University of Tehran Press

Industrial Management Journal

Print ISSN: 2008-5885
Online ISSN: 2423-5369

Homepage: https://imj.ut.ac.ir

Designing a Green Routing Network with an Optimized Heterogeneous Fleet through Constrained Clustering: A Case Study in the Food Industry

Mahsa Pashang [1], **Seyed Hossein Razavi Hajiagha**✉ [2], **and Saeed Alaei** [3]

1. M.A. in Industrial Management, Production and Operation Management, Department of Management, Faculty of Management and Financial Sciences, Khatam University, Tehran, Iran. E-mail: m.pashang@khatm.ac.ir
2. Corresponding author, Associate Prof., Department of Management, Faculty of Management and Financial Sciences, Khatam University, Tehran, Iran. E-mail: h.razavi@kkhatam.ac.ir
3. Assistant Prof., Department of Industrial Engineering, Faculty of Engineering, Khatam University, Tehran, Iran. E-mail: s.alaee@khatm.ac.ir

Article Info

Article type:
Research Article

Article history:
Received January 13, 2025
Received in revised form May 31, 2025
Accepted June 03, 2025
Available online June 10, 2025

Keywords:
green capillary network, supply chain routing, heterogeneous transportation fleet.

ABSTRACT

Objective: This research aims to propose a large-scale vehicle routing model for the distribution network of a food industry product and apply the model in a real-world case study.

Methods: A mathematical model is formulated to minimize the total variable transportation costs. Considering the complexity of the model, a constrained clustering algorithm is used to decompose the problem. Then, vehicles are assigned to demand clusters according to their capacity. Finally, each cluster's symmetric traveling salesman problem (TSP) is solved using a genetic algorithm. The parameters of the proposed genetic algorithm were calibrated based on its widespread application in solving symmetric TSPs. A conservative approach was adopted to ensure the solution's validity by evaluating a worst-case scenario considering the highest node demands.

Results: By applying the proposed algorithm to the case study, over 2,000 demand nodes across Tehran were grouped into 91 clusters. Then, based on the demand level of each cluster, the vehicles are assigned, consisting of 26 small and 65 large cars. Within each cluster, the assigned vehicle followed an optimized route among the nodes, designed based on the optimal tour generated by solving the cluster-specific TSP using the genetic algorithm, and then returned to the central warehouse.

Conclusion: Comparing the results with the current situation, the size of the proposed transportation fleet showed a 40% reduction. Additionally, reducing fleet size and optimizing the routes improved the total distribution network costs by 25%. Given the model's computational efficiency, this improvement is considered satisfactory.

Cite this article: Pashang, M., Razavi Hajiagha, H., & Alaei, S. (2025). Designing a green routing network with an optimized heterogeneous fleet through constrained clustering: A case study in the food industry. ***Industrial Management Journal***, *17*(2), 175-197. https://doi.org/10.22059/imj.2025.388687.1008224

Publisher: University of Tehran Press.

DOI: https://doi.org/10.22059/imj.2025.388687.1008224

Introduction

Food security is a critical global issue directly connected to various sustainable development goals (SDGs), including poverty alleviation, hunger eradication, health and wellness, climate change, and land conservation (Pandey & Pandey, 2023). The food supply chain is key to ensuring food security and maintaining societal sustainability within this context. Due to the perishable nature of food products, efficient design and management of distribution networks, particularly in densely populated urban areas, is essential for minimizing waste, preserving quality, and improving timely access to food resources (Marzban et al., 2023). In recent years, increased public awareness of the environmental damage caused by fossil fuel consumption has pushed organizations to develop green transportation systems to minimize emissions from logistics operations (Asgharizadeh et al., 2017). From an economic standpoint, distribution and logistics networks account for 5% to 20% of a country's GDP and approximately 20% of the total cost of goods sold (Andrejić et al., 2018). According to the International Energy Agency (IEA), about 20% of Iran's total CO_2 emissions in 2022 resulted from transportation-related fossil fuel consumption. Global reports indicate that most carbon monoxide (CO) pollution in Arctic Council member countries originates from transportation. As shown in Figure 1, the green bar highlights the significant difference between transportation and other sources. The data derived from this specific world region may also be generalized to the whole globe.

