisation provides a computationally feasible pathway to evaluate the model and extract insights.

This research seeks to answer critical questions for waste management planners: How does the statistical correlation of waste generation between different areas impact the optimal system configuration and its cost-effectiveness? What is the quantifiable economic value of designing a system that can adapt its configuration in response to changing parameters, compared to a static design? By solving the proposed model, our numerical experiments demonstrate that the covariance structure is a first-order driver of optimal system design, necessitating more flexible capacity and frequency planning in highly correlated environments. Furthermore, our findings highlight the substantial value of dynamic resource reallocation, showing that an adaptable system consistently outperforms static approaches. These insights provide valuable, data-driven guidance for designing robust and cost-effective MSW management systems capable of operating efficiently under uncertainty.

The remainder of this paper is organised as follows. Section 2 provides a structured review of the relevant literature, positioning our work within the broader context of waste management optimisation. Section 3 describes the problem faced and introduces notation and assumptions. Section 4 develops the mathematical model, formalising the optimisation problem and its underlying assumptions. Section 5 describes the hybrid solution methodology employed to solve the problem. Section 6 presents the numerical experiments, assessing the model's computational tractability, quantifying the value of the stochastic solution, and deriving actionable managerial insights. Finally, Section 7 concludes the paper, summarising key findings, discussing model limitations, and finally suggesting directions for future research.

2. Literature review

The optimisation of MSW management is a complex, multi-faceted challenge that has garnered significant attention in the operations research community. Effective planning requires integrated decisions spanning strategic, tactical, and operational levels to balance economic efficiency, environmental responsibility, and service quality. Strategic decisions typically concern the siting of major facilities like landfills and transfer stations (Eiselt & Marianov, 2015), while operational decisions focus on daily tasks like vehicle routing. This paper is situated at the tactical level, which bridges these two horizons by focusing on the design of the collection infrastructure itself. This review synthesises the extensive body of literature relevant to this tactical problem, structuring the discussion around three key pillars: foundational decision problems, the evolution towards integrated models, and the critical incorporation of uncertainty. The review culminates by identifying a distinct research gap that this work aims to fill.

2.1. Foundational decision problems in waste management

Historically, research in waste management logistics has often decomposed the tactical planning problem into two primary, independently solved subproblems: location-allocation and vehicle routing. The waste bin location problem (WBLP), a specialised form of the facility location problem, focuses on determining the optimal placement of collection containers. A comprehensive review by Rossit & Nesmachnow (2022) highlights the diversity of objectives in this domain, which range from minimising investment costs (Cavallin et al., 2020) and citizen travel distance (Bautista & Pereira, 2006) to maximising population coverage (Letelier et al., 2022). Methodologies vary widely, from classical mixed-integer linear programming (MILP) models to GIS-based approaches that leverage spatial data to identify suitable sites (Erfani et al., 2017). Recent work has introduced novel criteria, such as identifying litter hotspots through geotagged images to guide bin placement, thereby enhancing accessibility and encouraging recycling behaviour (Azri et al., 2023). While essential, these models often treat routing costs simplistically, for instance, by using collection frequency as a proxy (Rossit et al., 2020), without explicitly modelling the complex tour-building process.

Concurrently, the vehicle routing problem (VRP) has been extensively studied to optimise waste collection routes for a pre-determined set of collection points. Given the periodic nature of waste collection, variants such as the periodic vehicle routing problem (PVRP) are particularly relevant (Angelelli & Speranza, 2002). The seminal work of Beltrami & Bodin (1974) laid much of the groundwork for applying VRP heuristics to MSW collection. However, solving the VRP in isolation assumes that the set of bin locations and their required service frequencies are fixed inputs. This sequential approach–first locating bins, then designing routes–is inherently suboptimal, as the cost of routing is highly dependent on the spatial configuration of the bins, and vice-versa.

2.2. The evolution towards integrated models

Recognizing the limitations of sequential decision-making, a significant stream of research has focused on integrated models, most notably the location-routing problem (LRP). The LRP simultaneously determines facility locations (or in this context, bin locations) and vehicle routes, capturing the fundamental trade-off between the fixed costs of establishing sites and the variable costs of servicing them. The survey by Prodhon & Prins (2014) highlights the maturity of this field and the prevalence of heuristic and metaheuristic solution methods for these NP-hard problems.

In the MSW context, Hemmelmayr et al. (2014) presented a key integrated model for the waste bin allocation and routing problem (WBARP), developing a metaheuristic that combines a variable neighbourhood search (VNS) for routing with an exact MILP for allocation. More recently, Mahéo et al. (2023) tackled a similar integrated problem using a Benders decomposition approach, which proved highly effective for real-world instances. Other studies have integrated different facets of the problem; for instance, Olmez et al. (2022) developed a VNS-based metaheuristic for a waste cooking oil collection network that integrates locating collection centres, assigning households, determining bin numbers, and routing. These integrated approaches represent a significant advancement, as they provide a more holistic and system-wide optimal solution. However, a vast majority of the LRP literature in waste management relies on deterministic assumptions for key parameters, particularly the rate of waste generation. This overlooks the inherent variability and randomness that characterise real-world waste production, which can lead to system designs that are either inefficient or prone to failure.

2.3. Incorporating uncertainty in waste management optimisation

The inherently stochastic nature of waste generation, which fluctuates due to factors like seasonality, socioeconomic trends, and random human behaviour, is a critical challenge. Acknowledging this, a growing body of research has moved beyond deterministic models to explicitly incorporate uncertainty.

A primary paradigm for this is stochastic optimisation, which models uncertain parameters using probability distributions. A prominent framework in this area is the stochastic inventory routing problem (SIRP), which optimises collection (delivery) schedules and routes for locations with stochastic demand to avoid stockouts (overflows). Markov et al. (2020) made a significant contribution by developing an SIRP model for recyclable waste collection that includes dynamic, probability-based costs for container overflows and route failures derived from a non-stationary demand forecast. Following this trend, Spinelli et al. (2025) proposed a multi-stage stochastic optimisation model for an inventory routing problem (IRP) in recyclable waste collection, developing a rolling horizon heuristic to manage computational complexity while highlighting the significant value of the stochastic model over its deterministic counterpart. Similarly, Cuellar-Usaquén et al. (2025) addressed a dynamic recycling collection problem with uncertainty in both material availability and quality, proposing a stochastic lookahead method that anticipates future costs and significantly outperforms myopic policies. While these models are highly sophisticated in their treatment of uncertainty, they typically focus on optimising collection policies for a pre-existing and fixed network of bins. They address the operational and short-term tactical questions of when and how to collect, but not the primary tactical decision of where to locate the bins and what types of bins to use in the first place.

Other paradigms have also been employed. Robust optimisation, which hedges against worst-case scenarios without requiring full probability distributions, has been applied by Rossit & Bard (2025) to a bi-objective bin location problem with uncertain waste generation. Other methods include chance-constrained programming, used to handle stochasticity in smart city waste management (Akbarpour et al., 2021) and household waste collection (Jammeli et al., 2021), and fuzzy logic, used to model uncertainty in complex collection problems (Aliahmadi et al., 2021).

Furthermore, the rise of the internet of things (IoT) has spurred research into dynamic and real-time operational models. Studies leveraging data from "smart bins" to optimise daily collection routes exemplify this trend. For instance, the work by Roy et al. (2021) provides a clear example of this development, initially proposing a conceptual model for IoT-based smart bin allocation and routing with a basic heuristic, which was later expanded into a comprehensive case study featuring a sophisticated hybrid metaheuristic framework (Roy et al., 2022). Similarly, Kim et al. (2025) developed a clustered vehicle routing approach for smart waste collection. These IoT-based approaches are powerful for improving operational efficiency by reacting to real-time fill levels. However, they operate within the constraints of the existing infrastructure. The long-term efficiency of such smart systems is ultimately governed by the underlying tactical design of the network–the locations, types, and baseline collection frequencies of the bins.

2.4. Research gap and contributions

A review of the literature reveals a clear and compelling research gap. While the field has matured in three distinct streams–(1) integrated deterministic models like the LRP, (2) sophisticated stochastic models like the SIRP that optimise operations over fixed networks, and (3) dynamic operational models leveraging IoT data–a holistic framework that bridges these areas remains elusive. Specifically, deterministic models fail to capture the real-world variability that leads to costly overflows and inefficient resource use. Conversely, most stochastic models take the physical network of bins as a given, precluding the optimisation of the infrastructure itself.

This paper identifies the following specific gaps in the existing body of knowledge:

1. There is a lack of a unified model that simultaneously optimises the four fundamental, interdependent tactical decisions: bin location, waste allocation, bin type selection, and collection frequency.
2. While stochasticity in waste generation is increasingly considered, the covariance structure between different generation nodes–which captures the tendency for waste levels in different areas to fluctuate together–is almost universally ignored. This is a critical oversight, as high correlation can amplify system-wide risks.
3. The coordination of collection schedules for hundreds or thousands of bins is a major practical challenge. The application of a structured, inventory-theoretic approach like the JRP to establish economically grounded, coordinated collection frequencies within an integrated, stochastic location-allocation model is novel.

This paper addresses these gaps by making the following primary contributions to the literature:

- It proposes a novel, integrated stochastic optimisation framework that jointly determines bin location, allocation, type selection, and collection frequency to minimise the long-run expected total system cost.
- It explicitly models the stochastic nature of waste generation using its mean, variance, and, crucially, its covariance structure, allowing for a more realistic assessment of system-wide risks associated with correlated demand peaks.
- It introduces a JRP-inspired coordination mechanism to optimise collection frequencies, providing a practical and economically

grounded method for managing collection schedules across a large network of bins.

- It demonstrates that this highly complex MINLP can be reliably solved by embedding a JRP heuristic within a standard metaheuristic algorithm. The subsequent thorough computational analysis not only validates the tractability of the proposed formulation but also yields critical managerial insights into the significant impact of waste generation covariance and the substantial economic benefits of an adaptive system design over a static one.

3. Problem description, notation, and assumptions

An urban area contains many waste sources (nodes), each able to host a collection point (GAP). Installing a bin at a node requires upfront investment and periodic maintenance. Sources generate waste stochastically and incur delivery costs depending on quantity and distance to their bin.

The model considers the availability of multiple bin types. From a tactical planning perspective, a "bin type" is primarily defined by its key decision-making attributes: its capacity and its associated investment and maintenance costs. These types represent the range of real-world options available to municipal planners, such as standard 1100 L wheeled dumpsters, larger 3000 L static compactors, or modern semi-underground collection systems. The selection of a specific bin type is therefore a strategic decision that creates a fundamental trade-off: higher-capacity bins require greater initial investment but can reduce the required collection frequency and the risk of costly overflows, whereas smaller bins are cheaper to install but may necessitate more frequent service. Furthermore, bin types can be specialised for different waste streams (e.g., general municipal solid waste, mixed recyclables, or organic waste), allowing the framework to be applied to segregated collection systems. Each bin or homogeneous bin family is then emptied at a fixed interval by capacitated trucks that depart and return at the central disposal facility.

This study introduces a cost function covering bin placement, waste delivery to bins, collection task activation, facility operation, and penalties for exceeding or underutilising bin capacity, as well as for surpassing truck capacity. The objective is to determine the optimal bin location, type selection, waste allocation, and collection frequency that minimises the long-run expected total system cost rate.

3.1. Notation

The mathematical model is developed using the following notation:

Sets:	
$\mathbb{R}_+^*$	Positive real numbers
$\mathbb{N}$	Natural numbers
$\mathbb{I}$	Set of nodes
$\mathbb{I}'$	Subset of $\mathbb{I}$ that denotes the set of nodes where a bin is positioned
$\mathbb{K}$	Set of bin types
Indices:	
i, j	Node
k	Bin type
Decision variables:	
T	Basic collection period [unit time]
m_i	Integer multiplier of the bin located in node $i \in \mathbb{I}'$
x_{ik}	Binary variable that specifies whether the bin in node $i \in \mathbb{I}'$ is of type $k \in \mathbb{K}$ ($x_{ik} = 1$) or not ($x_{ik} = 0$)
z_{ij}	Binary variable that specifies whether the waste source in node $i \in \mathbb{I}$ commits garbage to the bin in node $j \in \mathbb{I}$ ($z_{ij} = 1$) or not ($z_{ij} = 0$)
Parameters:	
A	Major cost of a waste collection order, which may consist, e.g., in the cost of arranging a given sequence of collection tasks [$/collection order]
a_{ik}	Minor cost of arranging a waste collection order for the bin of type $k \in \mathbb{K}$ located in node $i \in \mathbb{I}'$ [$/collection order]
π_{0i}	Penalty cost per unit quantity exceeding the capacity of the bin located in node $i \in \mathbb{I}'$ [$/unit quantity]

(*continued on next column*)

(*continued*)

π_{1i}	Penalty cost per unit quantity below the capacity of the bin located in node $i \in \mathbb{I}'$ [$/unit quantity]
c_{1ij}	Cost per unit distance per unit quantity of waste committed from node $i \in \mathbb{I}$ to the bin located in node $j \in \mathbb{I}'$ [$/unit length/unit quantity]
c_{2ik}	Cost rate–that is, investment × interest rate–to position a bin of type $k \in \mathbb{K}$ in node $i \in \mathbb{I}$ [$/unit time]
r_{ij}	Distance between nodes $i \in \mathbb{I}$ and $j \in \mathbb{I}$ [unit length]
C_k	Capacity of the bin of type $k \in \mathbb{K}$ [unit quantity]
d_i	Average waste generation rate of node $i \in \mathbb{I}$ [unit quantity/unit time]
σ_i	Standard deviation of the waste generation rate of node $i \in \mathbb{I}$ [unit quantity/unit time]
σ_{ij}	Covariance of the waste generation rates of nodes i and j, with $\sigma_{ii} \equiv \sigma_i^2$ [unit quantity/unit time]^2
h	Operating cost rate per unit waste quantity of the central disposal facility [$/unit quantity/unit time]
θ	Penalty cost for exceeding the truck capacity, which may be related, e.g., to the cost of activating extra trips needed to complete the waste collection task [$/unit quantity]
Q	Truck capacity [unit quantity]
Functions and operators:	
f	Standard normal probability density function (p.d.f.)
F	Standard normal cumulative distribution function (c.d.f.)
ψ^+	Standard normal first-order loss function
ψ^-	Standard normal first-order complementary loss function
$E[\bullet]$	Mathematical expectation
$\mathrm{Var}(\bullet)$	Variance
$\mathrm{Cov}(X, Y)$	Covariance between the random variables X and Y
x^+	Maximum between 0 and x–that is, $x^+ \equiv \max\{0, x\}$
x^-	Maximum between 0 and $-x$–that is, $x^- \equiv \max\{0, -x\}$
$\lfloor x \rfloor$	Largest integer smaller than, or equal to, x
$\mathrm{lcm}(\mathbf{q})$	Least common multiple (LCM) of the integer vector $\mathbf{q}$
$a \bmod b$	Modulo–that is, the remainder after division of a by b, with a and b positive integers
$\mathrm{card}(\bullet)$	Cardinality of a set
Random variables:	
Z_i	Total quantity of waste accumulated in the bin located in node $i \in \mathbb{I}'$ in the time between two consecutive emptying–that is, in the collection period
X_i	Waste generated in the unit time in node $i \in \mathbb{I}$

Additional notation will be introduced as needed.

