Computers & Industrial Engineering 198 (2024) 110714

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

A location-sizing and routing model for a biomethane production chain fed by municipal waste

Anna Livia Croella [a,*], Luca Fraccascia [a,b]

[a] Department of Computer, Control, and Management Engineering "Antonio Ruberti", Sapienza University of Rome, Rome, Italy
[b] Department of High-Tech Business and Entrepreneurship, University of Twente, Enschede, Netherlands

ARTICLE INFO

Keywords:
Biomethane
Waste management
Circular economy
Location-routing
Mixed integer linear programming

ABSTRACT

This paper proposes an integrated approach for a biomethane supply chain from Organic Fraction of Municipal Solid Waste (OFMSW), addressing both strategic plant location-sizing and tactical vehicle routing. A two-stage iterative approach following a cluster-first location&sizing&route-second approach is implemented to optimize investment and transportation costs. A novel classification of waste producers into small and large categories allows for a tailored approach, minimizing transportation costs through waste pooling for small producers and allocating vehicles effectively for large producers. Extensive computational experiments on a case study on the Lazio Region demonstrate the cost-effectiveness of the proposed approach, which saves 30% of transportation costs with respect to a location-sizing model. Additionally, the sensitivity analysis highlights the significant impact of including huge-scale plants in the solution.

1. Introduction

The recent conflict between Russia and Ukraine has highlighted the vulnerability of European energy security, especially in countries like Italy, where natural gas constitutes 40% of the energy mix, and over 90% of this demand is met through imports (De Nicolò, Fraccascia, & Pontrandolfo, 2024; Eurostat, 2022). In response, the European Union (EU) has intensified efforts to reduce reliance on fossil fuels and transitiontowards renewable energy sources, as outlined in the European Commission's REPowerEU Plan (European Commission, 2022). Among the renewable energy alternatives, biomethane production from the OFMSW presents a promising solution. Biomethane, a renewable gas produced by Anaerobic Digestion (AD) of organic wastes, offers dual benefits: it enhances national energy security and contributes to sustainable waste management (Thrän et al., 2023). The ample availability of OFMSW throughout the year, coupled with accelerated generation rates resulting from urbanization and population growth (Nations, 2022), make it an invaluable resource for bio-fuel supply chains (Quddus, Chowdhury, Marufuzzaman, Yu, & Bian, 2018).

The existing studies regarding biomethane production from OFMSW primarily concentrate on technical aspects, such as process improvement (Yan et al., 2016; Zhu, Hsueh, & He, 2011), technical-economic feasibility (Ishaq & Ishaq, 2022; Wang, Chai, Shao, & Qian, 2021), and environmental benefits (de Jesús Vargas-Soplín, Prochnow, Herrmann, Tscheuschner, & Kreidenweis, 2022; Gross et al., 2021; Starr, Gabarrell, Villalba, Talens Peiro, & Lombardi, 2014). Nevertheless, to the best of our knowledge, there is a lack of comprehensive research addressing the design and optimization of its entire production chain—from determining the number, size, and location of plants to optimizing waste collection and transportation logistics (Atashbar, Labadie, & Prins, 2018). This gap underlines the need for models that can guide policymakers in making strategic decisions for establishing a resilient and efficient biomethane production network. In this regard, despite the significant potential for utilizing OFMSW in biomethane production, Italy currently processes only 5.1% of the available waste and operates a limited number of plants (ISPRA, 2022). Therefore, addressing this inefficiency by developing a more effective biomethane production chain from OFMSW is essential to optimize resource utilization and expand production capacity.

This paper addresses the above-mentioned gaps by proposing an integrated location-routing model that optimizes both the strategic and tactical dimensions of the biomethane production chain from OFMSW. Specifically, the model *(i)* identifies the optimal locations and sizes of biomethane production plants, *(ii)* efficiently allocates waste, and *(iii)* devises optimal collection routes, particularly for smaller producers who often underutilize truck capacity. Given the complexity of this problem, which involves numerous variables and constraints, exact optimization methods may be computationally infeasible for large-scale, real-world scenarios. Therefore, we employ heuristic techniques that balance computational efficiency with solution quality, offering near-optimal solutions that are practically viable.

* Corresponding author.
E-mail addresses: annalivia.croella@uniroma1.it (A.L. Croella), luca.fraccascia@uniroma1.it (L. Fraccascia).

https://doi.org/10.1016/j.cie.2024.110714
Received 12 May 2024; Received in revised form 9 September 2024; Accepted 6 November 2024
Available online 14 November 2024
0360-8352/

A two-stage iterative algorithm is proposed, following a cluster-first location& sizing&route-second approach. In the first stage, a modified implementation of the k-means clustering heuristic is employed to group small waste producers into clusters for pooling. In the second stage, the focus shifts to optimizing the locations and capacities of biomethane production plants, considering the possibility of upgrading existing plants to support the AD process for biomethane production. Waste is allocated to the designated plants, with large producers—those capable of filling truck capacities—assigned directly, while efficient vehicle routes are designed for waste collection from clustered producers. The optimization problem is addressed through a Mixed Integer Linear Programming (MILP) model based on a two-commodity flow formulation, tailored to manage the complexities of a multi-depot, multi-vehicle routing problem with integrated location-sizing features. The model's effectiveness is validated using a case study of the Lazio Region in Italy. As a decision-support tool, the model offers policymakers a comprehensive framework for designing Waste-to-Energy strategies. Indeed, by integrating strategic and tactical techno-economic considerations, this approach can guide regions in developing waste management solutions that are closely aligned with their unique challenges and objectives.

The remainder of this paper is organized as follows: Section 2 presents the theoretical background and reviews the existing literature, Section 3 outlines the problem formulation, and Section 4 applies the proposed model to a case study, detailing and analyzing the computational results. Finally, Section 5 concludes the paper by developing theoretical implications and future research directions.

2. Literature review

The literature on supply chain management for bioenergy production encompasses various aspects, with a primary focus on supply chain planning and design. Key activities, often treated separately in the literature, include determining optimal locations for feedstock hubs and biogas plants during the strategic-tactical design phase and optimizing transportation networks between these locations during the operational design phase (Sarker, Wu, & Paudel, 2019). This review provides an overview of significant research efforts within the broad field of biomass energy. It then narrows its focus to biomethane production and its diverse feedstock sources, with a particular emphasis on OFMSW, highlighting the need for further research into integrating logistical elements to enhance both the accuracy and practicality of biomethane production models.

A considerable body of literature addresses plant location and sizing for small-scale biomass design networks (Hong, Shen, & Lam, 2016; Iakovou, Karagiannidis, Vlachos, Toka, & Malamakis, 2010). Frequently, the biogas localization problem is approached using Geographical Information System (GIS) methods combined with multicriteria techniques (Jesus, Souza, Puglieri, Piekarski, & Francisco, 2021). For instance, spatial GIS-based analyses have been utilized by Valenti, Porto, Dale, and Liao (2018) and Zubaryeva, Zaccarelli, Giudice, and Zurlini (2012) to identify optimal locations and sizes for biogas plants in Italy, with a focus on the Lecce and Catania provinces, respectively. Additionally, over the past two decades, MILP models have been widely employed to address strategic issues in various bioenergy supply chains (Jesus et al., 2021; Ng, How, Lim, Ngan, & Lam, 2022). These models typically focus on bioethanol and biofuel production plants, biomass cultivation, and storage centers, using candidate locations and feedstock amounts as inputs (Akgul, Shah, & Papageorgiou, 2012; Balaman & Selim, 2014; Chen & Önal, 2014; Zhang et al., 2016). The sources of feedstock in these studies vary, including corn and stover (Giarola & Bezzo, 2011; Ng & Maravelias, 2017), straw, animal waste (Park, 2019), lignocellulosic biomass (Osmani & Zhang, 2014), algae, and microalgae (Ahn, Lee, Lee, & Han, 2015; Arabi, Yaghoubi, & Tajik, 2019a, 2019b; Asadi, Habibi, Nickel, & Sahebi, 2018), among others. The objectives of these MILP models are often defined by single-objective economic functions (Ng et al., 2022). However, many studies adopt a multi-objective approach, incorporating techniques such as Pareto optimization, principal component analysis, and fuzzy methods (Franco, Bojesen, Hougaard, & Nielsen, 2015; Ghaderi, Moini, & Pishvaee, 2018; How & Lam, 2018; Silva, Alçada-Almeida, & Dias, 2017). These models aim to balance financial goals — such as minimizing capital investment and operational costs — while also reducing environmental impacts like greenhouse gas emissions.

Although there is extensive literature on biomass and biogas supply chains, studies specifically focused on biomethane are relatively limited (An, Wilhelm, & Searcy, 2011; Ba, Prins, & Prodhon, 2016; Sharma, Ingalls, Jones, & Khanchi, 2013). For example, Hoo, Hashim, Ho, and Yunus (2019) use a techno-economic MILP model to study the profitability of biomethane plants in Southern Malaysia, optimizing grid transportation costs for biomethane derived from various biogas sources. O'Shea, Wall, Kilgallon, and Murphy (2016) examine the impact of incentives and scale on the location and sizing of biomethane facilities, aiming to identify advantageous grid injection points to enhance profitability. Jensen, Münster, and Pisinger (2017) propose a Mixed Integer Programming (MIP) for the network flow problem, optimizing internal product flows and investment plans in a biogas supply chain, with computational experiments conducted in Denmark. Similarly, Lyng, Bjerkestrand, Stensgård, Callewaert, and Hanssen (2018) highlight the profitability of AD at the regional level in Norway, emphasizing the reduction of greenhouse gas emissions.

While biomethane research primarily focuses on agricultural feedstocks, the potential of OFMSW management remains underexplored (Asefi, Shahparvari, & Chhetri, 2020). A recent study by Fraccascia, Spagnoli, Riccini, and Nastasi (2021) proposed a methodology to design the biomethane production chain from OFMSW at the regional level, encompassing waste allocation to production plants and assessing environmental and economic performance. However, the approach has significant limitations: it lacks the optimization of plant number, size, and location, requiring these parameters to be predefined, and it overlooks the optimization of truck routing, potentially missing operational efficiencies. Reviews by Ghiani, Laganà, Manni, Musmanno, and Vigo (2014) and Morrissey and Browne (2004) emphasize the need for greater attention to social acceptability within this framework. Despite these challenges, successful examples from European cities demonstrate the benefits of converting OFMSW into energy through well-implemented biomass supply chains (Eyl-Mazzega et al., 2019; Vrabie, 2021).

While there has been extensive research on bioenergy supply chains, few studies focus on the tactical and operational levels, such as feedstock allocation, transportation, and storage, which are crucial due to the high transportation costs. Lautala et al. (2015) emphasize integrating logistical aspects for a realistic biofuel supply chain model. Notable efforts include Bojesen, Birkin, and Clarke (2014)'s two-step location–allocation model for the Danish biogas sector and Höhn, Lehtonen, Rasi, and Rintala (2014)'s GIS-based optimization of biogas plant locations in Finland. Strategic-tactical models by Ekşioğlu, Acharya, Leightley, and Arora (2009) and Lin, Rodríguez, Shastri, Hansen, and Ting (2014) highlight the impact of feedstock demand and inventory on supply chain design. Additionally, Ng and Maravelias (2017) and Yang et al. (2022) present MIP and MILP models for optimally siting and sizing biomass plants, addressing the challenges of large-scale biofuel supply chains.

In this framework, the integration of routing decisions adds another layer of complexity. As a result, heuristic approaches are often used to tackle this optimization challenge (see surveys Nagy & Salhi, 2007; Prodhon & Prins, 2014). Studies like Cao, Zongxi, and Zhou (2020) and Wu, Sarker, and Paudel (2015) address the Vehicle Routing Problem (VRP) in biofuel supply chains, employing two-stage and iterative approaches to manage the interdependencies between routing and other decision variables, such as location and capacity.

Asadi et al. (2018) introduce a bi-objective stochastic MILP for the location-inventory-routing of a biofuel supply chain fed by microalgae, optimizing decisions such as the number, locations, and inventory levels of plants, as well as the allocation of extraction sites to plants and their respective routes. Finally, Li, Zhao, Yang, and Guo (2023) explore the use of evolutionary algorithms, including genetic algorithms and ant colony optimization, to solve a multi-objective model for the location-routing problem in biomass waste collection.