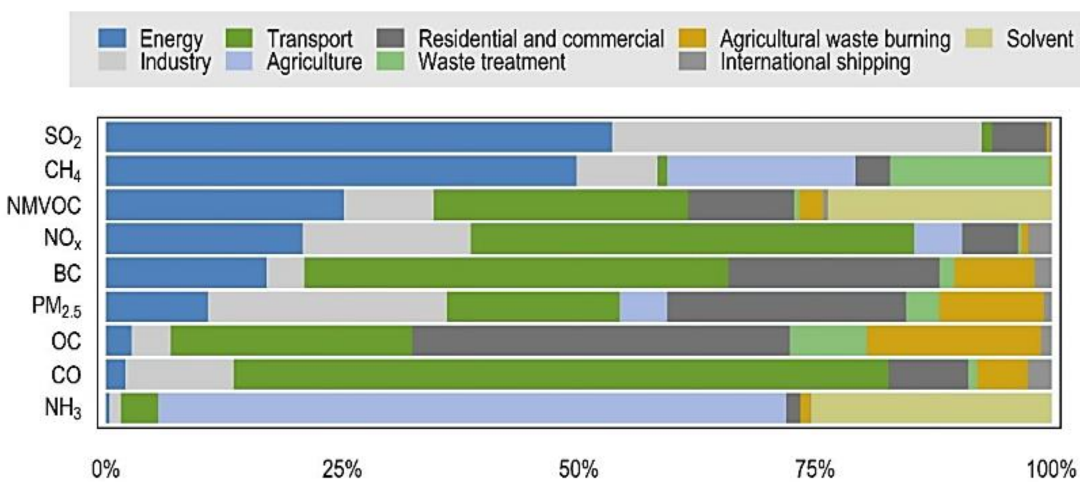


Figure 1. Emission rates of air contaminants among Arctic Council member countries (Valriberas, 2023)

Another study on greenhouse gas emissions reveals that the amount of GHGs generally released by food transportation accounts for 19% of emissions from the food industry. In the case of fruit and vegetables, the emission of transportation (36%) is approximately twice the amount released during their production (19%) (Li et al., 2022; European Commission report, 2023). This report concentrates on pollution caused by food transportation from farms to manufacturing and consumers. Sustainable development, rooted in the 1970s (Ruggerio, 2021), is based on three pillars—economic, social, and environmental—and aims to balance them in development processes (Mensah & Casadevall, 2019; Mangukiya & Sklarew, 2023). In the food industry, sustainability is defined by the pursuit of balance among economic productivity, environmental health, and social justice (Prasanna et al., 2024). Achieving sustainable development is unattainable without sustainable production and distribution (Selvan et al., 2023). Sustainable distribution is any combination of transportation tools and practices that causes minimal harm to the pillars of sustainability (Sommerauerova et al., 2018). Consequently, incorporating sustainability principles into the design of distribution networks is increasingly vital (Lin et al., 2014; Ganji et al., 2020). As a significant component of the supply chain, the distribution stage significantly influences the food industry's environmental, economic, and social dimensions. Sustainability in the food sector requires a holistic perspective that balances economic viability, environmental integrity, and social equity (Prasanna et al., 2024). Therefore, efficient distribution network design and management, particularly in urban settings, are crucial for improving the sustainability and efficiency of food systems (Marzban et al., 2023).

Distribution networks act as the interface between producers and consumers, playing a decisive role across the supply chain. Vehicle routing is one of the primary and ongoing problems of distribution networks optimization (Kazemi et al., 2021). A well-designed routing plan is crucial for improving distribution issues, directly affecting delivery speed, timing, and cost (Zhang et al., 2021).

Choosing a suitable and appropriate path for vehicles used for goods distribution is a strategic and vital subject. The Vehicle Routing Problem (VRP) is a complex combinatorial optimization problem aimed at finding feasible routes for a fleet of vehicles that deliver goods or services from a depot to a set of customers (Shi, 2024). With a limited capacity, each vehicle must return to the depot after completing its tour, and each customer is served by exactly one vehicle. Practically, this problem seeks to minimize total travel distance (Baker & Ayechew, 2003). VRP extends the traveling salesman problem (TSP) by considering multiple vehicles and deliveries. Over time, numerous VRP variants have emerged to address real-world complexities such as traffic congestion, time windows, dynamic demand, and heterogeneous fleets (Braekers et al., 2016).