3.2. Assumptions

The following main assumptions are adopted to develop the mathematical model:

1. All nodes generate waste, and the waste generation rate of node $i \in \mathbb{I}$ is the random variable
2. X_i with mean d_i and standard deviation σ_i. Moreover, considering any set $\mathscr{T}$ of disjoint intervals, $\left\{X_i^{(t)}\right\}_{t\in\mathscr{T}}$ is a family of independent and identically distributed random variables.
3. A bin can be positioned in any node $i \in \mathbb{I}$, and nodes where a bin is located keep generating wastes–that is, positioning a bin does not shut off the waste generation process. A node can have either a single bin (of a single type) or none, and each bin type can be positioned in any node. If $\mathbb{I}'$ is the set of nodes including a bin, then $\mathbb{I}' \subseteq \mathbb{I}$, and $i \in \mathbb{I}' \Leftrightarrow \sum_k x_{ik} > 0$.
4. Every T time units some costly activities connected to planning collection routes are carried out. These activities pertain to a major collection order, and the corresponding major ordering cost is A.
5. The bin located in node $i \in \mathbb{I}'$ is visited every $T_i = m_i T$ time units, and a_{ik} is the (minor waste collection ordering) cost incurred to visit it, being $k \in \mathbb{K}$ the type of the bin placed in the node.
6. Each node delivers waste to a single bin.
7. The total quantity of waste accumulated in the bin located in node $i \in \mathbb{I}'$ in the time between two consecutive emptying–that is, the random variable Z_i–is a Gaussian random variable. This assumption is grounded in the central limit theorem (CLT)

applied to the temporal domain; specifically, Z_i represents the aggregation of waste generation over the collection interval T_i. Even if generation rates (X_i) follow arbitrary distributions, their sum over the collection cycle tends toward normality, provided the interval is sufficiently long (Silver et al., 2017).

8. The bin of type $k \in \mathbb{K}$ located in node $i \in \mathbb{I}'$ has maximum capacity C_k. If the bin is saturated, any further waste delivered to it will generate the unit cost π_{0i}. Moreover, at the time the bin is visited by the collection truck to be emptied, the cost π_{1i} is paid for each unit of waste below the bin capacity. In other words, there exists a penalty for any inefficient use of bin capacity.
9. The model imposes no constraints on maximum truck travel distance or time. Constrained route optimisation (VRP) is not explicitly addressed. Consequently, while additional trips or trucks might be necessary to complete waste collection, their direct impact on overall system performance is not explicitly evaluated. Instead, truck capacity constraints provide an indirect assessment of routing feasibility. It is also assumed that bins are entirely emptied during each collection.
10. An infinite time horizon is considered.

4. Mathematical model and problem formulation

We address the tactical planning problem faced by a waste management planner aiming to design a cost-efficient waste collection system. The objective is to minimise the long-run expected total system cost rate by jointly optimising four decision variables:

1. Bin location: Selecting the specific nodes from a set of potential sites where collection points will be installed.
2. Bin type selection: Determining the capacity and associated cost profile for the bin at each chosen location.
3. Waste allocation: assigning each waste generation source to a single bin.
4. Collection frequency: Establishing a coordinated schedule where bins are emptied at integer multiples of a base period, modelled via a JRP structure to balance major routing costs against minor service costs.

To improve clarity and guide the reader through the mathematical formulation, Table 1 provides a high-level summary mapping the key decision variables to their corresponding strategic and operational planning decisions.

Our model bridges strategic infrastructure planning and operational logistics. A fundamental assumption is that waste generation rates are correlated random variables. This allows the model to explicitly account for inter-node covariance and quantify the expected penalty costs

Table 1
Mapping of key decision variables to planning decisions.

Decision variable type	Symbol	Planning decision represented
Tactical decisions (system design)		
Bin location	x_{ik}(specifically, which i has $\sum_k x_{ik} > 0$)	Where to physically place collection bins.
Bin type selection	x_{ik}, (specifically, which k is chosen for a given $i \in \mathbb{I}'$)	What size/capacity of bin to install at each location.
Waste allocation	z_{ij}	Which households/sources are assigned to which bin.
Operational decisions (system scheduling)		
Base collection period	T	The fundamental frequency of collection tours.
Collection multiplier	m_i	The specific collection interval for bin i (as a multiple of T)

associated with operational risks, such as bin overflows, capacity underutilisation, and truck overcapacity.

We begin the development of the model with the component of the cost function dependent on the waste collection frequency. This component is modelled using a JRP structure (Khouja & Goyal, 2008), reflecting the practical grouping of bins for coordinated collection at common intervals.

We first determine the mean and the variance of Z_i. Let W_i be the total amount of waste that is delivered to the bin located in node $i \in \mathbb{I}'$ in the time unit. Since $W_i = \sum_{j\in\mathbb{I}} z_{ji} X_j$, the mean and the variance of W_i are

$$D_i = E[W_i] = E\left[\sum_{j\in\mathbb{I}} z_{ji} X_j\right] = \sum_{j\in\mathbb{I}} z_{ji} d_j$$

and

$$\begin{aligned}\Sigma_i^2 = \mathrm{Var}(W_i) = \mathrm{Var}\left(\sum_{j\in\mathbb{I}} z_{ji} X_j\right) &= \sum_{j_1,j_2\in\mathbb{I}} z_{j_1 i} z_{j_2 i} \mathrm{Cov}\left(X_{j_1}, X_{j_2}\right)\\ &= \sum_{j\in\mathbb{I}} z_{ji}^2 \sigma_j^2 + \sum_{j_1,j_2\in\mathbb{I}, j_1\neq j_2} z_{j_1 i} z_{j_2 i} \mathrm{Cov}\left(X_{j_1}, X_{j_2}\right)\\ &= \sum_{j\in\mathbb{I}} z_{ji}^2 \sigma_j^2 + \sum_{j_1,j_2\in\mathbb{I}, j_1\neq j_2} z_{j_1 i} z_{j_2 i} \sigma_{j_1 j_2},\end{aligned}$$

respectively. It follows that the random variable Z_i–that is, the quantity of waste accumulated in the bin located in node $i \in \mathbb{I}'$ in the time between two consecutive emptying, which is $m_i T$–has mean

$$\mu_i = E[Z_i] = m_i T E[W_i] = m_i T D_i$$

and variance

$$\varsigma_i^2 = \mathrm{Var}(Z_i) = m_i T \mathrm{Var}(W_i) = m_i T \Sigma_i^2.$$

Fig. 1 gives an illustrative representation of the dynamics of a two-bin system with $m_1 = 4$ and $m_2 = 2$. This figure also shows the situation in which a bin becomes full, but keeps gathering waste in the sense that people, finding a totally full bin, tend to leave the garbage on the ground near the bin itself. This behaviour clearly leads to some additional cost that has to be considered in the formulation. The amount of waste that exceeds the capacity of the bin located in node $i \in \mathbb{I}'$ in the single collection cycle is $\left(Z_i - \sum_{k\in\mathbb{K}} x_{ik} C_k\right)^+$, whose expected value is

$$E\left[\left(Z_i - \sum_{k\in\mathbb{K}} x_{ik} C_k\right)^+\right] = \varsigma_i [f(p_i) + p_i(F(p_i) - 1)] = \varsigma_i \psi^+(p_i),$$

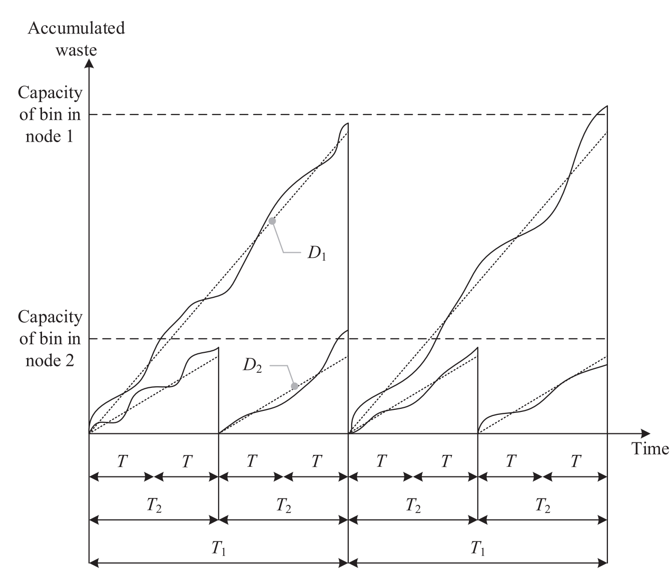

Fig. 1. Illustrative representation of the dynamics of a two-bin system with $m_1 = 4$ and $m_2 = 2$.

where $p_i = \left(\sum_{k\in\mathbb{K}} x_{ik}C_k - \mu_i\right)/\varsigma_i$, being $\varsigma_i = \sqrt{\varsigma_i^2}$. Then, the expected cost per collection cycle related to the quantity of waste exceeding the capacity of the hypothetical bin located in node $i \in \mathbb{I}'$ is $\pi_{0i}\varsigma_i\psi^+(p_i)$.

According to assumption n. 7, some cost related to the inefficient exploitation of the bin capacity may arise. If we consider the bin located in node $i \in \mathbb{I}'$, the unexploited bin capacity in the single collection cycle is $\left(Z_i - \sum_{k\in\mathbb{K}} x_{ik}C_k\right)^-$, whose expected value is

$$E\left[\left(Z_i - \sum_{k\in\mathbb{K}} x_{ik}C_k\right)^-\right] = \varsigma_i(f(p_i) + p_iF(p_i)) = \varsigma_i\psi^-(p_i).$$

Then, the expected cost per collection cycle related to the inefficient exploitation of the capacity of the bin located in node $i \in \mathbb{I}'$ is $\pi_{1i}\varsigma_i\psi^-(p_i)$.

The MSW system features a central disposal facility with an operating cost per unit time, denoted h'. This cost comprises both variable (quantity-dependent) and fixed components, analogous to warehouse operating cost assessment (Azzi et al., 2014; Castellano & Glock, 2021). The unit stockholding cost rate h is related to h' by $h = h'/V$, where V is the expected average quantity of waste processed by the facility. Assuming $V = \sum_{i\in\mathbb{I}'} m_iTD_i/2$, the facility's expected operating cost rate becomes $h' = h\sum_{i\in\mathbb{I}'} m_iTD_i/2$.

We now evaluate the cost related to exceeding the truck capacity in collection cycles. We first observe that, given the vector $\mathbf{m} = (m_i)_{i\in\mathbb{I}'}$ of integer multipliers, the collection process over all bins is a regenerative process that repeats itself probabilistically every MT time units, where $M = \text{lcm}(\mathbf{m})$. In the lth collection cycle, some bins are visited while others not. We introduce the following additional variable, defined for each $i \in \mathbb{I}'$ and $l = 1, 2, \cdots, M$:

$$y_{il} = \begin{cases} 1 & \text{if } l \bmod m_i = 0, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $y_{il} = 0$ if the bin in node $i \in \mathbb{I}'$ is not visited in cycle l, and $y_{il} = 1$ otherwise. Note that y_{il} is known once the multipliers m_i, with $i \in \mathbb{I}'$, is specified.

The excess waste over the truck capacity in the lth collection cycle is $\left(\sum_{i\in\mathbb{I}'} y_{il}Z_i - Q\right)^+$. To evaluate its expectation, we need to determine the mean and the variance of $\sum_{i\in\mathbb{I}'} y_{il}Z_i$, which is a Gaussian random variable, being a linear combination of Gaussian random variables. The mean of $\sum_{i\in\mathbb{I}'} y_{il}Z_i$, which is related to the lth collection cycle, is obtained as

$$\mu_{\Sigma l} = E\left[\sum_{i\in\mathbb{I}'} y_{il}Z_i\right] = \sum_{i\in\mathbb{I}'} y_{il}E[Z_i] = \sum_{i\in\mathbb{I}'} y_{il}\mu_i = T\sum_{i\in\mathbb{I}'} y_{il}m_iD_i,$$

while its variance is (please refer to Appendix A in the online supplement for its full derivation)

$$\varsigma_{\Sigma l}^2 = T\sum_{i\in\mathbb{I}'} y_{il}^2 m_i\Sigma_i^2 + T^2 \sum_{i_1, i_2\in\mathbb{I}', i_1\neq i_2} \left(y_{i_1 l}y_{i_2 l}m_{i_1}m_{i_2} \sum_{j_1, j_2\in\mathbb{I}} z_{j_1 i_1}z_{j_2 i_2}\text{Cov}\left(X_{j_1}, X_{j_2}\right)\right).$$

The expectation of $\left(\sum_{i\in\mathbb{I}'} y_{il}Z_i - Q\right)^+$ is thus given by

$$E\left[\left(\sum_{i\in\mathbb{I}'} y_{il}Z_i - Q\right)^+\right] = \varsigma_{\Sigma l}[f(p_{\Sigma l}) + p_{\Sigma l}(F(p_{\Sigma l}) - 1)] = \varsigma_{\Sigma l}\psi^+(p_{\Sigma l}),$$

where $p_{\Sigma l} = (Q - \mu_{\Sigma l})/\varsigma_{\Sigma l}$, being $\varsigma_{\Sigma l} = \sqrt{\varsigma_{\Sigma l}^2}$. Since the expected penalty cost due to exceeding the truck capacity in the lth collection cycle is $\theta\varsigma_{\Sigma l}\psi^+(p_{\Sigma l})$, then the expected penalty cost due to exceeding the truck capacity over the M collection cycles occurring in the interval between two regeneration instances is $\theta\sum_{l=1}^{M} \varsigma_{\Sigma l}\psi^+(p_{\Sigma l})$.