This study contributes to this field by implementing a two-stage iterative algorithm that follows a cluster-first, location-sizing-route-second approach. In the initial stage, a constrained clustering algorithm is applied (for an in-depth review of clustering methods, see Dao, Duong, & Vrain, 2017; Höppner & Klawonn, 2008). The second stage extends a two-commodity flow model (Baldacci, Hadjiconstantinou, & Mingozzi, 2004; Ramos, Gomes, & Barbosa-Povoa, 2019) by incorporating aspects such as location, sizing, and timing to achieve a more realistic representation of biomethane production scenarios. The model is designed to handle medium-scale instances, addressing a notable gap in the current biomethane literature.

Table 1 offers a summary of the key aspects and contributions of the cited studies. For detailed discussions on the solution methods, please refer to Sections 3.1.1 and 3.1.2.

3. Problem formulation and solution approach

The bmNDP focuses on optimizing the placement and size of biomethane production plants within a defined regional area. Potential sites for these plants include *(i)* new construction locations, *(ii)* existing biogas plants that can be upgraded to biomethane production facilities, and *(iii)* current plants already capable of processing organic waste into biomethane. The region consists of municipalities that generate a fixed daily amount of OFMSW, referred to as *producers*. Municipalities are categorized based on their waste generation into *small* or *large* producers.

Small producers' waste is aggregated until the vehicle reaches its maximum capacity, while waste from large producers is transported individually by one or more vehicles directly to a plant. Without loss of generality, the model considers two distinct vehicle capacities, each associated with a specific type of producer. Additionally, the model excludes internal routes for waste collection from road dumpsters within municipalities, focusing solely on macro-routing.

The main objectives of the bmNDP are:

- *(i)* determine the optimal locations, sizes, and types of all active plants;
- *(ii)* allocate the produced waste to the active plants effectively;
- *(iii)* plan the routes among small producers, ensuring each route starts and ends at an operational plant and serves a subset of small producers.

The goal is to minimize both investment and operating costs related to plant construction and transportation while adhering to various network constraints.

As an example, Fig. 1 illustrates the optimal solution for a bmNDP instance with four potential plant locations and twelve producers. Among these, three are large producers (indicated by cyan circles) generating fifteen, ten, and six units of waste, while the remaining nine are small producers, each generating one unit of waste (indicated by orange fill). The vehicles serving small and large producers have maximum capacities of four and seven units, respectively. In the figure, triangle and diamond symbols represent locations for plants that can be either upgraded or newly built, respectively, while the square location offers three potential plant sizes. The optimal solution is depicted with dashed-filled shapes denoting the active plants. At the square location, the smallest plant size is chosen.

To address the bmNDP, a two-stage iterative algorithm is proposed, utilizing a cluster-first location&sizing&route-second approach. In the initial stage, the algorithm assigns producers to vehicles by grouping small producers into n independent clusters. This is achieved using a K-means clustering heuristic with specific constraints, referred to as Bounded K-means (BKm) clustering. Details of this algorithm are provided in Section 3.1.1. Once the clustering is completed, a MILP model for the Location-Sizing Routing Problem with Fixed Clusters (LSRP-FC) is solved. This model extends the two-commodity flow formulation proposed by Ramos et al. (2019) for the multi-depot VRP with heterogeneous vehicle fleets and maximum travel times, by incorporating plant location and sizing aspects. The MILP model for the LSRP-FC is detailed in Section 3.1.2.

A flowchart illustrating the solving process is presented in Fig. 2.

3.1. Mathematical models

3.1.1. A bounded K-means algorithm

In the initial phase of the algorithm, the focus is on clustering small producers based on their proximity. Among the various clustering methods available, such as hierarchical clustering and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), the *k-means algorithm* is particularly notable for its simplicity and effectiveness (MacQueen, 1967). The k-means algorithm operates iteratively to minimize the sum of squared distances within each cluster for a fixed number k of clusters. The process involves three steps: *(i) Initialization Step*: k centroids, or cluster centers, are identified; *(ii) Assignment Step*: each data point is assigned to the nearest centroid based on distance; and *(iii) Update Step*: the centroids are recalculated based on the data points assigned to each cluster. The algorithm iterates through these steps until a stopping criterion is met.

While k-means clustering is widely used, various modifications and enhancements, in design and/or implementation, have been proposed to address specific needs and constraints (Ikotun, Ezugwu, Abualigah, Abuhaija, & Heming, 2023; Pérez-Ortega et al., 2019). In the last decade, Constrained Clustering has also gained attention for its ability to incorporate additional requirements into the clustering process, aligning results more closely with practical or domain-specific needs (Dao et al., 2017). The use of different kinds of constraints falls under the name of Constrained Clustering. These constraints can include limitations on cluster cardinality, such as uniform clustering, which aims to ensure that all clusters contain approximately the same number of data points (Genevay, Dulac-Arnold, & Vert, 2019; Höppner & Klawonn, 2008), and fair clustering, which seeks to balance the distribution of protected subgroups within each cluster (Chierichetti, Kumar, Lattanzi, & Vassilvitskii, 2018). Constrained Clustering can also address spatial or geometric properties of clusters and incorporate specific must-link and cannot-link requirements (Piccialli, Russo Russo, & Sudoso, 2022).

To address the practical requirements of the bmNDP, an implementation variant of the k-means algorithm, referred to as BKm algorithm, is employed. This variant integrates cardinality and geometric constraints into the clustering process. Moreover, the BKm algorithm uses a binary linear programming model during the *Assignment Step* to optimize the clustering outcome. The primary objective is to minimize the distances within each cluster from their centroids while adhering to the following constraints:

- *(i)* each data point must be assigned to exactly one cluster;
- *(ii)* each chosen centroid must belong to the corresponding cluster;
- *(iii)* each cluster must be populated and adhere to a maximum cardinality (*capacity constraints*);
- *(iv)* the total weight of each cluster must not exceed a maximum value (*geometric constraints*).

The algorithm proceeds until the maximum number of iterations is reached or converges to a local optimum. Since there is a higher likelihood of converging to local minima when using k-mean based

Table 1
Summary of related literature in bioenergy supply chain management.

Reference	Biomass feedstock	Energy/ fuel produced	Optimal decisions			Methodology	Solution method	Case study
			Facility location	Location/ allocation	Routing			
Ahn et al. (2015)	Algae and microalgae	Biofuel		×		MILP	Commercial solver	Korea
Akgul et al. (2012)	Wheat and wheat straw	Biofuel		×		MILP	Commercial solver	United Kingdom
Arabi et al. (2019a)	Algae	Biofuel		×		Multi-objective MIQP	Commercial solver	Iran
Arabi et al. (2019b)	Microalgae	Biobutanol		×		MILP	Data envelopment analysis	Iran
Asadi et al. (2018)	Microalgae	Biofuel		×	×	Bi-objective MILP	Metaheuristic algorithms	Iran
Balaman and Selim (2014)	Corn and animal manure	Biogas and electricity		×		MILP	Commercial solver	Izmir, Turkey
Bojesen et al. (2014)	Animal manure	Biogas	×			GIS-based analysis	Scenario analysis	Denmark
Cao et al. (2020)	Biomass Feedstock	Gas, heat and electricity		×	×	MIP	Hybrid tabu search algorithm	
Chen and Önal (2014)	Agricoltural waste and animal manure	Biofuel		×		MINLP	Backward-recursive heuristic	United States
Ekşioğlu et al. (2009)	Agricoltural waste	Biofuel		×		MIP	Commercial solver	Mississippi
Fraccascia et al. (2021)	OFMSW	Biomethane		×		Optimization model	Scenario analysis and commercial solver	Rome, Italy
Franco et al. (2015)	Agricoltural waste and animal manure	Biogas	×			Multicriteria GIS-based analysis	Fuzzy weighted overlap dominance	Ringkøbing-Skjern, Denmark
Ghaderi et al. (2018)	Switchgrass	Bioethanol		×		Multi-objective MILP	Possibilistic-fuzzy approach and commercial solver	Iran
Giarola and Bezzo (2011)	Corn and stover	Bioethanol	×			Multi-objective MILP	ϵ-constraint and commercial solver	Northern Italy
Hoo et al. (2019)	Agricoltural waste and animal manure	Biomethane		×		MILP	Commercial solver	Johor, Malaysia
How and Lam (2018)	Agricoltural food waste	Gas, heat and electricity		×		Multi-objective optimization model	Principal Component Aided approach	Johor, Malaysia
Höhn et al. (2014)	Agricoltural waste and animal manure	Biomethane	×			GIS-based analysis	Kernel Density maps	Southern Finland
Jensen et al. (2017)	Agricoltural waste and animal manure	Gas, heat and electricity	×			MIP	Commercial solver	North-West of Denmark
Li et al. (2023)	Biomass waste	Biogas		×	×	Multi-objective IP	Two evolutionary algorithms	
Lin et al. (2014)	Agricoltural waste	Biofuel		×		MILP	Commercial solver	Illinois
Lyng et al. (2018)	Animal manure	Biomethane		×		Optimization model	Scenario Analysis	Vestfold County, Norway
Ng and Maravelias (2017)	Corn and stover, switchgrass	Biofuel		×		MILP	Decomposition approach and Commercial solver	South Central Wisconsin
O'Shea et al. (2016)	OFMSW	Biomethane		×		Optimization model	Greedy algorithm	Ireland
Osmani and Zhang (2014)	Lignocellulosic biomass	Biofuel	×			Stochastic MILP	Sample average approximation algorithm	4-state in United States
Park (2019)	Animal manure	Biogas and biomethane		×		Multi-objective MILP	Commercial solver	North Dakota
Quddus et al. (2018)	Forest residues, corn stover and OFMSW	Biofuel		×		Stochastic MILP	Sample average approximation algorithm	Mississippi
Sarker et al. (2019)	Agricoltural waste	Biomethane		×		Non convex MINLP	Genetic algorithm-based heuristic	
Silva et al. (2017)	Animal waste	Biogas		×		Multi-objective MILP	Pareto approach and commercial solver	Entre-Douro-e-Minho, Portugal
Valenti et al. (2018)	OFMSW, agricoltural waste and animal manure	Biogas	×			GIS-based analysis	Scenario Analysis	Catania, Italy

(*continued on next page*)

Table 1 (*continued*).

Wu et al. (2015)	Forest residues, livestock manure and grass	Biomethane		×		MINLP	Heuristic approach	
Yang et al. (2022)	OFMSW, agricoltural waste and animal manure	Biogas		×		MISOCP	Commercial solver	Longshan County, China
Zhang et al. (2016)	Biomass feedstock	Biofuel		×		GIS simulations/optimization	Simulation/Exact algorithm	Michigan's Lower Peninsula
Zubaryeva et al. (2012)	OFMSW, agricoltural waste and animal manure	Biogas	×			Multicriteria GIS-based analysis	Scenario Analysis	Lecce, Italy
Our study	OFMSW	Biomethane		×	×	MILP	Two-stage iterative algorithm	Lazio, Italy

MIQP: Mixed Integer Quadratic Programming.
IP: Integer Programming.
MINLP: Mixed Integer Non-Linear Programming.
MISOCP: Mixed Integer Second Order Cone Programming.

algorithm (Peña, Lozano, & Larrañaga, 1999), the algorithm is executed multiple times with different (random) initial centroids.

The pseudo-code of the proposed procedure is presented below. For a complete list of notation, the reader is referred to Table 2.

Bounded K-means algorithm pseudo-code

Input: $\underline{\mathcal{M}} \subseteq \mathcal{M}$, $\mathcal{D} = (d_{m,m'}) \in \mathbb{R}^{|\underline{\mathcal{M}}|\times|\underline{\mathcal{M}}|}$, $q \in \mathbb{R}^{|\underline{\mathcal{M}}|}$, $\mathbb{Z} \ni n, s \geq 1$, $l \in \mathbb{R}_+$, $\bar{\imath}$

Output: $C = \{C_1, \ldots, C_n\}$

Initialization Step

Choose randomly n producers $\{m_1, \ldots, m_n\} \subseteq \underline{\mathcal{M}}$.
Set $\mu_c^0 := m_c$ and $C_c := \{\mu_c^0\}$ $\forall\, c = 1, \ldots, n$.