VRP solution techniques are generally categorized into three types: exact methods (e.g., branch and bound, minimum spanning tree, branch and cut, dynamic programming), heuristics (e.g., k-opt, nearest neighbor), and metaheuristics (e.g., genetic algorithms, simulated annealing, tabu search, particle swarm optimization, ant colony optimization) (Toth & Vigo, 2002; 2014). As problem size increases, exact methods become impractical for timely solutions, and heuristic or metaheuristic approaches—offering a balance between solution quality and computational efficiency—become preferable (Alshaar & Awad, 2019). Given the scale, complexity, and sensitivity of distribution networks in the real-world food industry, using robust models to handle the uncertainty and intelligent algorithms to deal with their complexity is critically important.

Based on these considerations, the current study focuses on optimizing the product distribution network of an Iranian food production company using a heterogeneous fleet. Currently, the company operates 151 vehicles, incurring annual transportation costs of over 55 billion IRR. Considering the large number of customers, the current study formulated a large-scale vehicle routing model and applied a multi-stage algorithm to solve it. First, customer nodes are clustered based on their demand and vehicle capacity. Then, an appropriate vehicle is assigned to each cluster. Finally, a TSP is formulated and solved for each cluster.

The remainder of this study is organized as follows: Section 2 reviews the related literature in two distinct fields. Section 3 presents the mathematical formulation and introduces the novel model proposed in this study. In Section 4, a real-world case study is discussed. Section 5 provides the discussion and conclusions, while Section 6 outlines the study's limitations. Finally, Section 7 offers recommendations for future research.

Literature Background

From a theoretical standpoint, research on the VRP can be divided into two categories: (1) studies that focus on developing various VRP models under different assumptions and characteristics (Campbell & Wilson, 2014), and (2) studies that propose different solution methods (Alesiani et al., 2022). Practically, VRP has wide applications. Vehicle assignment and routing models have a long-standing presence in the literature. In the food industry, notable studies include Chen et al. (2020), Giallanza and Puma (2020), Yagmur et al. (2021), Torabzadeh et al. (2022), and Wang et al. (2023).

Numerous studies on vehicle routing can be classified based on model type and the algorithms used. Eksioglu et al. (2009) conducted a comprehensive review, categorizing prior studies based on study type (theoretical, applied, or review), scenario features (e.g., number of stops, loading constraints, customer demand level), physical features (e.g., network design, customer location, number of depots), and data types. Building on this, Braekers et al. (2016) extended this classification framework to cover studies published between 2009 and 2016.

Generally, review studies in this field have contributed to developing base VRP models from various angles and aspects. For instance, Gansterer and Hartl (2018) reviewed collaborative VRP studies. Other researchers, such as Koç et al. (2020), Maghdani et al. (2021), and Asgari and Mirzapur Al-Hashem (2021), have focused on categorizing studies related to green VRP. Konstantakopoulos et al. (2022) reviewed VRP-related articles and distribution algorithms comprehensively. Sar and Ghadimi (2023) extended this review to reverse logistics operations. Alzate et al. (2024) examined VRP studies focusing on sustainability dimensions since 2013.

Large-scale VRPs are considered extremely complex, making them nearly impossible to solve using exact algorithms. Qin et al. (2021) proposed a mixed-integer linear programming model for routing a predefined heterogeneous fleet to minimize travel time; however, their method failed for large-scale problems. For such cases, they recommended a metaheuristic approach based on reinforcement learning.

Han (2023) addressed multi-depot VRPs in large logistics centers using a framework based on parallel OCS-k-means clustering and large-scale neighborhood search in a distributed computing environment, which doubled convergence speed compared to parallel algorithms. Li et al. (2022) developed a three-stage framework that showed strong performance in real—world applications, including customer clustering, cluster merging, and path search. Alesiani et al. (2022) proposed a constrained clustering algorithm for capacitated VRPs that balanced demand across clusters and managed complex constraints, performing well on problems with up to 16,000 nodes.

Table 1 summarizes recent large-scale VRP studies and situates the present study within this research landscape.

Table 1. Summary of reviewed studies

Study/ Author(s)		Kin et al.	Zhang et al.	Alesiani et al.	Rajaei et al.	Tao et al.	Haripriya and Gansan	Lee et al.	Han	Ho et al.	Liu et al.	Cavalier et al.	Yeh et al.	Present Study
Year		2021	2021	2022	2022	2022	2022	2022	2023	2023	2024	2024	2024	
Fleet type	Homogeneous			✓			✓	✓	✓	✓	✓	✓	✓	
	Heterogeneous	✓	✓		✓	✓								✓
Uncertainty Type	Deterministic			✓				✓	✓	✓	✓	✓	✓	
	Stochastic	✓	✓		✓	✓								✓
	Robust						✓							
Single-objective		Time		cost	cost	cost			distance	distance	cost	cost	cost	cost
Multi-objective	Cost		✓					✓						
	Distance						✓	✓						
	Traffic						✓							
	Delay penalty		✓											
Number of nodes		100	30	16000	850	120	40000	67	1300	5000	200	30000	100000	2000