We finally recall that the major waste collection ordering cost, A, is paid every T time units–that is, every collection period–and that the minor waste collection ordering cost, a_{ik}, is paid every time bin $i \in \mathbb{I}'$ is visited–that is, every m_iT time units. In conclusion, the component of the long-run expected total system cost rate that is dependent on the waste collection frequency is.

$$\begin{aligned} \mathfrak{K}_1 &= \mathfrak{K}_1(T, \mathbf{m}, \mathbf{X}, \mathbf{Z}) = \frac{A}{T} + \frac{\theta}{MT}\sum_{l=1}^{M} \varsigma_{\Sigma l}\psi^+(p_{\Sigma l}) \\ &+ \sum_{i\in\mathbb{I}'} \left(\frac{\sum_{k\in\mathbb{K}} x_{ik}a_{ik}}{m_iT} + h\frac{m_iTD_i}{2} + \pi_{0i}\frac{\varsigma_i\psi^+(p_i)}{m_iT} + \pi_{1i}\frac{\varsigma_i\psi^-(p_i)}{m_iT} \right), \end{aligned} \tag{1}$$

where $\mathbf{m} = (m_i)_{i\in\mathbb{I}'}$, $\mathbf{X} = (x_{ik})_{i\in\mathbb{I}', k\in\mathbb{K}}$, and $\mathbf{Z} = (z_{ij})_{i,j\in\mathbb{I}}$. Note that $\mathfrak{K}_1$ is also a function of $\mathbf{X}$ and $\mathbf{Z}$, the matrices of decision variables associated with the bin location-allocation problem. We also recall that $i \in \mathbb{I}' \Leftrightarrow \sum_{k\in\mathbb{K}} x_{ik} > 0$. Hence, all terms in $\mathfrak{K}_1$, except the former, are defined only for nodes at which a bin is positioned.

We now develop the remaining part of the long-run expected total system cost rate, which includes the costs concerning the location of bins and the allocation of waste to bins. The cost rate to position a bin of type k in node i is c_{2ik}. Hence, the total cost rate due to bin positioning is $\sum_{i\in\mathbb{I}}\sum_{k\in\mathbb{K}} c_{2ik}x_{ik}$.

Since the cost per unit distance per unit quantity of waste delivered from node $i \in \mathbb{I}$ to node $j \in \mathbb{I}$ is c_{1ij}, the long-run expected cost rate due to waste delivery from node $i \in \mathbb{I}$ is $d_i\sum_{j\in\mathbb{I}} c_{1ij}r_{ij}z_{ij}$, being r_{ij} the distance between nodes i and j. As a result, the long-run expected total cost rate related to waste delivery is $\sum_{i\in\mathbb{I}} d_i\sum_{j\in\mathbb{I}} c_{1ij}r_{ij}z_{ij}$. In conclusion, the long-run expected cost rate concerning the location and selection of bins and the allocation of waste to bins is.

$$\mathfrak{K}_2 = \mathfrak{K}_2(\mathbf{X}, \mathbf{Z}) = \sum_{i\in\mathbb{I}}\sum_{k\in\mathbb{K}} c_{2ik}x_{ik} + \sum_{i\in\mathbb{I}} d_i\sum_{j\in\mathbb{I}} c_{1ij}r_{ij}z_{ij}.$$

We finally get the long-run expected total system cost rate, $\mathfrak{K}$, as the sum of $\mathfrak{K}_1$ and $\mathfrak{K}_2$:

$$\begin{aligned} \mathfrak{K} &= \mathfrak{K}(T, \mathbf{m}, \mathbf{X}, \mathbf{Z}) = \sum_{i\in\mathbb{I}}\sum_{k\in\mathbb{K}} c_{2ik}x_{ik} + \sum_{i\in\mathbb{I}} d_i\sum_{j\in\mathbb{I}} c_{1ij}r_{ij}z_{ij} \\ &+ \frac{A}{T} + \frac{\theta}{MT}\sum_{l=1}^{M} \varsigma_{\Sigma l}\psi^+(p_{\Sigma l}) \\ &+ \sum_{i\in\mathbb{I}'} \left(\frac{\sum_{k\in\mathbb{K}} x_{ik}a_{ik}}{m_iT} + h\frac{m_iTD_i}{2} + \pi_{0i}\frac{\varsigma_i\psi^+(p_i)}{m_iT} + \pi_{1i}\frac{\varsigma_i\psi^-(p_i)}{m_iT} \right). \end{aligned}$$

The objective is to minimise the long-run expected total system cost rate, $\mathfrak{K}$, by optimising the decision variables related to bin location and type selection ($\mathbf{X}$), waste allocation ($\mathbf{Z}$), and collection frequency (T, $\mathbf{m}$):

$$\min_{(T, \mathbf{m}, \mathbf{X}, \mathbf{Z})} \mathfrak{K}.$$

This optimisation is subject to a set of constraints:

$$\sum_{j\in\mathbb{I}} z_{ij} = 1\ \forall i \in \mathbb{I}, \tag{2}$$

$$\sum_{k\in\mathbb{K}} x_{ik} - z_{ji} \geq 0\ \forall i, j \in \mathbb{I}, \tag{3}$$

$$\sum_{k\in\mathbb{K}} x_{ik} - \sum_{j\in\mathbb{I}} z_{ji} \leq 0\ \forall i \in \mathbb{I}, \tag{4}$$

$$\sum_{k\in\mathbb{K}} x_{ik} \leq 1\ \forall i \in \mathbb{I}, \tag{5}$$

$$z_{ij} \in \{0, 1\}\ \forall i, j \in \mathbb{I}, \tag{6}$$

$$x_{ik} \in \{0, 1\}\ \forall i \in \mathbb{I}, k \in \mathbb{K}, \tag{7}$$

$$m_i \in \mathbb{N}\ \forall i \in \mathbb{I}', \tag{8}$$

$$T \in \mathbb{R}_+^*. \tag{9}$$

Constraint (2) ensures a complete and unique assignment for every waste source. In practical terms, it dictates that each waste-generating node i must deliver its waste to exactly one collection bin at a location

j. Constraint (3) creates a logical link between bin placement and waste allocation. It stipulates that waste from any source j can only be assigned to a location i if a bin is actually present at i. Constraint (4), working in tandem with constraint (3), prevents the placement of unused "orphan" bins. It ensures that if a bin is installed at a location i, it must be allocated at least one waste source j. Constraint (5) ensures that each potential site i can host at most one type of bin. This reflects the physical limitation that a single collection point cannot simultaneously be, for example, a small bin and a large compactor. Constraint (6) defines the waste allocation variable z_{ij} as binary. A value of 1 signifies that waste source i is fully assigned to the bin at location j, while 0 signifies no assignment. Partial allocations are not permitted. Constraint (7) defines the bin placement and type selection variable x_{ik} as binary. A value of 1 indicates that a bin of type k is installed at location i; 0 indicates it is not. Constraint (8) establishes the structured, periodic nature of the collection schedule, a core feature of the JRP-inspired framework. It dictates that the collection interval for any given bin i must be a positive integer multiple (e.g., 1, 2, 3, …) of the base period T. Finally, constraint (9) ensures that the basic collection period T (e.g., measured in weeks or days) is a positive, real-valued number, as a zero or negative time interval would be physically meaningless.

5. Solution method

The optimisation of problem (P) under constraints (2)–(9) constitutes a MINLP. Since this formulation generalises the deterministic JRP, which is known to be NP-hard (Arkin et al., 1989), exact solution methods are computationally intractable for realistic problem sizes. Moreover, its non-convex structure poses significant challenges for commercial optimisation solvers. As we will demonstrate empirically in Section 6.1, even state-of-the-art exact methods fail to solve instances of a realistic size within a reasonable time frame.

Consequently, we propose a hybrid solution methodology that decomposes the problem into two hierarchical levels: an outer combinatorial search for discrete location, bin type selection, and allocation decisions, and an inner heuristic procedure to optimise continuous collection frequency parameters. It is important to note that the primary focus of this research is the integrated stochastic problem formulation itself. The solution procedure described herein is presented as a practical and effective mechanism to solve this complex model and extract managerial insights, rather than as a novel algorithmic contribution in its own right.

5.1. Overview of the hybrid solution method

The proposed solution framework relies on a master-subproblem decomposition. The master problem involves determining the optimal configuration of binary decision variables: bin location and type selection ($\mathbf{X}$) and waste allocation ($\mathbf{Z}$). Due to the combinatorial nature of these decisions, an iterative metaheuristic search strategy is employed as the "outer loop.".

For any candidate configuration $(\mathbf{X}, \mathbf{Z})$ generated by the metaheuristic, the problem reduces to finding the optimal collection frequency parameters–specifically the basic collection period T and the integer multipliers $\mathbf{m}$–that minimise the system cost. This constitutes the subproblem or "inner loop." The minimal cost obtained from this subproblem serves as the fitness value for the candidate solution, guiding the metaheuristic's subsequent search steps.

Fig. 2 illustrates the schematic workflow of this hybrid approach. The outer loop can be driven by any suitable metaheuristic algorithm capable of handling combinatorial constraints.

5.2. Heuristic algorithm for collection frequency optimisation

This section describes the algorithm employed in the inner loop to address the subproblem of minimising the long-run expected total sys-

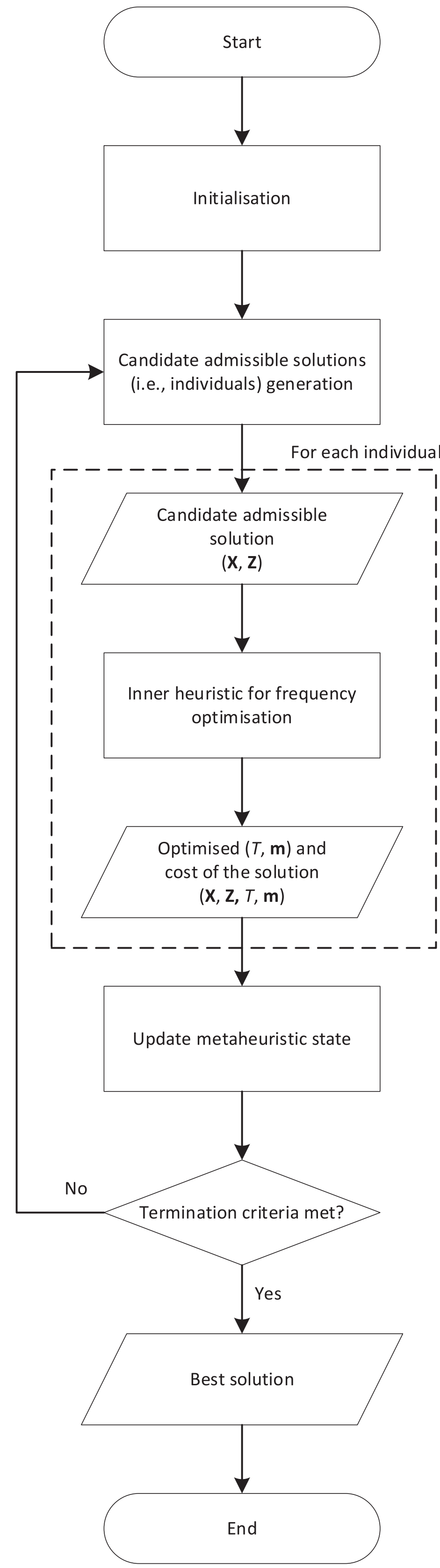

Fig. 2. Schematic of the hybrid solution method. The dashed line demarcates the "inner loop" algorithm.

tem cost rate $\mathfrak{R}$ with respect to T and $\mathbf{m}$, given a fixed admissible solution $(\mathbf{X},\mathbf{Z})$. For a fixed $(\mathbf{X},\mathbf{Z})$, the component of $\mathfrak{R}$ that depends on the collection frequency parameters $(T,\mathbf{m})$ is $\mathfrak{R}_1$ (see Eq. (1)). Function $\mathfrak{R}_1$ shares a structural similarity with the cost function used in the stochastic JRP–that is, the JRP with stochastic demand (see, e.g., Braglia et al., 2021). However, it is more complex due to the inclusion of additional cost components. The stochastic JRP is typically tackled using heuristic algorithms, as identifying the global optimum is known to be challenging. A similar situation arises in our case; therefore, we propose a heuristic procedure inspired by the improved version of Silver's algorithm developed by Kaspi & Rosenblatt (1983) for the deterministic JRP.

Algorithm 1 formalises the heuristic procedure to find the minimum of $\mathfrak{R}_1$ in $(\mathrm{T},\mathbf{m})$, for a given admissible solution in $(\mathbf{X},\mathbf{Z})$. Its derivation is presented in Appendix B in the online supplement.

Algorithm 1. Procedure to find the heuristic solution $(T^*,\mathbf{m}^*)$ to the problem of minimising $\mathfrak{R}_1$ in $(T,\mathbf{m})$, for a given admissible solution in $(\mathbf{X},\mathbf{Z})$, under constraints (8) and (9)..

initialise: *MaxIter_1*, *MaxIter_2*
set $\mathfrak{R}_1^* = +\infty$, $Counter\_1 = 1$, $ContrVar = 0$, $g = 1$, *and* $m_i = 1$ *and* $\widehat{m}_i^{(g)} = 0$ *for each* $i \in \mathbb{I}'$
calculate $\overline{T}_i$ *by minimising Eq. (A.1) in* m_iT *for each* $i \in \mathbb{I}'$
let $i^* = \operatorname{argmin}_{i\in\mathbb{I}'}\overline{T}_i$ *and rearrange indices so that* $i^* = 1$
set $m_1 = 1$
WHILE ($Counter\_1 \le MaxIter\_1$ *AND* $ContrVar == 0$) *DO*
_*set* $Counter\_2 = 1$
_*WHILE* ($Counter\_2 \le MaxIter\_2$ *AND* $m_i \ne \widehat{m}_i^{(g)}$ *for each* $i \in \mathbb{I}'$) *DO*
__*set* $\widehat{m}_i^{(g)} \leftarrow m_i$ *for each* $i \in \mathbb{I}'$
__*calculate* $\widehat{T}$ *by solving Eq. (A.5) in* T
__*FOR* (*each* $i \in \mathbb{I}'\backslash\{1\}$) *DO*
___*set* $q_i \leftarrow \max\{1, \lfloor \overline{T}_i/\widehat{T} \rfloor\}$
___*IF* ($\mathfrak{R}_{1i}(\widehat{T},q_i) \le \mathfrak{R}_{1i}(\widehat{T},q_i+1)$) *THEN*
____*set* $m_i \leftarrow q_i$
___*ELSE*
____*set* $m_i \leftarrow q_i + 1$
___*END IF*
__*END FOR*
__*set* $Counter\_2 \leftarrow Counter\_2 + 1$ *and* $g \leftarrow g + 1$
_*END WHILE*
_*calculate* $\overline{T}$ *by solving Eq. (A.3) in* T
_*IF* ($\mathfrak{R}_1(\overline{T},\mathbf{m}) \le \mathfrak{R}_1^*$) *THEN*
__*set* $T^* \leftarrow \overline{T}$ *and* $\mathbf{m}^* \leftarrow \mathbf{m}$
__*set* $m_1 \leftarrow m_1 + 1$
__*set* $\mathfrak{R}_1^* \leftarrow \mathfrak{R}_1(\overline{T},\mathbf{m})$
__*set* $Counter\_1 \leftarrow Counter\_1 + 1$
_*ELSE*
__*set* $ContrVar = 1$
_*END IF*
END WHILE

Since the metaheuristic algorithm, i.e., the outer loop, tackles the combinatorial problem over $(\mathbf{X},\mathbf{Z})$ and, for each admissible solution, it is necessary to run Algorithm 1 in order to get the corresponding cost, then it may be of interest to calculate the maximum number of times Algorithm 1 needs to run for a given pair (N,K). In other words, it may be useful determining the number of admissible solutions in $(\mathbf{X},\mathbf{Z})$. This calculation is presented in Appendix C in the online supplement.