Stopping Criteria

If ($\mu_c^{i+1} = \mu_c^i$ $\forall\, c = 1, \ldots, n$) $\vee$ ($i = \bar{\imath}$) **STOP** Else $i := i + 1$

i-th iteration

Assignment Step

Assign each $m \in \underline{\mathcal{M}}$ to a cluster by solving:

$$\min_{\varsigma} \sum_{c=1}^{n} \sum_{m\in\underline{\mathcal{M}}} \varsigma_{mc} \cdot d_{m\mu_c^i}$$

$$\begin{aligned}
s.t \quad & \sum_{c=1}^{n} \varsigma_{cm} = 1 && \forall\, m \in \underline{\mathcal{M}} && (i)\\
& \varsigma_{c\mu_c^i} = 1 && \forall\, c = 1, \ldots, n && (ii)\\
& 1 \leq \sum_{m\in\underline{\mathcal{M}}} \varsigma_{cm} \leq s && \forall\, c = 1, \ldots, n && (iii)\\
& \sum_{m\in\underline{\mathcal{M}}} q_{m'} \cdot \varsigma_{cm} \leq l && \forall\, c = 1, \ldots, n && (iv)\\
& \varsigma_{cm} \in \{0, 1\} && \forall\, c = 1, \ldots, n,\ m \in \underline{\mathcal{M}}
\end{aligned}$$

Update Step

Recompute cluster centroids:

$$\mu_c^{i+1} := \arg\min_{m\in C_c} \sum_{m'\in C_c, m\neq m'} d_{mm'} \quad \forall\, c = 1, \ldots, n$$

It is important to note that during the *Assignment Step* in the very first iteration, no solution might exist when considering certain sets of centroids. In such cases, the *Initialization Step* is repeated until a feasible assignment is found, if any.

3.1.2. A MILP model for the LSRP-FC

This section presents a formulation for the LSRP-FC problem. The problem consists of three main components: *(i)* the facility location-sizing problem, which aims to determine the optimal locations, sizes, and types of facilities based on system constraints; *(ii)* the allocation problem, which assigns large producers to plants; and *(iii)* the VRP, which focuses on finding the most efficient routes to supply the cluster of small producers.

The formulation is based on the idea of the two-commodity network flow, initially proposed by Baldacci et al. (2004) for the VRP with a single depot and a homogeneous fleet of capacitated vehicles serving a set of customers. This approach uses a depot copy and flow continuous variables y_{ij} to model two different "commodity flows". The first flow goes from the *real* depot to the *copy* one and represents the vehicle load, while the second flow goes in the opposite direction and indicates the empty space in the vehicle. Since both the amount transported and the empty space are bounded by the vehicle capacity l, the sum of the two flows is always equal to the capacity ($y_{ij} + y_{ji} = l$). An additional set of binary variables x_{ij} is used to specify which flows belong to the optimal solution. This two-commodity flow formulation enabled the authors to solve larger instances compared to the benchmark algorithm (the *two-index formulation* by Laporte, Nobert, & Desrochers, 1985), and it provided a new, tighter lower bound.

More recently, Ramos et al. (2019) extended the application of the two-commodity flow approach to the Multi-Depot VRP for solving larger and more realistic instances. In their formulation, they considered multiple depots and various types of vehicles with different capacities (heterogeneous vehicle fleet). The approach demonstrated superior performance in terms of solution quality and computational efficiency compared to the traditional formulation, such as the *three-index formulation* proposed by Golden, Magnanti, and Nguyen (1977). A survey of formulations for the capacitated VRP can be found at Letchford and Salazar González (2006).

In our model, we further expand the decision variables space to incorporate the location, sizing, and type of the depots (biomethane production plants) and the allocation of some customers (large waste producers) to vehicles. The routing decisions are confined to the small waste producers, which are grouped into clusters during the first step of the algorithm.

Notation. In the following, the set $\mathcal{F}'$ denotes all locations of real and copy plants, while the set $\mathcal{N}_C$ consists of all pairs of (real and copy) plant locations and small producers of cluster C, such that no pair consists of two plants from the set $\mathcal{F}'$, $\mathcal{N}_C := (C \cup \mathcal{F}') \times (C \cup \mathcal{F}') - (\mathcal{F}' \times \mathcal{F}')\ \forall C \in \mathcal{C}$. For detailed information on the notation used, consult Table 2.

Variables. We denote as $\chi = (\lambda, \gamma, w, x, y)$ the decision variables vector. Two set of binary variables λ and γ are used to model the location-sizing problem:

$$\lambda_p^k = \begin{cases} 1 & \text{if a plant is built at location } p \\ & \text{with capacity level } k \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, p \in \mathcal{P},\ k \in \mathcal{K}$$

$$\gamma_u = \begin{cases} 1 & \text{if an existing plant is upgraded at location } u \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, u \in \mathcal{U}$$

Two sets of binary variables w and x are used to assign large and small producers to plants:

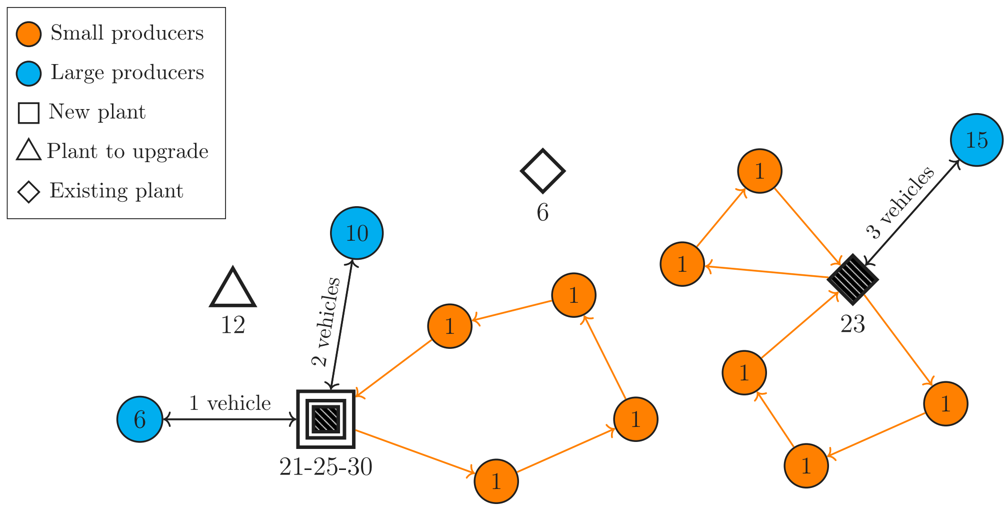


Fig. 1. Example of solution for a bmNDP istance. Vehicles serving small and large producers (waste units in labels) have a maximum capacity of four and seven units, respectively. Active plants are filled with dashed lines.

$$w_{mf} = \begin{cases} 1 & \text{if waste from large producers } m \\ & \text{is assigned to facility } f \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, m \in \overline{\mathcal{M}},\ f \in \mathcal{F}$$

$$x_{mm'} = \begin{cases} 1 & \text{if there is a flow from } m \text{ to } m' \\ 0 & \text{otherwise} \end{cases} \qquad \forall\,(m,m') \in \mathcal{N}_C,\ C \in \mathcal{C}$$

A set of continuous positive variables $y_{mm'}$, with $(m,m') \in \mathcal{N}_C,\ C \in \mathcal{C}$, represents the flow from plant to small producers and back.

Fig. 3 depicts the two-commodity flows for a cluster C containing four small producers, with a total waste of $l_C = 10$. Each small producer in the cluster is represented by a node, labeled with the amount of waste it generates. The vehicle's load flow (orange path) is shown as it departs from a real plant, visits all the nodes in the cluster, and arrives at the copy plant. Meanwhile, the vehicle's empty space flow (black dashed path) moves in the opposite direction.

Constraints. The following sets of constraints are related to the facility location problem:

$$\sum_{k\in\mathcal{K}} \lambda_p^k \le 1 \qquad \forall\, p \in \mathcal{P} \tag{1}$$

$$\sum_{u\in\mathcal{U}} \gamma_u \le \bar{u} \tag{2}$$

$$\sum_{k\in\mathcal{K}} \lambda_p^k + \sum_{k'\in\mathcal{K}} \lambda_{p'}^{k'} \le 1 \qquad \forall\,(p,p') \in \Lambda \tag{3a}$$

$$\gamma_u + \gamma_{u'} \le 1 \qquad \forall\,(u,u') \in \Gamma \tag{3b}$$

$$\sum_{k\in\mathcal{K}} \lambda_p^k + \gamma_u \le 1 \qquad \forall\,(p,u) \in \Omega \tag{3c}$$

Constraint (1) ensures that at most one new plant can be built at each candidate location newly opened plants. Inequality (2) sets the maximum number of plants that can be upgraded. Then, constraints (3a), (3b), and (3c) dictate that at most one plant is built (or upgraded) for each pair of incompatible locations.

Constraints (4a), (4b), and (4c) address the sizing aspect of the problem. Specifically, they ensure that the total waste from small and large producers assigned to each newly opened, upgraded, and existing plant does not exceed plant capacity.

$$\sum_{c\in\mathcal{C}} \sum_{m\in C} y_{pm} + \sum_{m\in\overline{\mathcal{M}}} q_m \cdot w_{mp} \le \sum_{k\in\mathcal{K}} k \cdot \lambda_p^k \qquad \forall\, p \in \mathcal{P} \tag{4a}$$

$$\sum_{c\in\mathcal{C}} \sum_{m\in C} y_{um} + \sum_{m\in\overline{\mathcal{M}}} q_m \cdot w_{mu} \le k_u \cdot \gamma_u \qquad \forall\, u \in \mathcal{U} \tag{4b}$$

$$\sum_{c\in\mathcal{C}} \sum_{m\in C} y_{em} + \sum_{m\in\overline{\mathcal{M}}} q_m \cdot w_{me} \le k_e \qquad \forall\, e \in \mathcal{E} \tag{4c}$$

The following constraints concern the large producers' waste allocation.

$$\sum_{f\in\mathcal{F}} w_{mf} = 1 \qquad \forall\, m \in \overline{\mathcal{M}} \tag{5}$$

$$\sum_{f\in\mathcal{F}} 2w_{mf} \cdot t_{mf} \le \bar{t} \qquad \forall\, m \in \overline{\mathcal{M}} \tag{6}$$

Constraints (5) assign the waste from each large producer to one and only one real plant, respectively, while (6) guarantee that the travel time for the disposal round trip does not exceed the maximum vehicle travel time.

The feasible cluster routes for small producers are instead defined by the following constraints.

$$\sum_{m'\in C\cup\mathcal{F}'} x_{m'm} = 2 \qquad \forall\, m \in C,\ C \in \mathcal{C} \tag{7}$$

$$x_{fm} = x_{mf} \qquad \forall\, f \in \mathcal{F}',\ m \in \underline{\mathcal{M}} \tag{8}$$

$$\sum_{m\in C} x_{mf} = \sum_{m\in C} x_{\tilde{f}m} \qquad \forall f \in \mathcal{F},\ C \in \mathcal{C} \tag{9}$$

$$x_{mm} = 0 \qquad \forall\, m \in \underline{\mathcal{M}} \tag{10}$$

In particular, constraints (7) ensure that each small producer in a cluster is connected to two different sites, which can be either another producer, a real plant, or a copy plant. This reflects the existence of two *commodity* flows, one from a real plant to its associated copy plant and one in the opposite direction. Constraints (8) requires that for every (real and copy) plant, the inflow must equal the outflow, while (9) ensures that, for each real plant f, the inflow is equal to the outflow from its copy plant $\tilde{f}$. Finally, constraints (10) prevent loops in the flow involving small producers.