According to Table 1, most research on large-scale vehicle routing has focused on single-objective problems aimed at minimizing cost, and researchers have primarily concentrated on developing efficient algorithms to solve such problems. Consequently, most studies have been organized around the classical VRP model. Even in multi-objective studies, cost relative to travel distance is consistently included as a primary objective, and occasionally, sustainability-related goals—such as delay penalties (related to customer satisfaction) or traffic reduction (via fewer vehicles)—have been considered. The innovation of the present study lies in modeling a heterogeneous fleet while simultaneously minimizing both transportation costs and environmental pollution from fuel consumption, which is a large-scale problem. To address the model's complexity and make it solvable, a three-stage decomposition algorithm is proposed: (1) constrained clustering of demand nodes based on fleet capacity, (2) vehicle assignment, and (3) application of a genetic algorithm to solve the problem within each cluster. Finally, the proposed model and algorithm are applied to a real-world case. According to the literature review, the main contribution of this study is its focus on large-scale VRP modeling and solution.

Materials and Methods

Mathematical Model

The investigated problem involves the distribution network of a food production company in Tehran. The company distributes products via a fleet of two vehicles—light and heavy—from a central warehouse on the city's outskirts. Deliveries are made twice weekly, on Sundays and Wednesdays, and the proposed model pertains to each delivery cycle.

Formally, the problem can be defined as a large-scale vehicle routing problem over a graph G = (V, A), where V = {0, 1, ..., n} represents the set of nodes (demand points/customers), and A represents the set of arcs. Each node j has a non-negative demand dj. The vehicle routing problem in this study is considered symmetric, i.e., $c_{ij} = c_{ji}$, where c_{ij} is the cost or distance of travelling from node i to node j. A heterogeneous fleet comprising K types of vehicles is available, each with a specific capacity, per-kilometer travel cost, and fuel consumption. The objective is to find the shortest Hamiltonian tour starting and ending at the depot, visiting each node exactly once (Stavropoulou, 2022).

- **Assumptions of the Mathematical Model:**

1. Vehicles start their tours from the depot.
2. Only one vehicle is assigned per tour.
3. Each customer is visited exactly once.
4. Each tour ends at the depot.

5. Multiple types of vehicles with varying capacities are available.
6. Customer demand is considered stochastic, defined over an interval based on historical data.
7. The model is designed for a single ordering cycle.

The notations used are as follows.

Sets and Indices

i, j	demand nodes, $i \in \{0,1,\dots,n\}$
k	Set of vehicles, $k \in \{1,2,\dots,K\}$

Parameters:

d_{ij}	Distance from node *i* to node *j*
q_j	Demand of customer node *j*
c_k	Variable travel cost per kilometer for vehicle type k
f_k	Fuel consumption rate per kilometer by vehicle type k
G	Per unit of fuel cost
Q_k	Maximum capacity of vehicle type *k*

Decision variables

q_{ijk}	Load transported by vehicle k from node *i* to node *j*
x_{ijk}	$\begin{cases} 1, & \text{if vehicle type } k \text{ travels from node } i \text{ to node } j \\ 0, & \text{otherwise} \end{cases}$
y_{jk}	$\begin{cases} 1, & \text{if vehicle type } k \text{ serves demand node } j \\ 0, & \text{otherwise} \end{cases}$
$CapIu_i$	Auxiliary continuous variable (real number)

The mathematical model of the problem is formulated as follows.

$$Min \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=1}^{m} (c_k + f_k \cdot g) d_{ij} \cdot x_{ijk} \tag{1}$$

$$\sum_{\substack{i=0 \\ i \neq j}}^{n} x_{ijk} = \sum_{\substack{i=0 \\ i \neq j}}^{n} x_{jik} \quad \forall k \in K. \forall j \in I \tag{2}$$

$$\sum_{i=0}^{n} x_{ijk} = y_{jk} \quad \forall k \in K.\ \ \forall j \in I \tag{3}$$

$$\sum_{k=1}^{K} y_{jk} = 1 \quad \forall j \in I \setminus \{0\} \tag{4}$$

$$\sum_{j=1}^{n} x_{0jk} \leq 1 \quad \forall k \in K \tag{5}$$

$$u_i - u_j + n\sum_{k=0}^{m} x_{ijk} \leq n-1 \qquad \forall i \in I. \forall j \in I \setminus \{0\} \tag{6}$$