6. Numerical experiments

This section presents a comprehensive set of numerical experiments designed to assess the proposed model, demonstrate its computational tractability, and derive actionable practical insights. The analysis is structured into five main parts. First, in Section 6.1, we provide a theoretical and empirical justification for our heuristic methodology by analysing the problem's computational complexity and demonstrating the limitations of exact commercial solvers. Second, in Section 6.2, we assess the efficiency and effectiveness of the proposed general solution framework. To this end, we evaluate its performance when exploiting two distinct standard metaheuristics (a genetic algorithm and a multistart method) to navigate the combinatorial search space. Third, in Section 6.3, we quantify the inherent benefit of our modelling approach by calculating the value of the stochastic solution (VSS), demonstrating the economic superiority of explicitly accounting for uncertainty versus relying on deterministic averages. Fourth, in Section 6.4, we conduct extensive scenario and sensitivity analyses to explore the system's behaviour under varying operational conditions and parameter fluctuations. Finally, in Section 6.5, we synthesise these numerical findings into strategic managerial implications and concrete policy recommendations for waste management planning.

All experiments were conducted on a machine equipped with an Intel® Core™ i9-14900HX processor (2.20 GHz) and 64 GB of RAM. The algorithms were implemented in MATLAB® R2024b, leveraging its parallel computing toolbox to utilise up to 24 cores.

6.1. Analysis of problem complexity and justification for the heuristic approach

The computational complexity of the problem stems primarily from the vast, non-convex search space of the discrete decision matrices $(\mathbf{X},\mathbf{Z})$. As derived in Appendix C in the online supplement, the number of admissible solutions grows at a super-exponential rate. For instance, for a moderately-sized problem with $N = 10$ nodes and $K = 3$ bin types, the number of feasible combinations of bin locations, type selection, and allocations exceeds 10^{10}. For the largest instances tested presented in Section 5.2 ($N = 40$, $K = 3$), this number becomes astronomically large, rendering any form of exhaustive enumeration computationally impossible. This combinatorial explosion is a key driver of the problem's NP-hard nature.

Before presenting the main computational results of our proposed algorithm, we first establish the problem's computational complexity to provide a clear justification for adopting a heuristic methodology. To this end, we benchmarked a necessarily adapted version of the problem using the state-of-the-art commercial solver Gurobi.

To render the problem tractable for an exact solver, several adaptations to the formulation were required, creating a non-convex mixed-integer quadratic programme (MIQP). These adaptations included:

1. relaxing the discontinuous truck capacity term, which relies on the non-linear LCM function, by assuming infinite truck capacity;
2. discretising the continuous base period T into a finite set of 50 potential values; and
3. determining the integer multipliers through a brute-force enumeration over a pre-defined range (e.g., [1, 20]) using a large set of binary variables.

This formulation choice, while necessary for the MIQP structure, contrasts with the efficient, iterative search performed by our heuristic (Algorithm 1) and significantly increases the model's combinatorial complexity.

Our analysis of this adapted model yielded a two-fold justification for a heuristic approach:

1. Scalability failure: the exact MIQP approach proved to be computationally intractable for realistically-sized problems. For a small instance with $N = 3$ nodes and $K = 2$ bin types (instance ID 3.2.1 included into the online dataset–see Bertolini et al., 2025), Gurobi found a proven optimal solution. However, when applied to a moderately-sized instance with $N = 10$ nodes and $K = 3$ bin types (instance ID 10.3.1 included into the online dataset –see Bertolini et al., 2025), the solver failed to prove optimality within a one-hour (3,600 s) time limit, terminating with a large optimality gap (92.5%). Given that the full, continuous problem is strictly harder than this

adapted formulation, this result confirms that exact methods are unsuitable for practical application.

2. Solution quality degradation via discretisation: the necessary adaptations required by the exact solver can lead to suboptimal solutions. For the small $N = 3$ instance, Gurobi identified a global optimum for its discretised model with a cost of \$54,379.06/year. In contrast, our proposed procedure, which optimises the full continuous model, discovered a superior configuration with a cost of \$39,769.51/year–a 26% improvement. This demonstrates that the necessary discretisation of the decision space can prevent the solver from identifying the true continuous optimum, which a specialised heuristic is designed to explore.

Taken together, these findings show that an exact solver is both too slow for large instances and potentially suboptimal for small ones due to its structural and formulation requirements. This provides a compelling, data-driven mandate for the specialised hybrid heuristic approach proposed in this paper.

6.2. Analysis of the solution procedure

As already observed, the search over $(\mathbf{X}, \mathbf{Z})$ is a combinatorial optimisation problem, in which the cost of a given admissible pair $(\mathbf{X}, \mathbf{Z})$ is evaluated through Algorithm 1. In these experiments, both parameters MaxIter_1 and MaxIter_2 in Algorithm 1 are equal to 30.

We adopt two metaheuristic algorithms to approach the outer optimisation loop discussed in Section 5.1, i.e., the combinatorial search over $(\mathbf{X}, \mathbf{Z})$. The first one is a standard genetic algorithm (GA), and the second one is a specifically developed memoryless, randomised multi-start method (MSM). We have chosen these algorithms as both are known to be effective for combinatorial optimisation (Martí et al., 2013; Sivanandam & Deepa, 2008). While the MSM is detailed in Appendix D, a full description of our GA implementation—including chromosome encoding, fitness evaluation, and operator selection—is provided in Appendix E. The parameters of GA and MSM were optimised with preliminary experiments, trying to seek an acceptable balance between computation time and solution quality. The optimised parameter values of GA and MSM are shown in Table 2.

We investigate the performance of both algorithms considering problems with varying dimension–that is, for different (N, K) pairs. For each (N, K) pair, we generated 5 random instances, drawing parameter values from the intervals shown in Table 3. To ensure a robust and statistically sound evaluation, we performed 30 independent runs for each instance with both the GA and MSM. The complete results are presented in Table 4. This table details, for each instance, the optimal solution from an exhaustive search–for instance sizes up to $N = 5$ and $K = 3$ due to computation limits of our platform–and provides a full statistical summary for both metaheuristics, including the best-found cost, the mean cost, the standard deviation of the cost, and the mean computation time with its standard deviation over the 30 runs. Fig. 3 shows how the average computation time varies with the number of entries in $(\mathbf{X}, \mathbf{Z})$ for GA and MSM in the instances we examined. The online dataset accompanying this paper and available on Mendeley Data (Bertolini et al., 2025) provides all problem instances utilised in this analysis.

Table 2
Optimised parameters of metaheuristic solvers.

Algorithm	Parameter	Value
GA	Population size	$15 \times N$
	Maximum generations number	$10 \times N$
	Maximum stall generations	$\lfloor 0.3 \times 10 \times N \rfloor$
	Elite count	$\lfloor 0.1 \times 15 \times N \rfloor$
	Crossover fraction	0.8
	Mutation probability	0.1
	Crossover function	Two-point
	Mutation function	Bit-flip
	Tournament size	$\lfloor 0.15 \times 15 \times N \rfloor$
MSM	Number of starts	$5 \times N$
	Maximum iterations	$8 \times N$

Even for problems with a small number of variables, the computation time of the exhaustive search is substantial, reaching values in the order of 10^4 to 105 s. This exponential increase in computational effort clearly demonstrates the impracticality of exhaustive search as the problem dimensionality increases, particularly when scalability is a critical factor.

In contrast, the GA shows a significantly improved performance profile in terms of both solution quality and computational efficiency. When benchmarked against the exhaustive search for smaller problem sizes, the GA is able to find the optimal solution in most cases while reducing the computational burden dramatically. As shown in Table 4, the GA demonstrates remarkable consistency and algorithmic stability. Exhibiting mean solution costs that are very close to the best-found values and a consistently low standard deviation (i.e., minimal variance across the 30 independent executions), this indicates that the GA reliably converges to high-quality solutions regardless of the initial random population. Furthermore, its computation times are not only dramatically lower than the exhaustive search but also exhibit low variability, making it a predictable and practical alternative for larger and more complex optimisation problems.

The MSM also offers a clear reduction in computation time compared to the exhaustive search, but its performance in terms of solution quality is more variable. In some instances, it achieves comparable speeds to the GA; however, its average performance is demonstrably weaker. Table 4 highlights this inconsistency and relative lack of stability; the MSM exhibits significantly higher mean costs and much larger standard deviations (high variance) for solution quality compared to the GA, particularly for larger instances. This points to a tendency for the multi-start approach to become trapped in suboptimal local minima. Such behaviour suggests that while the MSM may be advantageous in scenarios where rapid, approximate solutions are acceptable, its reliability in achieving near-optimal solutions diminishes with the increasing complexity of the problem.

To formally assess the performance difference between the two metaheuristics, we performed a Wilcoxon signed-rank test, a non-parametric test suitable for comparing paired samples. For each instance, the test compared the 30 solution costs obtained by the GA against the 30 costs from the MSM. The null hypothesis is that there is no median difference between the paired results. Table 5 presents the resulting p-values for each instance. A p-value below the significance level of $\alpha = 0.05$ indicates that the observed performance difference is statistically significant.

The results of the statistical tests are unambiguous. For the smallest problem set ($N = 3$, $K = 2$), where both algorithms always find the optimal solution, the performance difference is not statistically significant. However, as the problem size increases to $N = 4$, the GA's advantage becomes statistically significant in many instances. For all problems with $N \geq 5$, the p-values are exceptionally low ($p < 0.001$), providing overwhelming evidence to reject the null hypothesis. This rigorous, instance-level analysis confirms that the GA is not only more consistent but also provides solutions of a statistically superior quality for any problem of non-trivial size. Given these results and acknowledging that the GA is a well-established and standard metaheuristic, we conclude that it serves as a highly robust and appropriate search engine for our proposed framework. The GA's success demonstrates that complex, integrated stochastic MSW models can be effectively solved using available, high-performing metaheuristic paradigms combined with problem-specific heuristics (like Algorithm 1).

Taken together, these results underscore the trade-offs between computational efficiency and solution accuracy. For small-scale problems where computational resources are not a limiting factor, exhaustive search provides the benchmark for optimality. However, for larger

Table 3
Intervals in which parameter values are chosen randomly.

Parameter	Interval	Unit of measurement	Parameter	Interval	Unit of measurement
A	[90, 180]	\$/collection order	c_{1ij}	[0.5, 5.0]	\$/km/unit
h	[5, 30]	\$/unit/year	c_{2ik}	[20, 80]	\$/year
θ	[80, 160]	\$/unit	π_{0i}	[80, 160]	\$/unit
Q	[100, 500]	unit	π_{1i}	[80, 160]	\$/unit
C_1	[100, 300]	unit	r_{ij}	[1, 10]	km
C_k, for $k = 2,3,\cdots,K$	$[C_{k-1}, \lfloor 1.5 \times C_{k-1} \rfloor]$	unit	σ_i^2	[1, 400]	(unit/year)^2
a_{ik}	[30, 80]	\$/collection order	σ_{ij}, with $i \neq j$	[-400, 400]	(unit/year)^2
d_i	[3000, 6000]	unit/year			

problem instances, the GA emerges as the demonstrably superior method, offering a robust balance of speed, accuracy, and reliability that is critical in practical applications. The MSM, while relatively fast in some cases, lacks the consistency required for such a complex optimisation problem.

6.3. Value of the stochastic solution

A central premise of this paper is that explicitly modelling the stochastic nature of waste generation leads to more robust and cost-effective system designs. To empirically validate this premise and quantify the tangible benefit of our framework, we conducted an analysis to determine the VSS. The VSS represents the cost savings achieved by the stochastic model compared to a traditional deterministic model that operates solely on average values. To ensure the robustness of our findings and investigate how the VSS scales with problem size, we conducted this analysis across three representative small-scale instances of increasing complexity: instance ID 3.2.1 ($N = 3$, $K = 2$), instance ID 4.3.3 ($N = 4$, $K = 3$), and instance ID 5.3.3 ($N = 5$, $K = 3$).

The experiment was structured to ensure a rigorous comparison. First, for each instance, we defined and solved a deterministic version of the optimisation problem. The deterministic cost structure is formally equivalent to the limit of our stochastic cost function $\mathfrak{K}$ as all variance (σ_i^2) and covariance (σ_{ij}) parameters approach zero, effectively replacing stochastic penalties with deterministic ones that apply only when the mean accumulated waste exceeds capacity. To obtain the true optimal solutions for this analysis and eliminate any performance artifacts from metaheuristic randomness, both the deterministic and stochastic models for these three instances were solved using the exhaustive search method.

Next, to simulate the real-world performance of each "uncertainty-unaware" design, we evaluated its long-run expected total system cost rate using our original, full stochastic cost function. This cost reflects the financial consequences of implementing a plan designed for a world of averages. Finally, this cost was compared against the expected cost of the optimal stochastic solution.

The results of this comparative analysis are presented in Table 6. The findings consistently demonstrate a substantial VSS across all tested instances, confirming the practical value of the stochastic approach. Furthermore, the results strongly support the reasonable hypothesis that the benefit of explicitly modelling uncertainty grows with problem complexity. As the number of entries in $(\mathbf{X}, \mathbf{Z})$ increases, the relative VSS rises from 67.7% to 97.2%. This upward trend is logical: larger, more complex systems present more opportunities for stochastic variations to compound, amplifying the risks of simultaneous bin overflows and truck overcapacity across the network. A deterministic design, which optimises for a single average-case scenario, becomes progressively more brittle and prone to systemic failure as the number of interacting components grows. The stochastic model, by contrast, strategically incorporates capacity buffers and adjusts collection frequencies to hedge against these amplified risks, with the economic benefit of this foresight becoming more pronounced in larger systems. This analysis provides compelling evidence that our stochastic framework is not merely a refinement but a critical tool for designing cost-effective waste management systems, with its value increasing directly with the scale of the operational challenge.

6.4. Scenario and sensitivity analyses

This section focuses on examining the model's response to various scenarios and alterations in parameter values, aiming to derive valuable managerial insights. Experiments are made considering the case with $N = 3$ and $K = 2$, as this condition permits us (*i*) to exploit the exhaustive search and reach the optimum solution over $(\mathbf{X}, \mathbf{Z})$ in a relatively reasonable time, and (*ii*) to more easily uncover any patterns existing in results. Clearly, the implicit assumption is that configurations with a greater number of nodes behave similarly to smaller ones.