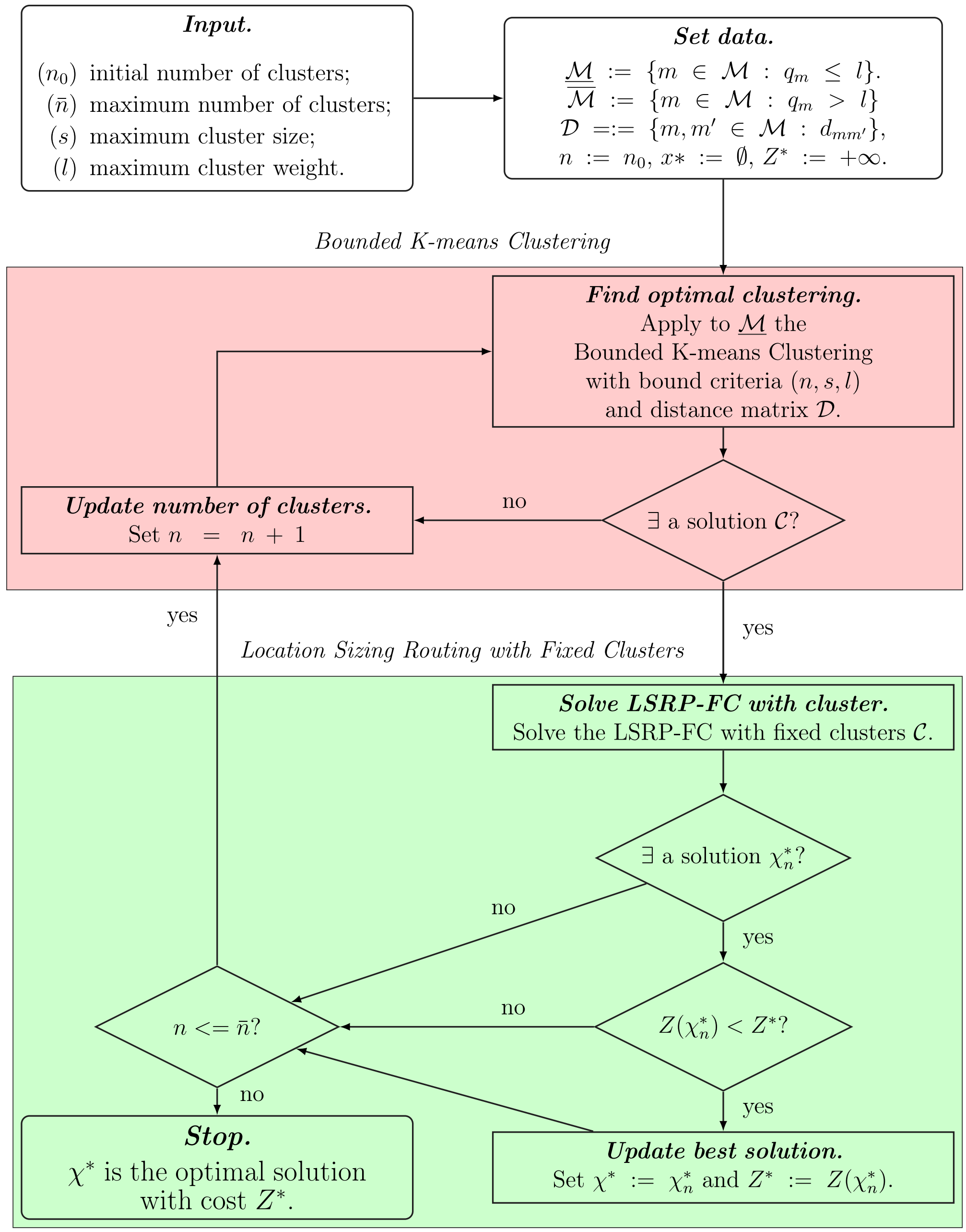


Fig. 2. Generalized flowchart of the iterative algorithm solving the bmNDP.

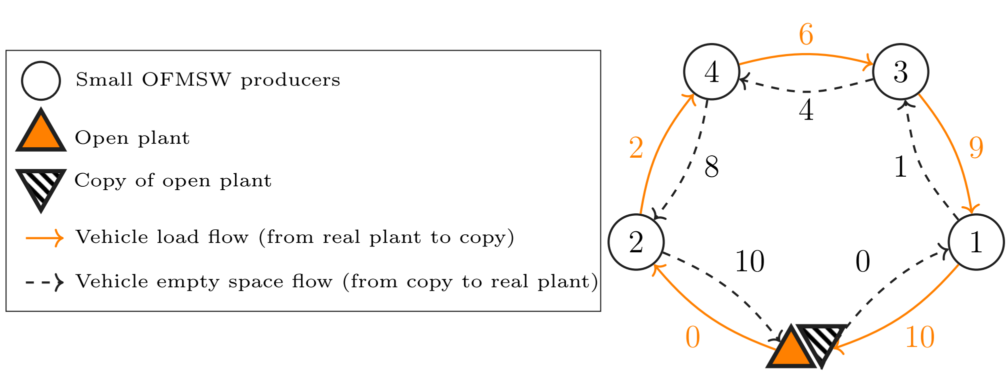


Fig. 3. Two-commodity flow representation for a cluster of four small producers with a total load equal to ten units. The values for the vehicle load and empty space are reported on solid and dashed edges, respectively.

Table 2
List of sets and parameters.

Sets	Description
Λ	Set of incompatible locations between new plants ($\Lambda \subseteq \mathcal{P} \times \mathcal{P}$)
Γ	Set of incompatible locations between upgraded plants ($\Gamma \subseteq \mathcal{P} \times \mathcal{R}$)
Ω	Set of incompatible locations for new and upgraded plants ($\Omega \subseteq \mathcal{P} \times \mathcal{R}$)
$\mathcal{C}$	Set of clusters for small OFMSW producers
$\mathcal{E}$	Set of locations of existing biomethane production plants
$\mathcal{F}$	Set of all potential biomethane production plants locations ($\mathcal{F} := \mathcal{P} \cup \mathcal{U} \cup \mathcal{E}$)
$\tilde{\mathcal{F}}$	Set of copy of all potential biomethane production plants locations
$\mathcal{F}'$	Set of all real and copy potential biomethane production plants locations
$\mathcal{N}_C$	Set of all pairs of plant locations and small producers of cluster C
$\mathcal{K}$	Set of new biomethane production plant capacity
$\mathcal{M}$	Set of all OFMSW producers
$\overline{\mathcal{M}}$	Set of OFMSW large producers
$\underline{\mathcal{M}}$	Set of OFMSW small producers
$\mathcal{P}$	Set of locations where a new biomethane production plant can be built
$\mathcal{U}$	Set of existing plant sites that can be upgraded to enable AD
Parameters	**Description**
β	Number of waste collections performed in a year
μ_c^i	Centroid of cluster c during iteration i of the BKm algorithm
$\bar{\imath}$	Number of iterations of the BKm algorithm
$\hat{\imath}$	Number of initialization of the starting centroids for the BKm algorithm
$\bar{a}$	Number of years of useful life of plants
CAP_k	Capital expenditures for plant capacity k
$d_{mm'}$	Distance from municipality m and m'
I_k	Investment cost for plant capacity k
k_f	Capacity of existing plant at locations f
L	Large vehicle capacity
l	Small vehicle capacity/Maximum cluster weight
l_C	Total OFMSW amount of cluster C
$\bar{n}$	Maximum number of clusters to test
n_0	Initial number of clusters to test
OP_f	Operating expenses for plant f
OP^{tr}	Operating expenses for transportation per vehicle
q_m	Quantity of OFMSW produced by municipality m
r	Discount rate
s	Maximum cluster size
$\bar{t}$	Maximum vehicle travel time
$t_{mm'}$	Vehicle Running time from municipality m and m'
$\bar{u}$	Maximum number of existing plants that can be upgraded to enable AD

As for the large producers, the waste from each small one has to be assigned to one and only one real plant (11) and the total travel time for collections and disposal must not be less or equal to the maximum vehicle travel time (12). In the final set of constraints, a factor of $\frac{1}{2}$ is introduced to account for the flow only in one direction.

$$\sum_{f\in\mathcal{F}}\sum_{m\in C} x_{mf} = 1 \qquad \forall\, C \in \mathcal{C} \tag{11}$$

$$\sum_{(m,m')\in\mathcal{N}_C} \frac{1}{2} x_{mm'} \cdot t_{mm'} \le \bar{t} \qquad \forall\, C \in \mathcal{C} \tag{12}$$

Finally, the following constraints model a feasible flow pattern (see also Fig. 3).

$$\sum_{m'\in C\cup\mathcal{F}'} (y_{mm'} - y_{m'm}) = 2q_m \qquad \forall\, m \in C,\ C \in \mathcal{C} \tag{13}$$

$$y_{m'm} + y_{mm'} = l_C \cdot x_{mm'} \qquad \forall\, (m,m') \in \mathcal{N}_C,\ C \in \mathcal{C} \tag{14}$$

$$\sum_{f\in\mathcal{F}'}\sum_{m\in\underline{\mathcal{M}}} y_{fm} = 0 \tag{15}$$

Flow Eqs. (13) ensure that the difference between the outflow and inflow for each small producer equals twice their waste amount. Additionally, Eqs. (14) ensure that the sum of the two opposite flows always matches the total waste of the cluster. Meanwhile, Eqs. (15) state that the outflows from all (real and copy) plants are set to 0. These last constraints reflect that the vehicles are fully utilized, leaving no remaining space after completing the cluster route.

Objective function. The objective function to be minimized is a linear function representing the sum of the investment costs:

$$Z_1(\chi) = \sum_{e\in\mathcal{E}} I_{k_e} + \sum_{p\in\mathcal{P}}\sum_{k\in\mathcal{K}} I_k \cdot \lambda_p^k + \sum_{u\in\mathcal{U}} I_{k_u} \cdot \gamma_u \tag{16}$$

and transportation operating costs:

$$Z_2(\chi) = OP^{tr}\left(\sum_{C\in\mathcal{C}}\sum_{(m,m')\in\mathcal{N}_C} \frac{1}{2} d_{mm'} \cdot x_{mm'} + \sum_{m\in\overline{\mathcal{M}}} 2\left\lceil \frac{q_m}{L} \right\rceil d_{mf} \cdot w_{mf}\right) \tag{17}$$

The term $Z_1(\chi)$ represents the investment costs associated with existing, upgraded, and newly opened plants, respectively. Investment costs for existing plants $\sum_{e\in\mathcal{E}} I_{k_e}$ are fixed and do not vary with the decision variables.

The term $Z_2(\chi)$ accounts for the transportation costs for both small and large producers. For each cluster, the actual distance traveled by the vehicle from a small producer to the next is counted once by using a $\frac{1}{2}$ factor. For large producers, the transportation cost is adjusted by a factor of 2 to account for the return journey and multiplied by the number of trips needed to handle the producer's waste.

4. Case study: Results and discussion

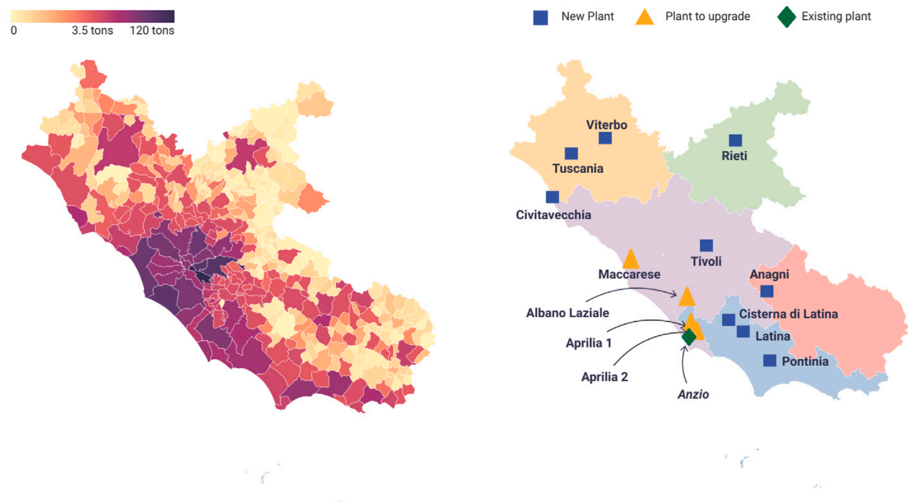
This section details the application of the proposed model to the Lazio Region, Italy's second most populous region, home to nearly six million residents across 378 municipalities, including the capital, Rome (ISTAT, 2022). Lazio generates around 600,000 tons of OFMSW annually, representing 21.73% of the total municipal solid waste produced in 2021 at the Italian level (ISPRA, 2022). Currently, the region has only one biomethane production plant utilizing OFMSW, with a portion of the organic waste used for compost production and approximately 250,000 tons transported to facilities outside the region each year (Fraccascia et al., 2021).