$$\sum_{j=1}^{n} q_j \cdot y_{jk} \leq Q_k \quad \forall k \in K \tag{7}$$

$$\sum_{i=0}^{n} q_{ijk} - \sum_{i=1}^{n} q_{jik} = q_j \cdot y_{jk} \quad \forall k \in K. \ \forall j \in I \setminus \{0\} \tag{8}$$

$$q_{ijk} \leq Q_k \cdot x_{ijk} \quad \forall i \in I. \forall j \in I. \forall k \in K \tag{9}$$

$$q_{ijk} \geq 0 \qquad \forall i \in I. \forall j \in I. \forall k \in K \tag{10}$$

$$u_i \geq 0 \qquad \forall i \in I \tag{11}$$

$$x_{ijk} \in \{0.1\} \quad \forall i \in I. \forall j \in I. \forall k \in K \tag{12}$$

$$y_{ik} \in \{0.1\} \quad \forall i \in I. \forall k \in K \tag{13}$$

The objective function in Equation (1) aims to minimize the total transportation cost of the fleet to serve customer demands, considering both the rental cost per kilometer and fuel consumption for the used vehicles. Minimizing this cost corresponds to identifying the shortest route in the network under the given constraints. Constraints (2) to (13) are defined to represent the feasible space of the problem. Constraints (2) to (6) are part of the classical VRP formulation (Toth & Vigo, 2002). Constraints (2) and (3) ensure that the exact vehicle that enters a node also exits it. Constraint (4) ensures that each node is visited exactly once. Constraint (5) ensures that once a vehicle is assigned to a node, no other vehicles are assigned to that same node. Constraint (6) eliminates sub-tours, while Constraint (7) reflects the vehicle capacity limit. Constraints (8) to (10) further reinforce this by regulating the flow of loads between nodes. Specifically, Constraint (8) balances incoming and outgoing flow, stating that the difference in vehicle load before and after visiting a node should equal that node's demand. Constraint (9) prevents load transfers over unused paths to avoid mismatched deliveries. Lastly, Constraint (10) ensures non-negativity of vehicle loads.

Solution Approach

Given the previous studies, the problem is classified as NP-hard. Therefore, an exact solution is impractical for large-scale instances. Most existing research addresses problems with fewer than 1,000 nodes. For this reason, a three-stage heuristic algorithm, illustrated in Figure 2, is proposed to solve the problem at the considered scale.

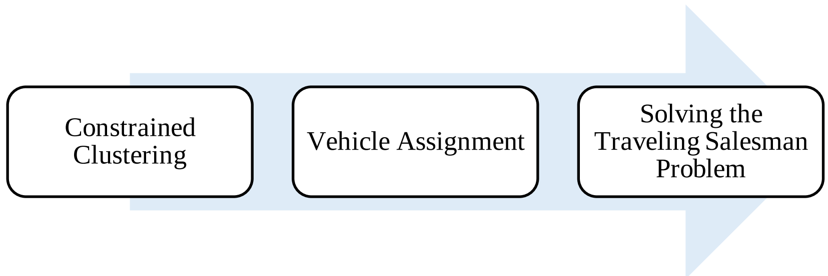


Figure 2. Three-Stage Heuristic Algorithm for Solving the Research Model

Based on previous studies, many researchers have employed clustering and decomposition techniques to reduce problem complexity. Similarly, the present study adopts this approach for problem-solving, with the subsequent sections detailing the procedural steps (Yeh et al., 2024).

Step 1: Constrained Clustering

The first step in making the model solvable involves decomposing it into smaller subproblems. In this phase, demand nodes are clustered such that the total demand of each cluster can be served by a single light or heavy vehicle. As mentioned before, VRP is an NP-hard problem. A Clustered Vehicle Routing Problem (CluVRP) approach is adopted to solve it. This was first introduced by Sevaux and Sorensen (2008) in the context of a real-world distribution network design problem. The initial step in solving CluVRP is generating clusters from demand nodes.

Due to the high computational complexity of clustering algorithms, repeated execution over many scenarios can lead to increased computation time, which reduces the model's practical applicability. Therefore, based on the clustering approach of Mirzapur Al-Hashem and Asgari (2023), this study assumes a worst-case scenario in which every demand node places at least one order, and no node in the network is demand-free. A solution that satisfies this worst-case condition will also be valid for all other scenarios.