Experiments are organised as follows. We consider four scenarios in terms of covariance matrix $(\sigma_{ij})_{i,j \in \mathbb{I}}$:

1. Low correlation scenario: the covariance matrix is nearly diagonal. This scenario assumes that variations in the waste generation rates at each node occur almost independently.
2. High correlation scenario: all pairs of nodes have high covariance. In this scenario, increases (or decreases) in generation rates tend to occur simultaneously across all nodes, increasing the overall risk of aggregated peaks.
3. Block correlation scenario: the three nodes are grouped into two clusters (e.g., two nodes in a densely populated urban area and one in a more peripheral area) and assume a high correlation within each cluster and a low correlation between clusters. This structure reflects a situation where local factors (such as weather or local events) strongly influence some nodes but not necessarily the entire area.
4. Mixed correlation scenario: some pairs of nodes exhibit a strong negative correlation (an increase in one area leads to a decrease in another due to compensatory mechanisms, potentially influenced by shifting population or economic factors), while others a strong positive correlation.

Regarding the other parameters, we consider a baseline condition as a reference. In each of the four scenarios, we analyse the impact of deviations from this baseline on the solution and various performance metrics by changing the value of one parameter at a time by $\pm$ 80%, keeping the value of the others fixed. Moreover, in each scenario, the comparison with the baseline condition is made considering two cases: (*i*) keeping the baseline solution in $(\mathbf{X}, \mathbf{Z})$ fixed, and (*ii*) allowing the solution in $(\mathbf{X}, \mathbf{Z})$ to adapt to the new parameter setting.

In the online supplement gives the parameter values that define the baseline condition. Along with the solution, the performance metrics we consider are the total system cost rate ($\mathfrak{K}$), the total bin location and waste allocation cost rate ($\mathfrak{K}_2$), the cost rate for exceeding truck capacity ($\mathfrak{K}_{\mathrm{tc}}$), the total cost rate for exceeding bin capacity ($\mathfrak{K}_{\mathrm{bc}}$), and the total cost rate for not fully exploiting bin capacity ($\mathfrak{K}_{\mathrm{nbc}}$)

Sections 6.4.1–6.4.4 report the results for each scenario. The solution

Table 4
Results of the comparison.

	# of entries in (X, Z)	Instance ID	Exhaustive		GA				MSM			
			Min. cost [$/year]	Time [s]	Best cost [$/year]	Mean cost [$/year]	St. dev. [$/year]	Time [s] Mean ± St. dev.	Best cost [$/year]	Mean cost [$/year]	St. dev. [$/year]	Time [s] Mean ± St. dev.
$N = 3$, $K = 2$	15	3.2.1	9.91×10^4	16.0	9.91×10^4	9.93×10^4	4.10×10^2	16.2 ± 0.4	9.91×10^4	1.00×10^5	1.53×10^3	16.5 ± 0.7
		3.2.2	4.63×10^4	44.9	4.63×10^4	4.64×10^4	3.02×10^2	20.1 ± 0.5	4.63×10^4	4.69×10^4	1.14×10^3	17.0 ± 0.8
		3.2.3	1.14×10^5	22.5	1.14×10^5	1.14×10^5	3.21×10^3	17.9 ± 0.4	1.14×10^5	1.16×10^5	2.15×10^4	26.5 ± 1.1
		3.2.4	8.39×10^4	10.5	8.39×10^4	8.41×10^4	4.14×10^2	13.6 ± 0.3	8.39×10^4	8.45×10^4	1.26×10^3	9.1 ± 0.5
		3.2.5	5.76×10^4	19.7	5.76×10^4	5.77×10^4	3.16×10^2	18.2 ± 0.5	5.76×10^4	5.83×10^4	1.42×10^3	38.6 ± 1.4
$N = 4$, $K = 2$	24	4.2.1	7.58×10^4	450.7	7.58×10^4	7.62×10^4	8.06×10^2	74.5 ± 1.8	7.58×10^4	7.64×10^4	2.50×10^3	37.1 ± 2.1
		4.2.2	7.95×10^4	437.0	7.95×10^4	8.01×10^4	9.54×10^2	42.7 ± 1.1	8.13×10^4	8.35×10^4	3.83×10^3	37.4 ± 2.3
		4.2.3	1.06×10^5	279.7	1.06×10^5	1.07×10^5	1.04×10^4	42.3 ± 1.0	1.06×10^5	1.09×10^5	4.11×10^4	29.7 ± 1.9
		4.2.4	1.16×10^5	251.4	1.16×10^5	1.17×10^5	1.11×10^4	46.9 ± 1.2	1.16×10^5	1.19×10^5	4.42×10^4	54.8 ± 2.8
		4.2.5	1.28×10^5	366.0	1.28×10^5	1.29×10^5	1.15×10^4	38.6 ± 0.9	1.30×10^5	1.34×10^5	4.51×10^4	70.9 ± 3.5
$N = 4$, $K = 3$	28	4.3.1	7.85×10^4	1.31×10^3	7.85×10^4	7.91×10^4	1.23×10^3	104.3 ± 2.5	7.85×10^4	7.94×10^4	4.90×10^3	95.2 ± 4.1
		4.3.2	6.71×10^4	718.6	7.13×10^4	7.21×10^4	1.48×10^3	27.2 ± 0.8	7.30×10^4	7.66×10^4	5.12×10^3	82.1 ± 3.9
		4.3.3	6.05×10^4	801.0	6.05×10^4	6.10×10^4	1.18×10^3	48.5 ± 1.3	6.05×10^4	6.12×10^4	4.81×10^3	91.3 ± 4.4
		4.3.4	6.23×10^4	1.12×10^3	6.23×10^4	6.28×10^4	8.89×10^2	60.7 ± 1.5	6.44×10^4	6.79×10^4	4.95×10^3	94.6 ± 4.7
		4.3.5	7.60×10^4	850.7	7.60×10^4	7.65×10^4	1.26×10^3	55.9 ± 1.4	7.60×10^4	7.67×10^4	5.06×10^3	86.3 ± 4.2
$N = 5$, $K = 3$	40	5.3.1	6.35×10^4	2.28×10^4	6.69×10^4	6.78×10^4	2.17×10^3	55.8 ± 1.8	7.00×10^4	7.41×10^4	6.50×10^3	155.9 ± 7.1
		5.3.2	6.33×10^4	1.99×10^4	6.86×10^4	6.95×10^4	2.29×10^3	44.9 ± 1.5	7.14×10^4	7.63×10^4	7.17×10^3	141.8 ± 6.8
		5.3.3	7.99×10^4	2.41×10^4	7.99×10^4	8.06×10^4	1.94×10^3	84.7 ± 2.5	8.42×10^4	8.81×10^4	7.28×10^3	107.8 ± 5.5
		5.3.4	1.04×10^5	2.17×10^4	1.14×10^5	1.16×10^5	2.52×10^4	53.0 ± 1.7	1.18×10^5	1.24×10^5	8.15×10^4	166.1 ± 8.0
		5.3.5	6.02×10^4	2.31×10^4	6.48×10^4	6.57×10^4	2.36×10^3	78.1 ± 2.4	6.40×10^4	6.75×10^4	6.29×10^3	94.2 ± 4.9
$N = 10$, $K = 3$	130	10.3.1	–	–	4.93×10^5	4.99×10^5	3.94×10^4	495.2 ± 15.4	5.45×10^5	5.92×10^5	2.01×10^5	1035.0 ± 45.4
		10.3.2	–	–	3.24×10^5	3.28×10^5	3.58×10^4	325.1 ± 11.5	4.11×10^5	4.53×10^5	1.95×10^5	881.1 ± 39.3
		10.3.3	–	–	2.34×10^5	2.37×10^5	2.87×10^4	235.5 ± 9.2	2.50×10^5	2.76×10^5	1.45×10^5	1105.5 ± 51.7
		10.3.4	–	–	2.72×10^5	2.76×10^5	3.15×10^4	273.8 ± 10.1	3.05×10^5	3.34×10^5	1.76×10^5	720.4 ± 35.1
		10.3.5	–	–	3.08×10^5	3.12×10^5	3.54×10^4	310.4 ± 12.7	3.30×10^5	3.65×10^5	1.95×10^5	1143.8 ± 55.6
$N = 20$, $K = 3$	460	20.3.1	–	–	1.57×10^6	1.60×10^6	2.61×10^5	1575.6 ± 58.3	1.90×10^6	2.15×10^6	1.01×10^6	1545.1 ± 81.9
		20.3.2	–	–	1.68×10^6	1.71×10^6	2.76×10^5	1683.5 ± 65.0	1.81×10^6	2.05×10^6	9.60×10^5	1894.7 ± 95.8
		20.3.3	–	–	1.47×10^6	1.51×10^6	2.82×10^5	1472.3 ± 55.6	1.85×10^6	2.12×10^6	1.03×10^6	1915.5 ± 102.0
		20.3.4	–	–	1.46×10^6	1.49×10^6	2.55×10^5	1466.1 ± 51.5	1.54×10^6	1.76×10^5	9.82×10^5	1493.4 ± 78.3
		20.3.5	–	–	1.89×10^6	1.93×10^6	3.04×10^5	1895.7 ± 71.9	2.23×10^6	2.51×10^6	1.12×10^6	1355.8 ± 72.6
$N = 30$, $K = 3$	990	30.3.1	–	–	4.94×10^6	5.03×10^6	6.67×10^5	4950.8 ± 180.6	6.16×10^6	6.98×10^6	2.85×10^6	6015.2 ± 290.4
		30.3.2	–	–	3.28×10^6	3.35×10^6	5.10×10^5	3290.1 ± 135.7	3.63×10^6	4.12×10^6	2.01×10^6	4565.5 ± 210.7
		30.3.3	–	–	5.63×10^6	5.72×10^6	7.57×10^5	5645.9 ± 210.1	7.73×10^6	8.66×10^6	3.21×10^6	5315.1 ± 255.2
		30.3.4	–	–	7.45×10^6	7.58×10^6	9.10×10^5	7460.2 ± 290.0	1.09×10^7	1.25×10^7	4.88×10^6	4465.3 ± 205.1
		30.3.5	–	–	5.11×10^6	5.19×10^6	7.11×10^5	5120.0 ± 195.4	5.78×10^6	6.51×10^6	3.01×10^6	5995.5 ± 280.6

(*continued on next page*)

Table 4 (*continued*)

	# of entries in (X, Z)	Instance ID	Exhaustive		GA				MSM			
			Min. cost [$/year]	Time [s]	Best cost [$/year]	Mean cost [$/year]	St. dev. [$/year]	Time [s] Mean ± St. dev.	Best cost [$/year]	Mean cost [$/year]	St. dev. [$/year]	Time [s] Mean ± St. dev.
$N = 40$, $K = 3$	1720	40.3.1	–	–	8.53×10^6	8.68×10^6	1.15×10^6	8540.3 ± 350.2	1.02×10^7	1.15×10^7	4.75×10^6	17310.4 ± 750.0
		40.3.2	–	–	9.46×10^6	9.61×10^6	1.25×10^6	9475.7 ± 410.3	1.00×10^7	1.14×10^7	4.90×10^6	11120.7 ± 510.3
		40.3.3	–	–	7.43×10^6	7.56×10^6	9.80×10^5	7440.4 ± 295.5	8.06×10^6	9.01×10^6	4.11×10^6	21450.6 ± 980.5
		40.3.4	–	–	1.02×10^7	1.04×10^7	1.41×10^6	10220.5 ± 480.6	1.17×10^7	1.32×10^7	5.65×10^6	17430.0 ± 810.1
		40.3.5	–	–	1.31×10^7	1.33×10^7	1.58×10^6	1315.0 ± 46.8	1.82×10^7	2.09×10^7	8.82×10^6	13210.5 ± 640.4

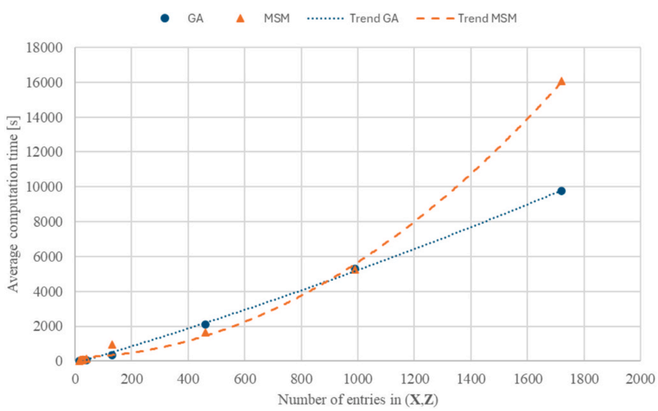

Fig. 3. Average computation time of GA and MSM.

and the value of performance metrics for the baseline condition under each scenario are given in Appendix G in the online supplement. Comments and discussion on results are provided in Section 6.4.5, while managerial insights are given in Section 6.5.

6.4.1. *Low correlation scenario*

In this scenario, we set $\sigma_{ij} = 0.1 \bullet \sigma_i \sigma_j$, for each $i \neq j$. Tables 7 and 8 present the results for case 1–the solution in $(\mathbf{X}, \mathbf{Z})$ of the baseline condition is kept fixed–and case 2–the solution in $(\mathbf{X}, \mathbf{Z})$ adapts to the new setting–, respectively. Percentage changes are calculated relative to the baseline condition. In case 2, if changing the value of a parameter does not lead to a variation in the solution in $(\mathbf{X}, \mathbf{Z})$, then the corresponding results are evidently identical to those obtained in case 1 and can be omitted from Table 8; the same approach is applied when presenting the results for the other scenarios.

6.4.2. *High correlation scenario*

This scenario considers $\sigma_{ij} = 0.8 \bullet \sigma_i \sigma_j$, for each $i \neq j$. Tables 9 and 10 give the results for case 1 and case 2, respectively.

6.4.3. *Block correlation scenario*

This scenario considers $\sigma_{ij} = 0.8 \bullet \sigma_i \sigma_j$, for $i, j = 1, 2$, and $\sigma_{ij} = 0.2 \bullet \sigma_i \sigma_j$, for the other pairs. Tables 11 and 12 give the results for case 1 and case 2, respectively.

6.4.4. *Mixed correlation scenario*

This scenario considers $\sigma_{ij} = 0.8 \bullet \sigma_i \sigma_j$, for $i, j = 1, 2$, and $\sigma_{ij} = -0.8 \bullet \sigma_i \sigma_j$, for the other pairs. Tables 13 and 14 give the results for case 1 and case 2, respectively.

6.4.5. *Discussion of results*

This discussion examines the effects of covariance structure and other parameters. Managerial insights are then presented in Section 6.5.

Table 5
Instance-by-instance statistical comparison of GA and MSM mean performance using the Wilcoxon signed-rank test. Significant p-values (< 0.05) are highlighted in bold.