4.1. Data and assumptions

Producers. The data on OFMSW was obtained from the Italian open-source portal "Catasto Rifiuti Sezione Nazionale" (National Waste Database) of "Istituto Superiore per la Protezione e la Ricerca Ambientale" (ISPRA). The dataset, sourced from 2021, details the annual organic waste production in the region with specifics for each municipality. For this study, we assumed three waste collections per week over 52 weeks, scaling the data by a factor of $\beta = 156$. Due to missing data for 43 out of 378 producers, estimates were made based on similar-sized producers. Additionally, the municipalities of Ponza and Ventotene were excluded due to their autonomous waste management systems. Based on Fraccascia et al. (2021), a single producer was considered for each of Rome's 15 districts, with waste shares calculated according to population data from Dipartimento Trasformazione Digitale U.O. Statistica - Open Data (2023).

In the study, 390 waste producers were identified, with 265 classified as small producers and 125 as large producers. The processed data detailing the waste generated by these producers is available in the Supplementary Materials.

Fig. 4(a) presents a choropleth map illustrating the daily waste production across the Lazio region. Small producers, generating less than 3.5 tons of waste per day, are shaded in yellow, while large producers are depicted in purple.



(a) Choropleth map of the daily OFMSW amount produced in Lazio region.

(b) Locations map of new plants, plants that can be upgraded to enable AD and existing plants. Each colored area represents a province.

Fig. 4. Lazio region choropleth and location maps. Created by Datawrapper.

Plants. The criteria for selecting potential locations for new biomethane plants are as follows:

- *(i)* New plants must be constructed within the borders of the Lazio region.
- *(ii)* A municipality qualifies as a potential location for a new plant only if it has been officially designated by the regional government.

Through an extensive web search of past plant construction proposals within the region, nine areas have been identified as potential locations.

For each potential location, three capacity levels are considered for the biomethane production plants:

- *(i)* Small-scale plants with a capacity to process 510 tons per day, equivalent to 80,000 tons per year.[1]
- *(ii)* Medium-scale plants with a capacity to process 640 tons per day, equivalent to 100,000 tons per year[1].
- *(iii)* Large-scale plants with a capacity to process 960 tons per day, equivalent to 150,000 tons per year[1].

Additionally, four existing composting facilities in the region are eligible for upgrades to incorporate AD technology and a plant in Anzio is currently operational with a processing capacity of 250 tons per day. Fig. 4(b) illustrates all potential plant locations. It is important to note that certain restrictions apply: specifically, plants located within 15 km of each other are considered incompatible. This restriction affects the pairs of locations at Cisterna di Latina and Latina, as well as Aprilia 1 and Aprilia 2. Furthermore, a maximum of two existing plants can be upgraded.

The investment costs are calculated by combining the CAPEX and Operating Expenses (OPEX) according to the cost model outlined by Fraccascia et al. (2021). For a detailed explanation, refer to Eq. (6), Eq. (8), and Table 1 in Fraccascia et al. (2021). The capacity levels of the plants influence both cost structures.

For plants requiring upgrades, certain installation costs — including those associated with anaerobic digestion, biomethane upgrading, gas grid connection, and fueling stations — are deducted from the CAPEX. Specifically, the investment cost is set at 50% of the cost of establishing a new plant, in line with recommendations from Italian industry experts. In contrast, for existing plants, only operating costs are considered.

To account for the time value of money, all costs are annualized using a depreciation rate of 2.7% over ten years (Fraccascia et al., 2021). The specific cost functions used in the MILP model can be expressed as follows:

$$
\begin{aligned}
Z(\chi) = Z_1(\chi) + Z_2(\chi) = {} & \sum_{e\in\mathcal{E}} \sum_{a=1}^{\bar{a}} \frac{OP_{k_e}}{(1+r)^a} + \sum_{p\in\mathcal{P}} \sum_{k\in\mathcal{K}} \\
& \times \left(CAP_k + \sum_{a=1}^{\bar{a}} \frac{OP_k}{(1+r)^a} \right) \lambda_p^k \\
& + \sum_{u\in\mathcal{U}} \left(CAP_{k_u} + \sum_{a=1}^{\bar{a}} \frac{OP_{k_u}}{(1+r)^a} \right) \gamma_u \\
& + \sum_{a=1}^{\bar{a}} \frac{\beta \cdot OP^{tr}}{(1+r)^a} \left(\sum_{C\in\mathcal{C}} \sum_{(m,m')\in\mathcal{N}_C} \frac{1}{2} d_{mm'} \cdot x_{mm'} \right. \\
& \left. + \sum_{m\in\mathcal{M}} 2 \left\lceil \frac{q_m}{L} \right\rceil d_{mf} \cdot w_{mf} \right)
\end{aligned} \tag{18}
$$

Transport. The origin–destination distance and time matrices were computed using the Open Source Routing Machine (OSRM) (Fitzgerald, 2023), a routing engine based on data from OpenStreetMap (OpenStreetMap, 2023).

In the model, a heterogeneous fleet is assumed, comprising both small and large vehicles. This reflects practical considerations, as waste management companies in smaller municipalities often utilize smaller vehicles. The usable capacity for small vehicles is set at 3500 kg, which is used to define the subset of small waste producers. For large vehicles, the maximum usable capacity is established at 7500 kg (Ama Roma, 2023). Additionally, the maximum travel time for both small

[1] This calculation assumes three waste collections per week across 52 weeks per year ($\beta = 156$).

and large vehicles is set at 8 hours, aligning with a double work shift for operators. The transportation costs per vehicle per kilometer are set at €0.60 (Fraccascia et al., 2021).

4.2. Settings

Experiments were conducted using various parameter sets to assess the model's performance.

The cluster size takes values from the set $s \in \{2, 3, 4, 5\}$. For comparison, a simplified location problem was solved where all producers were individually assigned to active plants without clustering, denoted as $s = 1$. The number of clusters was determined by considering a range from n_0 to the minimum of $\lfloor 1.5 \cdot n_0 \rfloor$ and the total number of small producers. The parameter n_0 was established through an optimization process to identify the minimum number of clusters that meet the constraints on cluster weight and size. Testing higher values for the number of clusters allowed for solutions with partially empty clusters, thus enabling exploration of various configurations.

For each cluster size and number of clusters, the algorithm was initialized with 10 randomly selected starting centroids and executed through 10 iterations of the BKm algorithm.

The experiments are conducted on an Ubuntu server equipped with an Intel(R) Xeon(R) Gold 6252N CPU and 96 GB RAM. The bmNDP algorithm is implemented in Python and utilizes *IBM ILOG CPLEX v12.10* as the MIP optimizer. A tolerance of 0.05% is set on the gap between the best integer objective and the best remaining node objective (MIPGap).

4.3. Computational results and discussion

This section presents the results obtained from applying the proposed model.

The computation time for the BKm clustering algorithm varies significantly based on the complexity of the instances. For the most challenging instances (with fewer clusters and larger maximum cluster sizes), the computation time ranges from 20 min to two hours. Simpler instances are resolved in about three minutes. After determining cluster compositions, solving the two-commodity flow formulation for the LSRP-FC takes, on average, one minute. Detailed metrics and statistics for all test instances are provided in the Supplementary Materials.

Fig. 5 illustrates the total cost trends across all tested models. Each line on the chart corresponds to a specific value of the parameter s, displaying the objective function values achieved with varying numbers of clusters, n.

The following observations are derived from the figure:

- Impact of cluster size: increasing the parameter s (the maximum cluster size) widens the cost gap between the simple location-sizing model (denoted as $s = 1$) and the LSRP-FC model.
- Effectiveness of larger clusters: results with maximum cluster sizes of 4 and 5 are similar, with neither consistently outperforming the other. This similarity arises because relaxing the maximum number of producers per cluster can lead to constraints on cluster weight becoming binding, which may necessitate smaller clusters in some cases.
- Number of clusters trend: the total costs generally increase with the parameter n, although not monotonically. This is primarily due to the rising transportation costs associated with longer truck routes.

Despite the observed trend, it is noteworthy that the cost difference between the best and worst solutions generated by the model is relatively small, at only 2%. Specifically, the best solution, achieved with parameters $(5, 79)$, and the worst solution, achieved with $(2, 199)$, show a total cost difference of 10.8 million euros, approximately 13,300 km, and 280 h per day.

Five plant locations are consistently selected across test scenarios. Among the 192 test instances, Anzio is always chosen due to its status as an existing facility, offering a cost-effective solution with only operating costs. Tivoli and Pontinia are selected in nearly all cases as large-scale plants due to their strategic regional positions. Additionally, in 65% of instances, a medium-sized plant is set up in Civitavecchia, and the plant in Albano Laziale is the only site upgraded.

4.3.1. Optimal solution

The optimal solution is obtained with parameter values $s = 5$ and $n = 79$, yielding a total cost of €513,743,301.

Fig. 6(a) illustrates the clustering results. The cluster sizes exhibit substantial variation, ranging from small clusters with just one producer (10%) to larger clusters accommodating up to five producers (29%). The average weight of the clusters is approximately 3300 kg, with half of the clusters ranging between 3450 and 3500 kg, which aligns with the maximum load capacity of small vehicles.

Figs. 6(b) and 6(c) compare the optimal with the one obtained using the simple location-sizing model with $s = 1$. The maps display the allocation of waste to the selected plants and the capacities assigned to each. The geographical concentration of producers around each active plant aligns with the goal of minimizing truck travel distances (see Section 3.1.2). However, some exceptions arise in the optimal solution due to two factors. First, the MILP model's positive tolerance gap allows for feasible solutions to be considered optimal, introducing flexibility in producer assignment. Second, cluster composition influences results: clusters may include producers who are closer to different plants. Thus, routing optimization may assign some producers to more distant plants if most producers in their cluster are nearer to those plants.

Although the plant locations differ between the solutions, both configurations utilize five active plants, each providing a total capacity of 3320 tons per day, which slightly exceeds the regional waste production of 3260 tons per day. The total investment costs for both solutions are identical, amounting to €476,461,889, with 40% attributed to CAPEX and 60% to OPEX. Despite having the same investment costs, the solutions differ in their total operational costs. The discrepancy of €16 million between the two solutions represents a 30% difference in transportation operating costs. The optimal solution offers significant improvements in transportation efficiency, resulting in a 30% reduction in daily kilometers (20,000 km) and a 26% decrease in daily travel time (3500 h). It is important to note that while our analysis assumes all vehicles are available from the start, clustering would actually reduce the number of required vehicles and drivers compared to a non-clustering scenario, contributing further to cost savings.

4.3.2. Sensitivity and scenario analysis

To evaluate the robustness of the optimal solution, a scenario analysis was conducted by introducing two alternative scenarios with variations in key parameters:

(i) *Economic Parameter* - CAPEX Costs of Plant Upgrades: This scenario tests the impact of altering the CAPEX costs associated with upgrading plants. Specifically, the costs are varied between 35% and 65% of the CAPEX costs for new plants, compared to the initial assumption of 50%.

(ii) *Technical Parameter* - Maximum Capacity of New Plants: Considering the observed minimal gap between the total disposable capacity of the plants and the required capacity, this scenario introduces the option of opening huge-scale plants with a capacity of 288 tons per day (or 450,000 tons per year), as proposed by Fraccascia et al. (2021).

To compare the results with the optimal solution, the clustering solution generated by the BKm algorithm with parameters $s = 5$ and $n = 79$ was used as the baseline. For each new scenario, the corresponding LSRP-FC model was then solved. The main cost metrics for all scenarios are summarized in Table 3.