	Instance ID	p-value		Instance ID	p-value
$N = 3$, $K = 2$	3.2.1	0.258	$N = 10$, $K = 3$	10.3.1	**< 0.001**
	3.2.2	0.179		10.3.2	**< 0.001**
	3.2.3	0.115		10.3.3	**< 0.001**
	3.2.4	0.317		10.3.4	**< 0.001**
	3.2.5	0.091		10.3.5	**< 0.001**
$N = 4$, $K = 2$	4.2.1	0.065	$N = 20$, $K = 3$	20.3.1	**< 0.001**
	4.2.2	**0.011**		20.3.2	**< 0.001**
	4.2.3	0.058		20.3.3	**< 0.001**
	4.2.4	0.061		20.3.4	**< 0.001**
	4.2.5	**0.015**		20.3.5	**< 0.001**
$N = 4$, $K = 3$	4.3.1	**0.046**	$N = 30$, $K = 3$	30.3.1	**< 0.001**
	4.3.2	**0.011**		30.3.2	**< 0.001**
	4.3.3	0.057		30.3.3	**< 0.001**
	4.3.4	**0.015**		30.3.4	**< 0.001**
	4.3.5	0.053		30.3.5	**< 0.001**
$N = 5$, $K = 3$	5.3.1	**< 0.001**	$N = 40$, $K = 3$	40.3.1	**< 0.001**
	5.3.2	**< 0.001**		40.3.2	**< 0.001**
	5.3.3	**< 0.001**		40.3.3	**< 0.001**
	5.3.4	**< 0.001**		40.3.4	**< 0.001**
	5.3.5	**< 0.001**		40.3.5	**< 0.001**

6.4.5.1. *Impact of covariance structure.*

The structure of the covariance matrix was found to exert a pronounced effect on the solution and the system's performance metrics:

- Bin capacity utilisation and overflow penalties: High and mixed positive correlations, contrasted with low correlation, increased simultaneous waste generation peaks. This amplified the likelihood

Table 6
VSS across instances of increasing complexity.

Instance ID	Cost of deterministic solution* [$/year]	Cost of stochastic solution [$/year]	VSS [$/year]	Relative VSS (% Cost Reduction)
3.2.1	3.07×10^5	9.91×10^4	2.08×10^5	67.7%
4.3.3	2.59×10^5	6.05×10^4	1.99×10^5	76.8%
5.3.3	2.82×10^6	7.99×10^4	2.74×10^6	97.2%

* This value represents the expected cost of the deterministic solution when evaluated under stochastic conditions.

of bins becoming overloaded or underutilised if unaddressed. Consequently, under high inter-node dependence, penalty costs for exceeding and under-exploiting bin capacity were more pronounced, especially in fixed solutions. An adaptable system often responded by favouring larger bin capacities, more frequent, or both, to mitigate these amplified risks. For instance, a significant increase in waste generation rates under high correlation, particularly within a fixed system, led to a more substantial rise in overflow penalties compared to a low correlation scenario, prompting the adaptable model to adjust bin types or collection multipliers more aggressively. Conversely, under low correlation, with more isolated variability, the impetus for extensive adjustments to manage simultaneous surges was diminished, leading to greater inherent stability in bin utilisation.

- Waste collection frequency and truck capacity: Strong positive correlation across nodes led to greater volatility in accumulated waste per collection cycle. This heightened the risk of exceeding truck capacity, incurring substantial $\mathfrak{R}_{\text{tc}}$ penalties, necessitating stricter collection frequency adjustments or meticulous node grouping. The block correlation scenario, emulating geographically correlated clusters, often led to region-specific collection frequency adjustments. Nodes in highly correlated urban clusters typically required more frequent emptying (lower m_i values selected by the adaptable model) than those in less interdependent peripheral areas. If truck capacity was substantially reduced under high correlation, particularly with a fixed system, the increase in truck overcapacity penalties was considerably larger than under low correlation, pushing an adaptable model towards significant frequency increases or changes in bin allocation.
- Solution robustness and adaptability: System responsiveness to changes in other parameters varied significantly with the covariance structure. Under high correlation, parameter variations had an amplified effect on overall system cost and its components compared to low correlation. This heightened sensitivity implies reduced system robustness to parameter uncertainty when waste dynamics are tightly coupled. Managerially, this underscores the need for vigilant monitoring and more frequent parameter updates in highly correlated environments. The benefits of an adaptable solution were also more pronounced under high correlation, as the model could make more impactful reconfigurations to mitigate amplified risks.

6.4.5.2. Effects of individual parameters. To visually summarise the relative impact of these parameter variations and provide an intuitive hierarchy of the system's cost drivers, Fig. 4 presents a sensitivity analysis chart for the highly correlated, adaptable scenario (synthesizing data from Tables 9 and 10). As illustrated, when the system is allowed to optimally reconfigure its network, the total expected cost rate ($\mathfrak{R}$) remains most acutely sensitive to severe reductions in truck capacity (Q) and changes in the disposal facility's operating cost (h), while parameters like the major collection cost (A) exert a comparatively modest influence.

The sensitivity analysis further explored the impact of individual parameters on the optimal solution and associated cost rates:

- Major collection cost (A): Variations in A directly influenced the optimal basic collection period T. Increased A marginally raised overall system cost; the adaptable solution generally responded by increasing T. The relative impact on bin penalties was often less pronounced. Even significant A changes often resulted in only moderate percentage changes in overall costs and T, suggesting that if A is stable, optimisation efforts might be better directed elsewhere.
- Truck capacity (Q): Q strongly influenced costs from exceeding truck capacity ($\mathfrak{R}_{\text{tc}}$). Reduced Q, especially in fixed solutions, sharply increased these penalties. The adaptable solution responded with significant frequency adjustments or bin modifications, highlighting the critical role of proper fleet sizing. Increasing Q substantially mitigated $\mathfrak{R}_{\text{tc}}$, though overall system cost sometimes remained largely unchanged, suggesting strategic benefits for higher-capacity trucks in volatile or highly correlated areas.

Table 7
Results obtained keeping the baseline solution in $(\mathbf{X}, \mathbf{Z})$ fixed under the low correlation scenario.

Parameter	% change	% change							
		$\mathfrak{R}$	$\mathfrak{R}_2$	$\mathfrak{R}_{\text{tc}}$	$\mathfrak{R}_{\text{bc}}$	$\mathfrak{R}_{\text{nbc}}$	T	m_2	m_3
A	+80%	+8.2%	0%	+17.5%	+3.3%	−3.2%	+0.2%	0%	0%
	−80%	−8.3%	0%	−15.1%	−3.3%	+3.3%	−0.2%	0%	0%
Q	+80%	≈ 0%	0%	−100%	≈ 0%	≈ 0%	≈ 0%	0%	0%
	−80%	+289.8%	0%	≫ 100%	−33.7%	+43.1%	−2.6%	0%	0%
h	+80%	+32.1%	0%	−47.4%	−12.3%	+13.3%	−0.9%	0%	0%
	−80%	−32.4%	0%	+88.5%	+13.7%	−12.5%	+0.9%	0%	0%
θ	+80%	≈ 0%	0%	+79.7%	≈ 0%	≈ 0%	≈ 0%	0%	0%
	−80%	≈ 0%	0%	−80.0%	≈ 0%	≈ 0%	≈ 0%	0%	0%
$d_i, \forall i$	+80%	+34.5%	+73.1%	−98.9%	+41.6%	+26.9%	−44.2%	0%	0%
	−80%	−41.2%	−73.1%	≫ 100%	−73.7%	−26.4%	+358.3%	0%	0%
$\sigma_i, \forall i$	+80%	+20.9%	0%	≫ 100%	+82.4%	+79.0%	−0.5%	0%	0%
	−80%	−20.2%	0%	−100%	−81.3%	−77.3%	+0.4%	0%	0%
$\pi_{0i}, \forall i$	+80%	+7.5%	0%	−88.7%	+14.9%	+47.2%	−2.8%	0%	0%
	−80%	−13.8%	0%	≫ 100%	−55.6%	−66.6%	+6.7%	0%	0%
$\pi_{1i}, \forall i$	+80%	+8.5%	0%	≫ 100%	+53.8%	+8.8%	+3.2%	0%	0%
	−80%	−17.8%	0%	−100%	−81.3%	−42.2%	−8.7%	0%	0%
$a_{ik}, \forall i,k$	+80%	+1.9%	0%	+3.8%	+0.8%	−0.8%	+0.1%	0%	0%
	−80%	−1.9%	0%	−3.7%	−0.8%	+0.8%	−0.1%	0%	0%
$C_k, \forall k$	+80%	+136.2%	0%	≫ 100%	−91.4%	+151.2%	+66.0%	0%	0%
	−80%	+47.9%	0%	−100%	+223.2%	+42.4%	−78.9%	0%	0%

Table 8
Results obtained by allowing the solution in $(\mathbf{X}, \mathbf{Z})$ to vary under the low correlation scenario.

Parameter	% change	% change						Node	Target bin	Bin type	m_i
		$\mathfrak{R}$	$\mathfrak{R}_2$	$\mathfrak{R}_{tc}$	$\mathfrak{R}_{bc}$	$\mathfrak{R}_{nbc}$	T				
A	−80%	−14.5%	−38.5%	−100%	+21.7%	+17.2%	−72.1%	1	2	–	–
								2	2	1	2
								3	3	1	5
Q	−80%	+45.7%	−88.0%	≫ 100%	+49.0%	+12.8%	−79.3%	1	1	1	6
								2	2	1	5
								3	3	1	7
h	+80%	+22.4%	+50.5%	−100%	+6.0%	+8.9%	−40.3%	1	2	–	–
								2	2	2	1
								3	2	–	–
d_1	+80%	+13.7%	−34.1%	−98.1%	+63.0%	+59.6%	−36.4%	1	1	2	1
								2	1	–	–
								3	3	1	2
	−80%	–22.0%	−80.8%	+267.8%	−5.8%	−3.3%	+33.0%	1	2	–	–
								2	2	2	1
								3	3	1	1
π_{03}	+80%	+3.1%	+50.5%	−100%	+14.0%	+1.6%	−40.0%	1	2	–	–
								2	2	2	1
								3	2	–	–
π_{13}	+80%	+3.1%	+50.5%	−100%	+14.0%	+1.6%	−40.0%	1	2	–	–
								2	2	2	1
								3	2	–	–
π_{01}	−80%	−13.5%	−35.0%	−100%	+6.7%	−43.5%	–33.3%	1	1	1	1
								2	1	–	–
								3	3	1	2
π_{11}	−80%	−11.9%	−34.1%	−98.6%	−40.2%	−11.1%	−25.5%	1	1	2	1
								2	1	–	–
								3	3	1	2
C_k, $\forall k$	+80%	+1.8%	+50.1%	−100%	+0.3%	−3.9%	−28.2%	1	2	–	–
								2	2	1	1
								3	2	–	–
	−80%	+29.9%	−87.6%	−100%	+271.5%	+85.1%	−70.8%	1	1	1	1
								2	2	2	1
								3	3	1	1

Table 9
Results obtained keeping the baseline solution in $(\mathbf{X}, \mathbf{Z})$ fixed under the high correlation scenario.

Parameter	% change	% change							
		$\mathfrak{R}$	$\mathfrak{R}_2$	$\mathfrak{R}_{tc}$	$\mathfrak{R}_{bc}$	$\mathfrak{R}_{nbc}$	T	m_2	m_3
A	+80%	+8.0	0%	+15.2%	+3.3%	−3.3%	+0.3%	0%	0%
	−80%	−8.0%	0%	−13.4%	−3.3%	+3.4%	−0.3%	0%	0%
Q	+80%	≈ 0%	0%	−100%	+0.2%	−0.2%	≈ 0%	0%	0%
	−80%	+280.1%	0%	≫ +100%	−34.0%	+44.1%	−3.0%	0%	0%
h	+80%	+31.1%	0%	−43.1%	−12.3%	+13.6%	−1.0%	0%	0%
	−80%	−31.4%	0%	+74.1%	+13.7%	−12.7%	+1.0%	0%	0%
θ	+80%	≈ 0%	0%	+78.5%	−0.2%	+0.2%	≈ 0%	0%	0%
	−80%	≈ 0%	0%	−79.8%	+0.2%	−0.2%	≈ 0%	0%	0%
d_i, $\forall i$	+80%	+34.4%	+73.1%	−96.2%	+42.3%	+26.2%	−44.2%	0%	0%
	−80%	−41.7%	−73.1%	≫ +100%	−73.9%	−25.7%	+353.1%	0%	0%
σ_i, $\forall i$	+80%	+23.3%	0%	≫ +100%	+77.6%	+84.0%	−0.8%	0%	0%
	−80%	–22.2%	0%	−100%	−80.6%	−78.2%	+0.5%	0%	0%
π_{0i}, $\forall i$	+80%	+8.3%	0%	−85.4%	+14.4%	+48.3%	−3.2%	0%	0%
	−80%	−15.1%	0%	≫ +100%	−57.6%	−65.0%	+7.2%	0%	0%
π_{1i}, $\forall i$	+80%	+9.2%	0%	+547.2%	+53.5%	+7.7%	+3.7%	0%	0%
	−80%	−19.5%	0%	−100%	−81.7%	−41.8%	−9.8%	0%	0%
a_{ik}, $\forall i, k$	+80%	+1.8%	0%	+3.3%	+0.8%	−0.8%	+0.1%	0%	0%
	−80%	−1.8%	0%	−3.2%	−0.8%	+0.8%	−0.1%	0%	0%
C_k, $\forall k$	+80%	+129.2%	0%	≫ +100%	−49.1%	+558.6%	−38.2%	+200%	+100%
	−80%	+49.8%	0%	−100%	+224.7%	+39.5%	−78.7%	0%	0%

- Operating cost at the central disposal facility (h): Increased h raised overall system cost ($\mathfrak{R}$) due to its linear dependence on processed waste quantity. Responses in $\mathfrak{R}$ were often relatively symmetric for $\pm$ 80% changes in h. However, impacts on specific penalties could differ based on baseline conditions. Municipalities must account for h fluctuations when planning.
- Penalty cost parameters (θ, π_{0i}, and π_{1i}): θ (truck overcapacity penalty) proportionally influenced $\mathfrak{R}_{tc}$; a substantial increase in θ in fixed solutions led to a comparable $\mathfrak{R}_{tc}$ rise without majorly altering other metrics. Bin capacity penalties–π_{0i} (exceeding) and π_{1i} (underutilisation)–exhibited more complex interactions. In highly correlated scenarios, increased bin penalties prompted the adaptable model to adjust bin capacities, collection frequencies, or waste allocation. This sensitivity underscores the need for careful cost calibration.
- Waste generation rate (d_i) and variability (σ_i): Mean generation rates (d_i) and standard deviations (σ_i) are fundamental. Increased d_i raised

Table 10
Results obtained by allowing the solution in $(\mathbf{X},\mathbf{Z})$ to vary under the high correlation scenario.