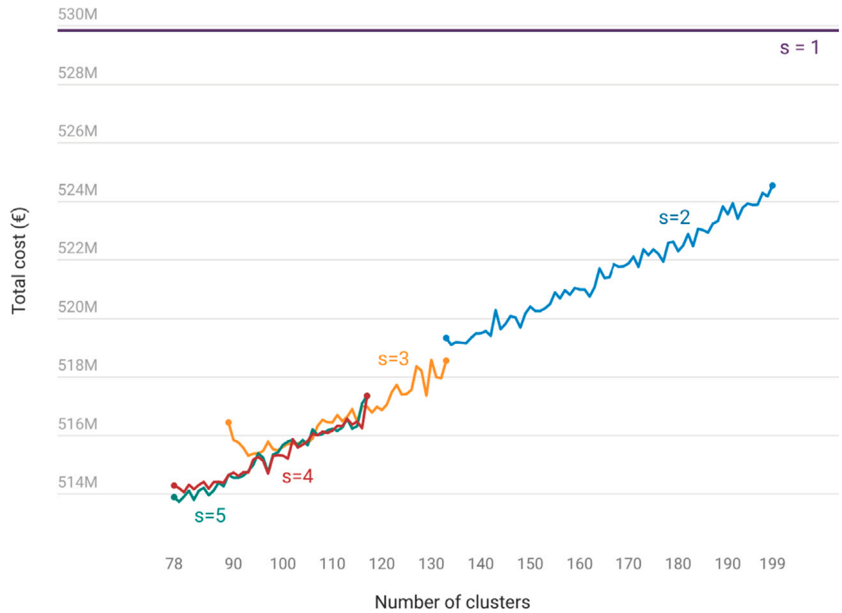
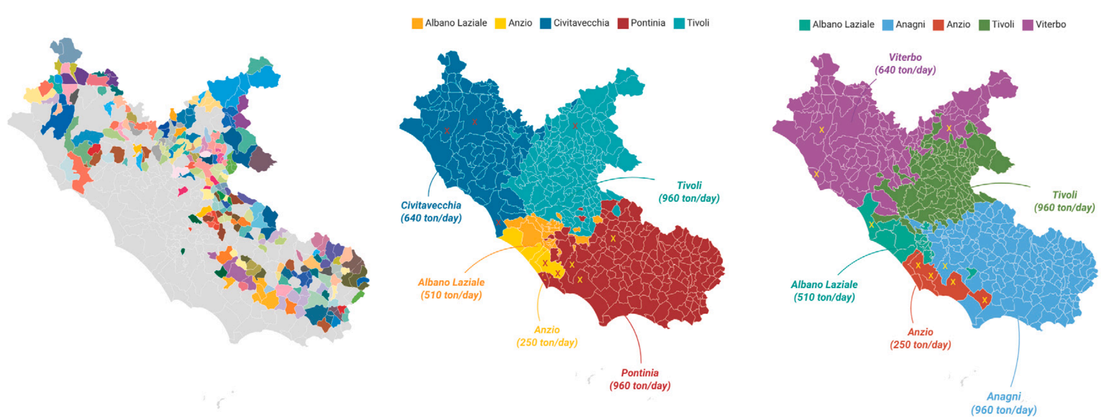


Fig. 5. Line chart of objective function values obtained for all instances. Each line represents the computational results obtained by fixing the maximum cluster size to s and the number of clusters to n. Created by Datawrapper.



(a) Lazio region clusters map for all waste small producers. Each cluster presents a different color. Producers shaded in grey represent large producers (not clustered).

(b) Lazio region map representing the best solution allocation ($s = 5$, $n = 79$).

(c) Lazio region map representing the simple location-sizing model solution allocation.

Fig. 6. Solution obtained by the algorithm and the simple location-sizing model. In subfigures (b) and (c), "x"-marked sites denote locations that were not selected. Created by Datawrapper.

Table 3
Sensitivity analysis main KPI and metrics.

New plants size	up CAPEX [%]	Obj [€]	CAPEX [€]	OPEX [€]	Transportation costs [€]	km [km]
	35	506,007,353	194,500,000	274,161,889	37,345,465	46,061
S/M/L	*50*	513,743,301	202,300,000	274,161,889	37,281,412	45,981
	65	521,589,516	210,100,000	274,161,889	37,327,627	46,039
	35	418,385,192	117,700,000	254,238,592	46,446,600	57,286
S/M/L/H	*50*	426,185,192	125,500,000	254,238,592	46,446,600	57,286
	65	433,992,894	133,300,000	254,238,592	46,454,302	57,295

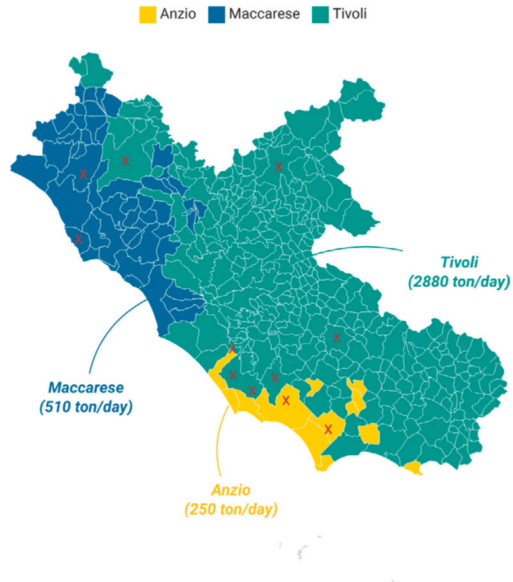


Fig. 7. Lazio region map representing the best solution found in the scenario with huge-scale plants and CAPEX of upgradable plants set to 35% of the CAPEX costs of new plants. Created by Datawrapper.

When examining scenarios with different CAPEX costs, a direct proportionality is observed between the CAPEX percentage and total costs. For instance, compared to the baseline scenario, scenarios with CAPEX at 35% and 65% result in minimal differences in total costs (−0.1% and +0.11%, respectively). This indicates that, under a fixed plant scale scenario, the selection of active plants remains stable across varying CAPEX costs. The variations in transportation costs are solely due to differences in routing strategies, which are influenced by the set MIPGap of 5% (see Section 4.2).

On the other hand, the inclusion of huge-scale plants significantly alters the solution provided by the model. Fig. 7 shows the allocation map and selected plant locations for the scenario with huge-scale plants and CAPEX of upgradable plants set at 35% of the CAPEX costs of new plants. In this scenario, the objective function value is approximately 18.5% lower than that of the baseline scenario. The optimal solution shifts from three medium/large plants to a single huge plant (Tivoli), reducing the total number of active plants from five to three (Anzio, Maccarese, and Tivoli).

Although economies of scale lower investment costs, they also lead to a 25% increase in daily kilometers traveled, raising transportation costs. However, the savings from economies of scale outweigh these additional transportation expenses, aligning with findings from previous studies (Fraccascia et al., 2021; Yue, You, & Snyder, 2014). It is important to note that the longer travel distances in this scenario could lead to higher greenhouse gas emissions, highlighting a trade-off between economic and environmental performance. Furthermore, relying on a huge plant introduces vulnerability, as the system is less resilient to disruptions that could impact the Tivoli plant's operations. This underscores the trade-off between efficiency and resilience in supply chain management, a concept also reflected in natural ecosystems (Ivanov, Sokolov, & Dolgui, 2014; Tukamuhabwa, Stevenson, Busby, & Zorzini, 2015; Ulanowicz, Goerner, Lietaer, & Gomez, 2009). Additionally, from a social perspective, the Not In My Back Yard (NIMBY) phenomenon—where local communities resist the development of large-scale energy facilities due to concerns about environmental impact, quality of life, and increased traffic (Kulla et al., 2022)—must be considered. This opposition has been increasingly evident in Italy in recent years, and larger plants are often perceived more negatively than smaller ones, making it a crucial factor in planning and decision-making.

5. Conclusions and future works

This paper represents a significant advancement in the integration of tactical and operational aspects within a biomethane supply chain fed by municipal solid waste. The main objective is to address both the strategic decision of plant location-sizing and the tactical challenge of vehicle routing.

The implications of the work lie in the successful application of the BKm clustering algorithm for waste pooling and the subsequent MILP formulation of the LSRP-FC model, which integrates plant selection and sizing, transportation routing, and economic considerations. The results indicate that the proposed two-stage approach can significantly reduce the supply chain costs. The trade-offs between investment costs and transportation expenses were carefully considered, with the inclusion of scenarios featuring huge-scale waste treatment plants. The inclusion of such plants significantly impacts the overall solution, underscoring the need for a comprehensive evaluation of technical and social implications. While presenting clear advantages in terms of economies of scale, caution must be exercised due to potential technical vulnerabilities and social resistance from local communities.

The proposed model demonstrates its flexibility and adaptability, allowing efficient solutions for medium–large instances across networks of varying scales. In our case study, the model successfully handled 390 waste producers, with 265 classified as small producers and served through waste pooling, and 14 different plant locations. These locations encompassed nine new plants with various capacities, four plants requiring upgrading for AD implementation, and one existing biomethane production plant. Remarkably, even in the most challenging instance, the model's solution was obtained within a short time frame of approximately one hour. The solution consistently outperforms a simple location-sizing model in terms of cost-effectiveness, establishing it as a valuable decision-support tool for real-life waste management systems. Furthermore, it is noteworthy that constraints presented in Section 3.1.2 might be revised or integrated according to the specific area considered, for instance, to take into account the regional legislation. This is a further element of the flexibility of the model developed in this paper.

However, some limitations need to be acknowledged. From a technical perspective, the consideration of amounts as deterministic may overlook potential variations due to factors like peak seasonality. Regular reevaluation of specific routes and pooling arrangements is advisable to adapt to practical day-to-day needs and ensure optimal system performance. Furthermore, internal collecting routes within municipalities were not modeled, and future optimization efforts should extend to this operational level as well. Additionally, the model primarily focuses on economic considerations, overlooking the environmental and social impacts of waste transportation. Future developments could include a broader range of factors, such as the use of biomethane as truck fuel to promote circularity, and the inclusion of resistance and opposition from local communities to address social perspectives. This would lead to a more comprehensive optimization process.

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.

Acknowledgments

This paper was written in the framework of the project titled "La simbiosi industriale come driver per la transizione energetica circolare: progettazione della filiera produttiva del biometano a livello regionale in Italia", supported by Sapienza University of Rome. The work of Anna Livia Croella has been supported by PNRR MUR project PE0000013-FAIR.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cie.2024.110714.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials. The data do not violate the protection of human subjects, or other valid ethical, privacy, or security concerns.

References

Ahn, Y.-C., Lee, I.-B., Lee, K.-H., & Han, J.-H. (2015). Strategic planning design of microalgae biomass-to-biodiesel supply chain network: Multi-period deterministic model. *Applied Energy*, *154*, 528–542. http://dx.doi.org/10.1016/j.apenergy.2015.05.047, URL https://www.sciencedirect.com/science/article/pii/S0306261915006649.

Akgul, O., Shah, N., & Papageorgiou, L. G. (2012). Economic optimisation of a UK advanced biofuel supply chain. *Biomass and Bioenergy*, *41*, 57–72. http://dx.doi.org/10.1016/j.biombioe.2012.01.040, URL https://www.sciencedirect.com/science/article/pii/S0961953412000505.

Ama Roma, S. p. A. (2023). Flotta e logistica. https://www.amaroma.it/mobile/pages/826-i-nuovi-mezzi-ama-per-la-pulizia-della-citta.html?news=1.

An, H., Wilhelm, W. E., & Searcy, S. W. (2011). Biofuel and petroleum-based fuel supply chain research: A literature review. *Biomass and Bioenergy*, *35*(9), 3763–3774. http://dx.doi.org/10.1016/j.biombioe.2011.06.021, URL https://www.sciencedirect.com/science/article/pii/S096195341100345X.

Arabi, M., Yaghoubi, S., & Tajik, J. (2019a). Algal biofuel supply chain network design with variable demand under alternative fuel price uncertainty: A case study. *Computers & Chemical Engineering*, *130*, Article 106528. http://dx.doi.org/10.1016/j.compchemeng.2019.106528, URL https://www.sciencedirect.com/science/article/pii/S0098135419301024.

Arabi, M., Yaghoubi, S., & Tajik, J. (2019b). A mathematical model for microalgae-based biobutanol supply chain network design under harvesting and drying uncertainties. *Energy*, *179*, 1004–1016. http://dx.doi.org/10.1016/j.energy.2019.04.219, URL https://www.sciencedirect.com/science/article/pii/S036054421930859X.

Asadi, E., Habibi, F., Nickel, S., & Sahebi, H. (2018). A bi-objective stochastic location-inventory-routing model for microalgae-based biofuel supply chain. *Applied Energy*, *228*, 2235–2261. http://dx.doi.org/10.1016/j.apenergy.2018.07.067, URL https://www.sciencedirect.com/science/article/pii/S030626191831095X.

Asefi, H., Shahparvari, S., & Chhetri, P. (2020). Advances in sustainable integrated solid waste management systems: lessons learned over the decade 2007–2018. *Journal of Environmental Planning and Management*, *63*, 1–26. http://dx.doi.org/10.1080/09640568.2020.1714562.

Atashbar, N. Z., Labadie, N., & Prins, C. (2018). Modelling and optimisation of biomass supply chains: a review. *International Journal of Production Research*, *56*(10), 3482–3506. http://dx.doi.org/10.1080/00207543.2017.1343506.

Ba, B. H., Prins, C., & Prodhon, C. (2016). Models for optimization and performance evaluation of biomass supply chains: An operations research perspective. *Renewable Energy*, *87*, 977–989. http://dx.doi.org/10.1016/j.renene.2015.07.045, URL https://www.sciencedirect.com/science/article/pii/S0960148115301403, Optimization Methods in Renewable Energy Systems Design.

Balaman, Y. S., & Selim, H. (2014). A network design model for biomass to energy supply chains with anaerobic digestion systems. *Applied Energy*, *130*, 289–304. http://dx.doi.org/10.1016/j.apenergy.2014.05.043.

Baldacci, R., Hadjiconstantinou, E., & Mingozzi, A. (2004). An exact algorithm for the capacitated vehicle routing problem based on a two-commodity network flow formulation. *Operations Research*, *52*, 723–738. http://dx.doi.org/10.1287/opre.1040.0111.

Bojesen, M., Birkin, M., & Clarke, G. (2014). Spatial competition for biogas production using insights from retail location models. *Energy*, *68*, 617–628. http://dx.doi.org/10.1016/j.energy.2013.12.039, URL https://www.sciencedirect.com/science/article/pii/S0360544213010967.

Cao, J. X., Zongxi, Z., & Zhou, Y. (2020). A location-routing problem for biomass supply chains. *Computers & Industrial Engineering*, *152*, Article 107017. http://dx.doi.org/10.1016/j.cie.2020.107017.

Chen, X., & Önal, H. (2014). An economic analysis of the future U.S. biofuel industry, facility location, and supply chain network. *Transportation Science*, *48*(4), 575–591. http://dx.doi.org/10.1287/trsc.2013.0488.

Chierichetti, F., Kumar, R., Lattanzi, S., & Vassilvitskii, S. (2018). Fair clustering through fairlets. CoRR abs/1802.05733 URL http://arxiv.org/abs/1802.05733.

Dao, T.-B.-H., Duong, K.-C., & Vrain, C. (2017). Constrained clustering by constraint programming. *Artificial Intelligence*, *244*, 70–94. http://dx.doi.org/10.1016/j.artint.2015.05.006, URL https://www.sciencedirect.com/science/article/pii/S0004370215000806, Combining Constraint Solving with Mining and Learning.

de Jesús Vargas-Soplín, A., Prochnow, A., Herrmann, C., Tscheuschner, B., & Kreidenweis, U. (2022). The potential for biogas production from autumn tree leaves to supply energy and reduce greenhouse gas emissions – A case study from the city of Berlin. *Resources, Conservation and Recycling*, *187*, Article 106598. http://dx.doi.org/10.1016/j.resconrec.2022.106598, URL https://www.sciencedirect.com/science/article/pii/S0921344922004323.

De Nicolò, M., Fraccascia, L., & Pontrandolfo, P. (2024). How the energy procurement switching strategies (driven by the Russia-Ukraine conflict) impact the global sustainability? The global sustainability dashboard. *Ecological Economics*, *225*, Article 108328. http://dx.doi.org/10.1016/j.ecolecon.2024.108328, URL https://www.sciencedirect.com/science/article/pii/S0921800924002258.

Dipartimento Trasformazione Digitale U. O. Statistica - Open Data (2023). La popolazione di roma. Struttura e dinamica demografica. https://www.comune.roma.it/web-resources/cms/documents/La_popolazione_a_Roma2019.pdf. (Accessed April 2023).

Ekşioğlu, S. D., Acharya, A., Leightley, L. E., & Arora, S. (2009). Analyzing the design and management of biomass-to-biorefinery supply chain. *Computers & Industrial Engineering*, *57*(4), 1342–1352. http://dx.doi.org/10.1016/j.cie.2009.07.003, URL https://www.sciencedirect.com/science/article/pii/S0360835209001909.

European Commission (2022). REPowerEU: A plan to rapidly reduce dependence on Russian fossil fuels and fast forward the green transition. https://ec.europa.eu/commission/presscorner/detail/en/IP_22_3131.

Eurostat (2022). EU natural gas import dependency down to 83% in 2021. https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220419-1.

Eyl-Mazzega, M.-A., Mathieu, C., Boesgaard, K., Daniel-Gromke, J., Denysenko, V., Liebetrau, J., et al. (2019). *Biogas and bio-methane in Europe: Lessons from Denmark, Germany and Italy*: *Technical report*, France: URL http://inis.iaea.org/search/search.aspx?orig_q=RN:50050009.

Fitzgerald, P. (2023). Open source routing machine. https://github.com/Project-OSRM/osrm-backend. (Accessed April 2023).

Fraccascia, L., Spagnoli, M., Riccini, L., & Nastasi, A. (2021). Designing the biomethane production chain from urban wastes at the regional level: An application to the Rome Metropolitan Area. *Journal of Environmental Management*, *297*, Article 113328. http://dx.doi.org/10.1016/j.jenvman.2021.113328.

Franco, C., Bojesen, M., Hougaard, J. L., & Nielsen, K. (2015). A fuzzy approach to a multiple criteria and geographical information system for decision support on suitable locations for biogas plants. *Applied Energy*, *140*, 304–315. http://dx.doi.org/10.1016/j.apenergy.2014.11.060, URL https://www.sciencedirect.com/science/article/pii/S030626191401229X.

Genevay, A., Dulac-Arnold, G., & Vert, J.-P. (2019). Differentiable deep clustering with cluster size constraints. ArXiv abs/1910.09036 URL https://api.semanticscholar.org/CorpusID:204800383.

Ghaderi, H., Moini, A., & Pishvaee, M. S. (2018). A multi-objective robust possibilistic programming approach to sustainable switchgrass-based bioethanol supply chain network design. *Journal of Cleaner Production*, *179*, 368–406. http://dx.doi.org/10.1016/j.jclepro.2017.12.218, URL https://www.sciencedirect.com/science/article/pii/S0959652617332018.

Ghiani, G., Laganà, D., Manni, E., Musmanno, R., & Vigo, D. (2014). Operations research in solid waste management: A survey of strategic and tactical issues. *Computers & Operations Research*, *44*, 22–32. http://dx.doi.org/10.1016/j.cor.2013.10.006, URL https://www.sciencedirect.com/science/article/pii/S0305054813003018.

Giarola, S., & Bezzo, F. (2011). Spatially explicit multi-objective optimisation for design and planning of hybrid first and second generation biorefineries. *Computers & Chemical Engineering*, *35*, 1782–1797. http://dx.doi.org/10.1016/j.compchemeng.2011.01.020.

Golden, B. L., Magnanti, T. L., & Nguyen, H. Q. (1977). Implementing vehicle routing algorithms. *Networks*, *7*(2), 113–148.

Gross, T., Breitenmoser, L., Kumar, S., Ehrensperger, A., Wintgens, T., & Hugi, C. (2021). Anaerobic digestion of biowaste in Indian municipalities: Effects on energy, fertilizers, water and the local environment. *Resources, Conservation and Recycling*, *170*, Article 105569. http://dx.doi.org/10.1016/j.resconrec.2021.105569, URL https://www.sciencedirect.com/science/article/pii/S0921344921001762.

Höhn, J., Lehtonen, E., Rasi, S., & Rintala, J. (2014). A geographical information system (GIS) based methodology for determination of potential biomasses and sites for biogas plants in southern Finland. *Applied Energy*, *113*, 1–10. http://dx.doi.org/10.1016/j.apenergy.2013.07.005, URL https://www.sciencedirect.com/science/article/pii/S0306261913005710.

Hong, B. H., Shen, H., & Lam, H. (2016). Overview of sustainable biomass supply chain: from concept to modelling. *Clean Technologies and Environmental Policy*, *18*, http://dx.doi.org/10.1007/s10098-016-1155-6.

Hoo, P. Y., Hashim, H., Ho, W. S., & Yunus, N. (2019). Spatial-economic optimisation of biomethane injection into natural gas grid: The case at southern Malaysia. *Journal of Environmental Management*, *241*, http://dx.doi.org/10.1016/j.jenvman.2018.11.092.

Höppner, F., & Klawonn, F. (2008). Clustering with size constraints. In *Computational intelligence paradigms: innovative applications* (pp. 167–180). Berlin, Heidelberg: Springer Berlin Heidelberg, http://dx.doi.org/10.1007/978-3-540-79474-5_8.

How, B. S., & Lam, H. L. (2018). Sustainability evaluation for biomass supply chain synthesis: Novel principal component analysis (PCA) aided optimisation approach. *Journal of Cleaner Production*, *189*, 941–961. http://dx.doi.org/10.1016/j.jclepro.2018.03.104, URL https://www.sciencedirect.com/science/article/pii/S0959652618307650.

Iakovou, E., Karagiannidis, A., Vlachos, D., Toka, A., & Malamakis, A. (2010). Waste biomass-to-energy supply chain management: A critical synthesis. *Waste Management*, *30*(10), 1860–1870. http://dx.doi.org/10.1016/j.wasman.2010.02.030, URL https://www.sciencedirect.com/science/article/pii/S0956053X10001169, Anaerobic Digestion (AD) of Solid Waste.

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*, 178–210. http://dx.doi.org/10.1016/j.ins.2022.11.139, URL https://www.sciencedirect.com/science/article/pii/S0020025522014633.

Ishaq, M., & Ishaq, H. (2022). Performance assessment of biogas-fed solid oxide fuel cell system for municipal solid waste treatment. *Journal of Cleaner Production*, *354*, Article 131702. http://dx.doi.org/10.1016/j.jclepro.2022.131702.

ISPRA (2022). Rapporto rifiuti urbani 2021. https://www.isprambiente.gov.it/files2022/pubblicazioni/rapporti/rapportorifiutiurbani_ed-2021-n-355-conappendice_agg18_01_2022.pdf.

ISTAT (2022). Bilancio demografico mensile e popolazione residente per sesso, anno 2022. https://demo.istat.it/app/?i=D7B.

Ivanov, D., Sokolov, B., & Dolgui, A. (2014). The ripple effect in supply chains: trade-off 'efficiency-flexibility-resilience' in disruption management. *International Journal of Production Research*, *52*(7), 2154–2172. http://dx.doi.org/10.1080/00207543.2013.858836.

Jensen, I. G., Münster, M., & Pisinger, D. (2017). Optimizing the supply chain of biomass and biogas for a single plant considering mass and energy losses. *European Journal of Operational Research*, *262*(2), 744–758. http://dx.doi.org/10.1016/j.ejor.2017.03.071, URL https://www.sciencedirect.com/science/article/pii/S0377221717303028.

Jesus, R., Souza, J., Puglieri, F., Piekarski, C., & Francisco, A. (2021). Biodigester location problems, its economic-environmental-social aspects and techniques: Areas yet to be explored. *Energy Reports*, *7*, 3998–4008. http://dx.doi.org/10.1016/j.egyr.2021.06.090.