Parameter	% change	% change						Node	Target bin	Bin type	m_i
		$\mathfrak{R}$	$\mathfrak{R}_2$	$\mathfrak{R}_{\text{tc}}$	$\mathfrak{R}_{\text{bc}}$	$\mathfrak{R}_{\text{nbc}}$	T				
A	−80%	−10.9%	−38.5%	−100%	+32.6%	+24.1%	−72.1%	1	2	–	–
								2	2	1	2
								3	3	1	5
Q	−80%	+40.9%	−88.0%	$\gg$ +100%	+30.7%	+0.3%	−79.3%	1	1	1	6
								2	2	1	5
								3	3	1	7
h	+80%	+30.6%	+50.5%	−100%	+39.6%	+38.5%	−40.5%	1	2	–	–
								2	2	2	1
								3	2	–	–
d_1	+80%	+17.3%	−34.1%	−91.3%	+81.5%	+56.4%	−36.1%	1	1	2	1
								2	1	–	–
								3	3	1	2
	−80%	−21.0%	−73.5%	−100%	+1.5%	+23.8%	+5.0%	1	3	–	–
								2	2	1	1
								3	3	1	1
π_{01}	−80%	−15.0%	−35.0%	−100%	+4.5%	−51.8%	−32.3%	1	1	1	1
								2	1	–	–
								3	3	1	2
π_{11}	−80%	−13.4%	−34.1%	−98.2%	−48.3%	−8.7%	−26.6%	1	1	2	1
								2	1	–	–
								3	3	1	2
C_k, $\forall k$	+80%	+9.6%	+50.1%	−100%	+31.6%	+22.3%	−28.3%	1	2	–	–
								2	2	1	1
								3	2	–	–
	−80%	+25.6%	−87.6%	−100%	+225.7%	+64.5%	−70.8%	1	1	1	1
								2	2	2	1
								3	3	1	1

Table 11
Results obtained keeping the baseline solution in $(\mathbf{X},\mathbf{Z})$ fixed under the block correlation scenario.

Parameter	% change	% change							
		$\mathfrak{R}$	$\mathfrak{R}_2$	$\mathfrak{R}_{\text{tc}}$	$\mathfrak{R}_{\text{bc}}$	$\mathfrak{R}_{\text{nbc}}$	T	m_2	m_3
A	+80%	+8.2%	0%	+17.2%	+3.3%	−3.3%	+0.2%	0%	0%
	−80%	−8.2%	0%	−14.9%	−3.3%	+3.4%	−0.2%	0%	0%
Q	+80%	≈ 0%	0%	−100%	+0.1%	−0.1%	≈ 0%	0%	0%
	−80%	+288.2%	0%	$\gg$ +100%	−33.9%	+43.4%	−2.6%	0%	0%
h	+80%	+32.0%	0%	−46.8%	−12.3%	+13.4%	−0.9%	0%	0%
	−80%	−32.2%	0%	+86.4%	+13.7%	−12.6%	+0.9%	0%	0%
θ	+80%	≈ 0%	0%	+79.6%	≈ 0%	≈ 0%	≈ 0%	0%	0%
	−80%	≈ 0%	0%	−80.0%	≈ 0%	≈ 0%	≈ 0%	0%	0%
d_i, $\forall i$	+80%	+34.5%	+73.1%	−98.6%	+41.7%	+26.8%	−44.2%	0%	0%
	−80%	−41.3%	−73.1%	$\gg$ +100%	−73.8%	−26.2%	+357.4%	0%	0%
σ_i, $\forall i$	+80%	+21.3%	0%	$\gg$ +100%	+81.6%	+79.7%	−0.5%	0%	0%
	−80%	−20.5%	0%	−100%	−81.2%	−77.4%	+0.4%	0%	0%
π_{0i}, $\forall i$	+80%	+7.6%	0%	−88.2%	+14.7%	+47.5%	−2.9%	0%	0%
	−80%	−14.0%	0%	$\gg$ +100%	−55.9%	−66.6%	+6.8%	0%	0%
π_{1i}, $\forall i$	+80%	+8.6%	0%	+726.8%	+53.9%	+8.4%	+3.3%	0%	0%
	−80%	−18.1%	0%	−100%	−81.5%	−42.1%	−8.9%	0%	0%
a_{ik}, $\forall i,k$	+80%	+1.9%	0%	+3.7%	+0.8%	−0.8%	+0.1%	0%	0%
	−80%	−1.9%	0%	−3.6%	−0.8%	+0.8%	−0.1%	0%	0%
C_k, $\forall k$	+80%	+135.0%	0%	$\gg$ +100%	−91.5%	+151.7%	+65.7%	0%	0%
	−80%	+48.2%	0%	−100%	+223.6%	+41.8%	−78.9%	0%	0%

mean waste accumulation, influencing delivery costs and significantly increasing $\mathfrak{R}_{\text{bc}}$ if unaddressed. The adaptable solution would typically select larger bins or increase collection frequency. Increased σ_i generally led to higher bin fill uncertainties, compelling the adaptable model to adjust collection intervals and bin types/multipliers. In high inter-node correlation scenarios, these effects were magnified, requiring accurate, current data for policy fine-tuning.

- Minor collection cost (a_{ik}): Though a smaller cost proportion, a_{ik} influenced allocation and frequency, especially in dense or high-frequency areas. Slight percentage changes in a_{ik} had modest repercussions on optimal configuration and costs. Maintaining competitive a_{ik} rates remains important.
- Bin capacity (C_k): C_k significantly shaped the cost profile, directly affecting $\mathfrak{R}_{\text{bc}}$ and $\mathfrak{R}_{\text{nbc}}$. Increased C_k, especially in adaptable solutions, generally reduced $\mathfrak{R}_{\text{bc}}$, unless causing substantial under-utilisation. Reduced C_k necessitated more frequent collections, increasing operational costs and potentially $\mathfrak{R}_{\text{bc}}$. This balance underscores the importance of matching bin capacities to local profiles and frequencies.

6.4.5.3. Interplay between parameter changes and solution adaptability. A consistent finding is the dynamic interplay between fixed versus adaptable solution structures. This comparison effectively serves as a benchmark of a static versus an adaptable system design philosophy, directly addressing the value of re-optimising strategic configurations in

Table 12
Results obtained by allowing the solution in $(\mathbf{X}, \mathbf{Z})$ to vary under the block correlation scenario.

Parameter	% change	% change						Node	Target bin	Bin type	m_i
		$\Re$	$\Re_2$	$\Re_{tc}$	$\Re_{bc}$	$\Re_{nbc}$	T				
A	−80%	−9.6%	−0.38%	−100%	+13.7%	+8.4%	−66.6%	1	3	–	–
								2	2	1	3
								3	3	1	2
Q	−80%	+44.9%	−88.0%	≫ +100%	+45.8%	+10.6%	−79.3%	1	1	1	6
								2	2	1	5
								3	3	1	7
h	+80%	+28.3%	+50.5%	−100%	+30.5%	+31.1%	−40.4%	1	2	–	–
								2	2	2	1
								3	2	–	–
d_1	+80%	+20.6%	−34.1%	−56.4%	+102.5%	+72.6%	−36.1%	1	1	2	1
								2	1	–	–
								3	3	1	2
	−80%	−21.0%	−73.5%	−100%	+10.3%	+22.5%	+5.6%	1	3	–	–
								2	2	1	1
								3	3	1	1
π_{01}	−80%	−12.5%	−35.0%	−100%	+16.6%	−46.8%	−32.3%	1	1	1	1
								2	1	–	–
								3	3	1	2
π_{11}	−80%	−10.9%	−34.1%	−90.8%	−42.3%	+0.7%	−26.6%	1	1	2	1
								2	1	–	–
								3	3	1	2
$C_k, \forall k$	+80%	+7.2%	+50.1%	−100%	+23.3%	+15.7%	−28.2%	1	2	–	–
								2	2	1	1
								3	2	–	–
	−80%	+29.2%	−87.1%	−100%	+263.4%	+81.5%	−70.8%	1	1	1	1
								2	2	2	1
								3	3	1	1

Table 13
Results obtained keeping the baseline solution in $(\mathbf{X}, \mathbf{Z})$ fixed under the mixed correlation scenario.

Parameter	% change	% change							
		$\Re$	$\Re_2$	$\Re_{tc}$	$\Re_{bc}$	$\Re_{nbc}$	T	m_2	m_3
A	+80%	+8.8%	0%	+16.9%	+2.6%	−2.4%	+0.1%	0%	0%
	−80%	−8.9%	0%	−14.6%	−2.5%	+2.5%	−0.2%	0%	0%
Q	+80%	≈ 0%	0%	−100%	≈ 0%	≈ 0%	≈ 0%	0%	0%
	−80%	+312.0%	0%	≫ +100%	−26.2%	+32.6%	−1.5%	0%	0%
h	+80%	+34.6%	0%	−46.2%	−9.6%	+10.1%	−0.5%	0%	0%
	−80%	−34.8%	0%	+84.4%	+10.5%	−9.4%	+0.5%	0%	0%
θ	+80%	≈ 0%	0%	+80.0%	≈ 0%	≈ 0%	≈ 0%	0%	0%
	−80%	≈ 0%	0%	−80.0%	≈ 0%	≈ 0%	≈ 0%	0%	0%
$d_i, \forall i$	+80%	+34.5%	+73.1%	−100%	+37.6%	+30.7%	−44.3%	0%	0%
	−80%	−39.7%	−73.1%	≫ +100%	−69.6%	−32.2%	+373.4%	0%	0%
$\sigma_i, \forall i$	+80%	+16.1%	0%	≫ +100%	+89.3%	+72.5%	−0.2%	0%	0%
	−80%	−15.7%	0%	−100%	−84.5%	−73.4%	+0.3%	0%	0%
$\pi_{0i}, \forall i$	+80%	+6.2%	0%	−87.4%	+29.4%	+35.9%	−1.6%	0%	0%
	−80%	−10.1%	0%	≫ +100%	−58.0%	−55.0%	+4.2%	0%	0%
$\pi_{1i}, \forall i$	+80%	+7.2%	0%	+767.3%	+42.4%	+42.4%	+1.8%	0%	0%
	−80%	−13.1%	0%	−100%	−70.6%	−44.8%	−6.0%	0%	0%
$a_{ik}, \forall i, k$	+80%	+2.0%	0%	+3.7%	+0.6%	−0.6%	≈ 0%	0%	0%
	−80%	−2.0%	0%	−3.5%	−0.6%	+0.6%	≈ 0%	0%	0%
$C_k, \forall k$	+80%	+140.8%	0%	≫ +100%	−74.6%	+688.7%	−39.8%	+100%	+200%
	−80%	+43.5%	0%	−100%	+205.7%	+59.0%	−79.4%	0%	0%

response to changing operational parameters.

- Fixed solution scenario: With fixed **X** and **Z** matrices–representing a static infrastructure–the analysis captured the direct, often unmitigated, effect of parameter variations. While operational frequencies could be adjusted, the physical layout was immutable. As seen in results, this rigidity often led to sharp increases in penalty costs, highlighting the vulnerability of such systems.
- Variable solution scenario: Allowing the full system configuration (**X**, **Z**, T, **m**) to adapt revealed systemic responses. Parameter changes often triggered bin reallocation, type/capacity adjustments, waste source reassignment, and frequency modifications. This generally yielded substantially better performance (lower $\Re$ or less severe increases) as the model actively mitigated adverse effects. For instance, increased σ_i under high correlation led the adaptable model to more conservative bin strategies and higher frequencies, a response impossible in fixed scenarios. This consistent outperformance demonstrates the significant economic value of an adaptable framework and supports the integration of dynamic re-evaluation into waste management planning.

The undeniable economic value of this adaptability is visually captured in Fig. 5. By isolating extreme capacity stress scenarios from the high correlation environment–specifically, drastic reductions or increases in available truck (Q) and bin (C_k) capacities–the graph vividly contrasts the operational consequences. While the rigid, fixed configuration suffers massive, unmitigated cost escalations (e.g., a 280.1% spike when truck capacity drops), the adaptable framework successfully

Table 14
Results obtained by allowing the solution in $(\mathbf{X}, \mathbf{Z})$ to vary under the mixed correlation scenario.

Parameter	% change	% change						Node	Target bin	Bin type	m_i
		$\Re$	$\Re_2$	$\Re_{tc}$	$\Re_{bc}$	$\Re_{nbc}$	T				
A	−80%	−13.4%	+50.1%	−100%	+41.0%	+27.1%	−60.0%	1	2	–	–
								2	2	1	1
								3	2	–	–
Q	−80%	+56.6%	−88.0%	≫ +100%	+107.3%	+52.8%	−79.3%	1	1	1	6
								2	2	1	5
								3	3	1	7
h	+80%	+22.6%	+50.1%	−100%	+45.1%	+23.6%	−60.0%	1	2	–	–
								2	2	1	1
								3	2	–	–
d_1	+80%	+14.7%	+4.0%	−100%	+26.1%	+70.3%	−52.6%	1	1	1	1
								2	2	1	2
								3	1	–	–
	−80%	−19.7%	−73.5%	−100%	+70.4%	+6.1%	+8.2%	1	3	–	–
								2	2	1	1
								3	3	1	1
π_{01}	−80%	−6.7%	+57.9%	−100%	−23.5%	−72.9%	−35.8%	1	1	2	1
								2	1	–	–
								3	1	–	–
π_{11}	−80%	−6.4%	+57.9%	−100%	−62.5%	−31.0%	−42.8%	1	1	2	1
								2	1	–	–
								3	1	–	–
$C_k, \forall k$	+80%	+2.4%	+50.1%	−100%	+1.8%	−2.6%	−28.0%	1	2	–	–
								2	2	1	1
								3	2	–	–
	−80%	+39.6%	−87.6%	−100%	+416.5%	+150.8%	−70.8%	1	1	1	1
								2	2	2	1
								3	3	1	1

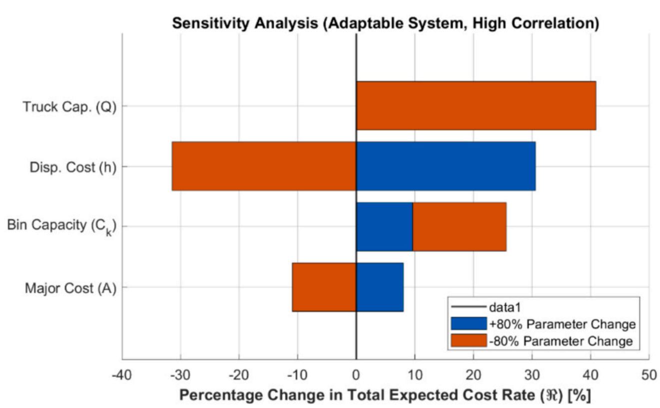

Fig. 4. Sensitivity analysis of the expected total system cost rate ($\Re$) to ± 80% variations in some core parameters.

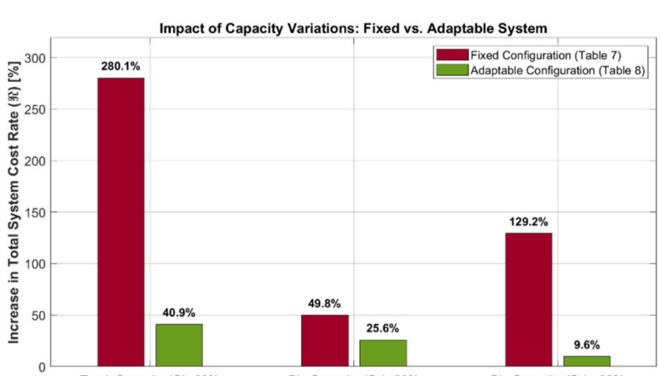

Fig. 5. Comparison of the percentage increase in expected total system cost rate between a fixed and an adaptable network configuration in the high correlation scenario.

cushions the shock through dynamic resource reallocation, restricting the cost increase to just 40.9%. This visually underscores the critical vulnerability of static MSW systems when facing high uncertainty.

6.5. Managerial implications and policy recommendations

The comprehensive scenario and sensitivity analyses presented in this study yield several explicit and actionable insights for practitioners and policymakers in municipal waste management. This section translates the numerical results into strategic recommendations, directly addressing the key questions of how system design should respond to uncertainty and parameter changes.

6.5.1. How correlations in waste generation affect system design

A one-size-fits-all approach to waste collection is demonstrably suboptimal. Our results show that the covariance structure of waste generation is a first-order driver of system design:

- Highly correlated regions, prone to simultaneous waste surges (e.g., commercial districts, dense urban centres), require a strategy of flexible capacity and higher service frequency. Planners in these areas must prioritise investing in larger-capacity bins or designing more frequent collection schedules (lower m_i values) to mitigate the amplified risk of system-wide overflows.
- Low-correlation areas (e.g., heterogeneous residential suburbs) are more stable, as local variations tend to average out. These areas can support more cost-effective designs with standard-sized bins and less frequent, fixed collection schedules. A nuanced, data-driven understanding of these local dynamics is therefore crucial for effective system design.

6.5.2. When adaptive reconfiguration outperforms static setups

Our analysis consistently shows that an adaptive system–one where bin locations, types, and allocations can be re-evaluated–significantly outperforms a static one. The policy implication is clear: authorities should establish mechanisms for frequent, data-driven re-evaluation of the network. The value of this adaptability is most pronounced under two conditions:

- When underlying parameters are volatile: If costs (fuel, disposal) or waste generation rates are expected to change, a static design becomes a significant liability. The ability to re-run the optimisation and reconfigure the network allows the system to evolve and maintain cost-effectiveness.
- In highly correlated systems: Because high covariance amplifies the impact of any parameter change, an adaptive framework is even more critical in these environments. For instance, a small increase in waste generation rates can cause a massive spike in overflow penalties that only a reconfiguration can effectively mitigate.

6.5.3. Key sensitivities and thresholds from a policy perspective

From a policy perspective, our sensitivity analysis reveals the critical levers that planners must monitor and manage:

- Infrastructure investment (truck and bin capacity): The system's high sensitivity to truck and bin capacity highlights the need for well-calibrated investments. Under-investing is a false economy, as the resulting overflow and overcapacity penalties can far outweigh capital savings. A detailed cost-benefit analysis, guided by a stochastic model like ours, should inform these strategic trade-offs between capital expenditure and operational flexibility.
- Operational cost monitoring (disposal, collection, penalties): Operational costs and penalty parameters must be accurately calibrated and periodically reviewed. Agencies must update these figures due to external factors like fuel prices or new regulations. Post-update, model re-runs are essential to assess impacts on network design and schedules, ensuring continued cost-effectiveness.
- Regional segmentation: Findings from block and mixed correlation scenarios suggest that municipalities with heterogeneous waste patterns benefit from regional clustering. Identifying sub-areas with distinct local correlations allows for tailored collection strategies, improving both efficiency and customer satisfaction.
- Data-driven planning (scenario and real-time analytics): The analysis underscores that waste generation uncertainty, especially when amplified by high correlation, has major cost implications. Developing contingency plans for a range of plausible scenarios is prudent. In the long term, integrating real-time data analytics from sources like "smart bins" offers a significant advantage, enabling continuous updating of model inputs (covariance, generation rates) to pre-empt cost escalations from unpredictable waste surges.

By embracing these insights, MSW authorities can advance towards more efficient, resilient, and economically sustainable waste management systems.

7. Conclusions, limitations, and future perspectives

This study presented a comprehensive framework for jointly optimising key MSW management decisions–bin location, allocation, type selection, and collection frequency–under stochastic waste generation, explicitly considering its covariance structure. The primary contribution lies in this integrated formulation, which quantifies the profound impact of uncertainty and inter-node correlation on tactical and operational planning.

A principal finding is that waste generation interdependence across nodes, captured by the covariance matrix, significantly dictates optimal system configuration and cost-effectiveness. High inter-node correlation necessitates flexible capacity planning and collection strategies to manage aggregated peaks and mitigate amplified penalty risks, contrasting with the more stable designs viable under low correlation. The model also showed sensitivity to parameters like truck/bin capacities and various costs, underscoring the need for accurate data.

Crucially, comparing fixed versus adaptable solution frameworks highlighted the substantial benefits of dynamic resource reallocation. Allowing system adaptation to parameter changes consistently yielded superior performance, emphasising the value of moving beyond static plans towards responsive, perhaps data-driven, MSW management. Furthermore, we demonstrated that this complex MINLP can be reliably solved using a standard GA embedded with a JRP-based heuristic, providing a practical tool for decision-makers.

Managerially, these findings provide a clear roadmap for modern MSW system design, advocating for: (1) strategies tailored to local waste dynamics, particularly the covariance structure; (2) investment in adaptive frameworks that allow for periodic network re-evaluation over static plans; (3) strategic infrastructure balancing informed by stochastic risk analysis; and (4) diligent monitoring of key cost parameters and data inputs. Integrating real-time data and analytics offers further potential for enhanced responsiveness.

It is important to acknowledge certain limitations inherent to the modelling assumptions. First, the assumption that accumulated waste follows a Gaussian distribution represents a trade-off between mathematical tractability and granular precision. While this choice is justified by the CLT, especially for bins with longer collection cycles where waste accumulates over multiple periods, it may be less accurate for bins with very high collection frequencies. In such cases, the temporal aggregation is insufficient for the distribution to converge to normality, potentially leading to an underestimation of "heavy-tail" overflow events. Second, the absence of explicit routing constraints represents a trade-off to maintain computational tractability. While our JRP-like structure–utilising major and minor costs–provides a robust proxy for tactical planning, it does not solve the detailed VRP. Consequently, the model assumes that collection capacity is the primary constraint, potentially overlooking operational nuances such as exact path sequencing, time windows, or specific road network restrictions.

Future research should address these limitations and expand the model's scope in several key directions. To overcome the distributional constraints and the assumption of stationary stochasticity, integrating simulation-optimisation (SimOpt) techniques would allow the system to evaluate costs using arbitrary probability distributions or historical data traces. This would also facilitate the modelling of dynamic uncertainty, such as seasonality, long-term trends in waste generation, or non-stationary variance, potentially through robust optimisation or multi-period stochastic programming approaches. To bridge the gap between tactical planning and operations, future work could explicitly integrate detailed vehicle routing algorithms to optimise collection paths, thereby validating the estimated logistical costs with operational reality. Furthermore, the objective function could be expanded into a multi-objective formulation to explicitly trade off economic costs against environmental performance (e.g., carbon emissions, noise pollution) and social service levels (e.g., equity of access). Finally, exploring the integration of real-time sensor data from "smart bins" could lead to the development of a dynamic, data-driven decision support system, while a real-world case study would further validate the model's operational effectiveness.

CRediT authorship contribution statement

Massimo Bertolini: Writing – review & editing, Visualization, Supervision, Conceptualization. **Davide Castellano:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Davide Mezzogori:** Writing – review & editing, Visualization, Supervision, Conceptualization.

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.

Acknowledgements

This work was partially supported by the European Union and the Government of Italy Ministry of University and Research through the project SusTex: Sustainable Textile, id 2022X2788M, grant number E53D23010440006, and by the European Union through the National Recovery and Resilience Plan (NRRP), Mission 04 Component 2 Investment 1.5 – NextGenerationEU, Call for tender n. 3277 dated 30/12/2021, Award Number: 0001052 dated 23/06/2022.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cie.2026.112057.

References

Akbarpour, N., Salehi-Amiri, A., Hajiaghaei-Keshteli, M., & Oliva, D. (2021). An innovative waste management system in a smart city under stochastic optimization using vehicle routing problem. *Soft Computing, 25*(8), 6707–6727.
Aliahmadi, S. Z., Barzinpour, F., & Pishvaee, M. S. (2021). A novel bi-objective credibility-based fuzzy model for municipal waste collection with hard time windows. *Journal of Cleaner Production, 296*, Article 126364.
Angelelli, E., & Speranza, M. G. (2002). The application of a vehicle routing model to a waste-collection problem: Two case studies. *Journal of the Operational Research Society, 53*(9), 944–952.
Arkin, E., Joneja, D., & Roundy, R. (1989). Computational complexity of uncapacitated multi-echelon production planning problems. *Operations Research Letters, 8*(2), 61–66.
Azri, S. B., Ujang, U., & Abdullah, N. S. (2023). Within cluster pattern identification: A new approach for optimizing recycle point distribution to support policy implementation on waste management in Malaysia. *Waste Management and Research, 41*(3), 687–700.
Azzi, A., Battini, D., Faccio, M., Persona, A., & Sgarbossa, F. (2014). Inventory holding costs measurement: A multi-case study. *International Journal of Logistics Management, 25*(1), 109–132.
Bautista, J., & Pereira, J. (2006). Modeling the problem of locating collection areas for urban waste management. an application to the metropolitan area of Barcelona. *Omega, 34*(6), 617–629.
Beltrami, E. J., & Bodin, L. D. (1974). Networks and vehicle routing for municipal waste collection. *Networks, 4*(1), 65–94.
Bertolini, M., Castellano, D., & Mezzogori, D. (2025). Test instances for analysing the solution procedure. *Mendeley Data, V1*.
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.
Braglia, M., Castellano, D., Santillo, L., & Song, D. (2021). Controlling lead times and minor ordering costs in the joint replenishment problem with stochastic demands under the class of cyclic policies. *International Transactions in Operational Research, 28*(1), 376–400.
Castellano, D., & Glock, C. H. (2021). The average-cost formulation of lot sizing models and inventory carrying charges: A technical note. *Operations Management Research, 14*(1–2), 194–201.
Cavallin, A., Rossit, D. G., Herrán Symonds, V., Rossit, D. A., & Frutos, M. (2020). Application of a methodology to design a municipal waste pre-collection network in real scenarios. *Waste Management and Research, 38*(1_suppl), 117–129.
Cuellar-Usaquén, D., Ulmer, M. W., Antons, O., & Arlinghaus, J. C. (2025). Dynamic multi-period recycling collection routing with uncertain material quality. *OR Spectrum, 47*(3), 699–742.
Dyson, B. E., & Chang, N. (2005). Forecasting municipal solid waste generation in a fast-growing urban region with system dynamics modeling. *Waste Management, 25*(7), 669–679.
Eiselt, H. A., & Marianov, V. (2015). Location modeling for municipal solid waste facilities. *Computers and Operations Research, 62*, 305–315.
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.
Ghiani, G., Laganà, D., Manni, E., & Triki, C. (2012). Capacitated location of collection sites in an urban waste management system. *Waste Management, 32*(7), 1291–1296.
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.
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*(3), 1613–1639.
Kaspi, M., & Rosenblatt, M. J. (1983). An improvement of silver's algorithm for the joint replenishment problem. *IIE Transactions (Institute of Industrial Engineers), 15*(3), 264–267.
Khouja, M., & Goyal, S. (2008). A review of the joint replenishment problem literature: 1989-2005. *European Journal of Operational Research, 186*(1), 1–16.
Kim, J., Manna, A., Roy, A., & Moon, I. (2025). Clustered vehicle routing problem for waste collection with smart operational management approaches. *International Transactions in Operational Research, 32*(2), 863–887.
Kumar, A., & Agrawal, A. (2020). Recent trends in solid waste management status, challenges, and potential for the future Indian cities – a review. *Current Research in Environmental Sustainability, 2*, Article 100011.
Letelier, C., Blazquez, C., & Paredes-Belmar, G. (2022). Solving the bin location–allocation problem for household and recycle waste generated in the commune of Renca in Santiago. *Chile. Waste Management and Research, 40*(2), 154–164.
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*(1), 441–465.
Markov, I., Bierlaire, M., Cordeau, J.-F., Maknoon, Y., & Varone, S. (2020). Waste collection inventory routing with non-stationary stochastic demands. *Computers and Operations Research, 113*, Article 104798.
Martí, R., Resende, M. G. C., & Ribeiro, C. C. (2013). Multi-start methods for combinatorial optimization. *European Journal of Operational Research, 226*(1), 1–8.
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.
Olmez, O. B., 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.
Prodhon, C., & Prins, C. (2014). A survey of recent research on location-routing problems. *European Journal of Operational Research, 238*(1), 1–17.
Ravindra, K., & Mor, S. (2019). E-waste generation and management practices in Chandigarh, India and economic evaluation for sustainable recycling. *Journal of Cleaner Production, 221*, 286–294.
Rossit, D., & Bard, J. (2025). Solving the waste bin location problem with uncertain waste generation rate: A bi-objective robust optimization approach. *Waste Management and Research, 43*(3), 421–437.
Rossit, D. G., & Nesmachnow, S. (2022). Waste bins location problem: A review of recent advances in the storage stage of the Municipal Solid Waste reverse logistic chain. *Journal of Cleaner Production, 342*, Article 130793.
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.
Roy, A., Manna, A., Kim, J., & Móon, I. (2021). Integrated planning of IoT-based smart bin allocation and vehicle routing in solid waste management. *In IFIP Advances in Information and Communication Technology*.
Roy, A., Manna, A., Kim, J., & Móon, I. (2022). IoT-based smart bin allocation and vehicle routing in solid waste management: A case study in South Korea. *Computers and Industrial Engineering, 171*.
Silver, E. A., Pyke, D. F., & Thomas, D. J. (2017). *Inventory and production management in supply chains*. Boca Raton, FL: CRC Press.
Singh, A. (2019). Managing the uncertainty problems of municipal solid waste disposal. *Journal of Environmental Management, 240*, 259–265.
Sivanandam, S. N., & Deepa, S. N. (2008). *Introduction to genetic algorithms*. Berlin, Germany: Springer-Verlag.
Spinelli, A., Maggioni, F., Ramos, T. R. P., Barbosa-Póvoa, A. P., & Vigo, D. (2025). A rolling horizon heuristic approach for a multi-stage stochastic waste collection problem. *European Journal of Operational Research, 323*(1), 276–296.
Tavares, G., Zsigraiova, Z., Semiao, V., & Carvalho, M. G. (2009). Optimisation of MSW collection routes for minimum fuel consumption using 3D GIS modelling. *Waste Management, 29*(3), 1176–1185.