Kulla, M., Novotný, L., Pregi, L., Dvořák, P., Martinát, S., Klusáček, P., et al. (2022). The good, the bad, and the nobody: Exploring diversity of perceptions of anaerobic digestion plants in Central and Eastern Europe. *Energy Research & Social Science*, *89*, Article 102644. http://dx.doi.org/10.1016/j.erss.2022.102644.

Laporte, G., Nobert, Y., & Desrochers, M. (1985). Optimal routing under capacity and distance restrictions. *Operations Research*, *33*(5), 1050–1073. http://dx.doi.org/10.1287/opre.33.5.1050.

Lautala, P., Hilliard, M., Webb, E., Busch, I., Hess, J., Roni, M., et al. (2015). Opportunities and challenges in the design and analysis of biomass supply chains. *Environmental Management*, *56*, http://dx.doi.org/10.1007/s00267-015-0565-2.

Letchford, A., & Salazar González, J. J. (2006). Projection results for vehicle routing. *Mathematical Programming*, *105*, 251–274. http://dx.doi.org/10.1007/s10107-005-0652-x.

Li, Y., Zhao, Q., Yang, S., & Guo, Y. (2023). Tailoring evolutionary algorithms to solve the multi-objective location-routing problem for biomass waste collection. *IEEE Transactions on Evolutionary Computation*, 1. http://dx.doi.org/10.1109/TEVC.2023.3265869.

Lin, T., Rodríguez, L. F., Shastri, Y. N., Hansen, A. C., & Ting, K. (2014). Integrated strategic and tactical biomass–biofuel supply chain optimization. *Bioresource Technology*, *156*, 256–266. http://dx.doi.org/10.1016/j.biortech.2013.12.121, URL https://www.sciencedirect.com/science/article/pii/S0960852413019615.

Lyng, K.-A., Bjerkestrand, M., Stensgård, A., Callewaert, P., & Hanssen, O. (2018). Optimising anaerobic digestion of manure resources at a regional level. *Sustainability*, *10*, 286. http://dx.doi.org/10.3390/su10010286.

MacQueen, J. B. (1967). Some methods for classification and analysis of MultiVariate observations. In L. M. L. Cam, & J. Neyman (Eds.), *Vol. 1*, *Proc. of the fifth berkeley symposium on mathematical statistics and probability* (pp. 281–297). University of California Press.

Morrissey, A., & Browne, J. (2004). Waste management models and their application to sustainable waste management. *Waste Management*, *24*(3), 297–308. http://dx.doi.org/10.1016/j.wasman.2003.09.005, URL https://www.sciencedirect.com/science/article/pii/S0956053X03001818, Landfill Process Modelling.

Nagy, G., & Salhi, S. (2007). Location-routing: Issues, models and methods. *European Journal of Operational Research*, *177*(2), 649–672. http://dx.doi.org/10.1016/j.ejor.2006.04.004, URL https://www.sciencedirect.com/science/article/pii/S0377221706002670.

Nations, U. (2022). World cities report 2022. Envisaging the future of cities. https://unhabitat.org/sites/default/files/2022/06/wcr_2022.pdf.

Ng, W. P. Q., How, B. S., Lim, C. H., Ngan, S. L., & Lam, H. L. (2022). Biomass supply chain synthesis and optimization. In *Value-chain of biofuels* (pp. 445–479). Elsevier, http://dx.doi.org/10.1016/B978-0-12-824388-6.00002-6, URL https://www.sciencedirect.com/science/article/pii/B9780128243886000026.

Ng, R. T., & Maravelias, C. T. (2017). Design of biofuel supply chains with variable regional depot and biorefinery locations. *Renewable Energy*, *100*, 90–102. http://dx.doi.org/10.1016/j.renene.2016.05.009, URL https://www.sciencedirect.com/science/article/pii/S0960148116304190, Special Issue: Control and Optimization of Renewable Energy Systems.

OpenStreetMap (2023). Openstreetmap. https://www.openstreetmap.org/copyright. (Accessed April 2023).

O'Shea, R., Wall, D., Kilgallon, I., & Murphy, J. D. (2016). Assessment of the impact of incentives and of scale on the build order and location of biomethane facilities and the feedstock they utilise. *Applied Energy*, *182*, 394–408. http://dx.doi.org/10.1016/j.apenergy.2016.08.063, URL https://www.sciencedirect.com/science/article/pii/S0306261916311503.

Osmani, A., & Zhang, J. (2014). Economic and environmental optimization of a large scale sustainable dual feedstock lignocellulosic-based bioethanol supply chain in a stochastic environment. *Applied Energy*, *114*, 572–587. http://dx.doi.org/10.1016/j.apenergy.2013.10.024.

Park, Y. S. (2019). Optimal location of biogas plants in supply chains under carbon effects: Insight from a case study on animal manure in north dakota. *Journal of Advanced Transportation*, *2019*, 13. http://dx.doi.org/10.1016/j.resconrec.2011.04.012.

Peña, J., Lozano, J., & Larrañaga, P. (1999). An empirical comparison of four initialization methods for the K-means algorithm. *Pattern Recognition Letters*, *20*(10), 1027–1040. http://dx.doi.org/10.1016/S0167-8655(99)00069-0, URL https://www.sciencedirect.com/science/article/pii/S0167865599000690.

Pérez-Ortega, J., Almanza-Ortega, N. N., Vega-Villalobos, A., Pazos-Rangel, R., Zavala-Díaz, C., & Martínez-Rebollar, A. (2019). The K-means algorithm evolution. *Introduction to Data Science and Machine Learning*, 69–90.

Piccialli, V., Russo Russo, A., & Sudoso, A. M. (2022). An exact algorithm for semi-supervised minimum sum-of-squares clustering. *Computers & Operations Research*, *147*, Article 105958. http://dx.doi.org/10.1016/j.cor.2022.105958.

Prodhon, C., & Prins, C. (2014). A survey of recent research on location-routing problems. *European Journal of Operational Research*, *238*(1), 1–17. http://dx.doi.org/10.1016/j.ejor.2014.01.005, URL https://www.sciencedirect.com/science/article/pii/S0377221714000071.

Quddus, M. A., Chowdhury, S., Marufuzzaman, M., Yu, F., & Bian, L. (2018). A two-stage chance-constrained stochastic programming model for a bio-fuel supply chain network. *International Journal of Production Economics*, *195*, 27–44. http://dx.doi.org/10.1016/j.ijpe.2017.09.019, URL https://www.sciencedirect.com/science/article/pii/S0925527317303031.

Ramos, T., Gomes, M., & Barbosa-Povoa, A. (2019). Multi-depot vehicle routing problem: a comparative study of alternative formulations. *International Journal of Logistics Research and Applications*, *23*, 1–18. http://dx.doi.org/10.1080/13675567.2019.1630374.

Sarker, B. R., Wu, B., & Paudel, K. P. (2019). Modeling and optimization of a supply chain of renewable biomass and biogas: Processing plant location. *Applied Energy*, *239*, 343–355. http://dx.doi.org/10.1016/j.apenergy.2019.01.216, URL https://www.sciencedirect.com/science/article/pii/S0306261919302363.

Sharma, B., Ingalls, R., Jones, C., & Khanchi, A. (2013). Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews*, *24*, 608–627. http://dx.doi.org/10.1016/j.rser.2013.03.049, URL https://www.sciencedirect.com/science/article/pii/S1364032113002086.

Silva, S., Alçada-Almeida, L., & Dias, L. (2017). Multiobjective programming for sizing and locating biogas plants: a model and an application in a region of Portugal. *Computers & Operations Research*, *83*, http://dx.doi.org/10.1016/j.cor.2017.02.016.

Starr, K., Gabarrell, X., Villalba, G., Talens Peiro, L., & Lombardi, L. (2014). Potential CO2 savings through biomethane generation from municipal waste biogas. *Biomass and Bioenergy*, *62*, 8–16. http://dx.doi.org/10.1016/j.biombioe.2014.01.023.

Thrän, D., Deprie, K., Dotzauer, M., Kornatz, P., Nelles, M., Radtke, K. S., et al. (2023). The potential contribution of biogas to the security of gas supply in Germany. *Energy, Sustainability and Society*, *13*(12), http://dx.doi.org/10.1186/s13705-023-00389-1.

Tukamuhabwa, B. R., Stevenson, M., Busby, J., & Zorzini, M. (2015). Supply chain resilience: definition, review and theoretical foundations for further study. *International Journal of Production Research*, *53*(18), 5592–5623. http://dx.doi.org/10.1080/00207543.2015.1037934, URL https://doi.org/10.1080/00207543.2015.1037934.

Ulanowicz, R. E., Goerner, S. J., Lietaer, B., & Gomez, R. (2009). Quantifying sustainability: Resilience, efficiency and the return of information theory. *Ecological Complexity*, *6*(1), 27–36. http://dx.doi.org/10.1016/j.ecocom.2008.10.005, URL https://www.sciencedirect.com/science/article/pii/S1476945X08000561.

Valenti, F., Porto, S. M., Dale, B. E., & Liao, W. (2018). Spatial analysis of feedstock supply and logistics to establish regional biogas power generation: A case study in the region of sicily. *Renewable and Sustainable Energy Reviews*, *97*, 50–63. http://dx.doi.org/10.1016/j.rser.2018.08.022, URL https://www.sciencedirect.com/science/article/pii/S1364032118305963.

Vrabie, C. (2021). Converting municipal waste to energy through the biomass chain, a key technology for environmental issues in (smart) cities. *Engineering Sustainability*, *13*, http://dx.doi.org/10.3390/su13094633.

Wang, J., Chai, Y., Shao, Y., & Qian, X. (2021). Techno-economic assessment of biogas project: a longitudinal case study from Japan. *Resources, Conservation and Recycling*, *164*, Article 105174. http://dx.doi.org/10.1016/j.resconrec.2020.105174, URL https://www.sciencedirect.com/science/article/pii/S0921344920304912.

Wu, B., Sarker, B. R., & Paudel, K. P. (2015). Sustainable energy from biomass: Biomethane manufacturing plant location and distribution problem. *Applied Energy*, *158*, 597–608. http://dx.doi.org/10.1016/j.apenergy.2015.08.080, URL https://www.sciencedirect.com/science/article/pii/S0306261915010168.

Yan, N., Ren, B., Wu, B., Bao, D., Zhang, X., & Wang, J. (2016). Multi-objective optimization of biomass to biomethane system. *Green Energy & Environment*, *1*, http://dx.doi.org/10.1016/j.gee.2016.05.001.

Yang, H., Dou, X., Pan, F., Wu, Q., Li, C., Zhou, B., et al. (2022). Optimal planning of local biomass-based integrated energy system considering anaerobic co-digestion. *Applied Energy*, *316*, Article 119075. http://dx.doi.org/10.1016/j.apenergy.2022.119075, URL https://www.sciencedirect.com/science/article/pii/S0306261922004652.

Yue, D., You, F., & Snyder, S. W. (2014). Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering*, *66*, 36–56. http://dx.doi.org/10.1016/j.compchemeng.2013.11.016, URL https://www.sciencedirect.com/science/article/pii/S0098135413003670, Selected papers from ESCAPE-23 (European Symposium on Computer Aided Process Engineering - 23), 9-12 June 2013, Lappeenranta, Finland.

Zhang, F., Johnson, D., Johnson, M., Watkins, D., Froese, R., & Wang, J. (2016). Decision support system integrating GIS with simulation and optimisation for a biofuel supply chain. *Renewable Energy*, *85*, 740–748. http://dx.doi.org/10.1016/j.renene.2015.07.041, URL https://www.sciencedirect.com/science/article/pii/S0960148115301373.

Zhu, Z., Hsueh, M., & He, Q. (2011). Enhancing biomethanation of municipal waste sludge with grease trap waste as a co-substrate. *Renewable Energy*, *36*, 1802–1807. http://dx.doi.org/10.1016/j.renene.2010.11.014.

Zubaryeva, A., Zaccarelli, N., Giudice, C., & Zurlini, G. (2012). Spatially explicit assessment of local biomass availability for distributed biogas production via anaerobic co-digestion: Mediterranean case study. *Renewable Energy*, *39*, 261–270. http://dx.doi.org/10.1016/j.renene.2011.08.021.