Clustering is based on geographic coordinates (latitude and longitude) and demand levels. The first step is determining the minimum number of clusters (k) so that each cluster's total demand is less than a large vehicle's capacity.

The clustering algorithm used in this study is a constrained version of the k-means algorithm. Alesiani et al. (2022) proposed an algorithm incorporating the constraint on total cluster demand. In each iteration, for a given number of clusters, a k-means problem is solved. The key distinction here is that when evaluating whether to assign a demand node i to cluster k (based on the shortest distance to the cluster centroid), an additional constraint is enforced: the total demand of the cluster, after including node i, must not exceed the capacity of a large vehicle. Formally: $qi + \Sigma\ (i' \in k)\ qi' \leq U1$

If this condition holds, node i is added to cluster k. Otherwise, cluster k is removed from the list of candidate clusters for node i. The pseudocode for this constrained clustering algorithm is illustrated in Figure 3. This figure, dik denotes the distance between node i and cluster k's center. In Step 5, demand nodes are sorted based on their distance to cluster centers. Algorithm 2 determines the feasibility of assigning nodes to clusters under the demand constraint. The cluster center coordinates are updated upon assigning a customer to a cluster. Algorithm 2 terminates when all demand nodes are assigned. Algorithm 3 involves updating the cluster centers after removing assigned customers and repeating Steps 1 and 2 until convergence is achieved across two consecutive iterations. If some customers remain unassigned or clusters contain no assigned customers, the number of clusters k is adjusted, and the steps are repeated. This iterative process guarantees the assignment of all demand nodes to clusters while respecting the demand constraints.

Algorithm 1: Obtaining the Sorted List of Customers (Demand Nodes):

1: input: $\{d_{i,k}\}, U'$

2: for $i \in U'$ do

3: Compute D_i by solving $\min_k d_{i,k}$

4: set P = (b_1, …, b_n) such that $D_{b_1} \leq \cdots \leq D_{b_n}$

5: output: P = (b_1, …, b_n)

Algorithm 2: Assigning Demand Nodes to Clusters Starting from Customer b1 in the Priority Queue:

1: input: $\{u_i\}, W, R, \{\mu_i\}, K, \{c_k\}, \{d_{i,k}\}, U$

2: Execute Algorithm 1 to compute P

3: for $b_i \in P$ do

4: set $C_{b_i} = K$

5: while $b_i \in U \wedge C_{b_i} \neq \emptyset$ do

6: set $k^* = \arg\min_k d_{b_i,k}$

7: if $|a_{k^*}| < W \wedge d_{b_i,k^*} \leq R \wedge \mu_{b_1} + \sum_{j \in a_{k^*}} \mu_j \leq Q$ then

8: $a_{k^*} \leftarrow a_{k^*} \cup \{b_i\}$

9: $c_k := \frac{1}{|a_{k^*}|} \sum_{i \in a_{k^*}} u_i$

10: $U = U\backslash\{b_i\}$

11: else

12: $C_{b_i} \leftarrow C_{b_i}\backslash\{k^*\}$

13: output: $\{a_k\}, \{c_k\}$

Algorithm 3: Constrained Clustering Algorithm

1: while $a_k \neq \emptyset \forall k \in K \wedge U \neq \emptyset$ do

2: $c_k \sim U(u_1, \dots, u_n), \forall k \in K$, such that $c_k \neq c_j, \forall j \in K$

3: $U \leftarrow \{1, \dots, n\}$

4: $a_k \leftarrow \emptyset, \forall k \in K$

5: while $\left(\{a_k\}_1^{|K|}\right)_{\text{previous}} \neq \left(\{a_k\}_1^{|K|}\right)$ do

6: $\left(\{a_k\}_1^{|K|}\right)_{\text{previous}} \leftarrow \left(\{a_k\}_1^{|K|}\right)$

7: $\left(\{a_k\}_1^{|K|}\right) \leftarrow \text{Alg. 2}$

8: if $U \neq \emptyset$ then

9: set $|K| \leftarrow |K| + 1, \text{if there are unassigned customers}$

10: if $\exists k | a_k = \emptyset$ then

11: set $|K| \leftarrow |K| - 1, \text{if there are clusters without customers:}\ a_k = \emptyset, \text{for some}\ k \in K$

12: output:

Figure 3. Pseudocode of the Constrained Clustering Algorithm by Alesiani et al. (2022)

Step Two: Vehicle Allocation

After identifying the demand clusters, the next step involves allocating vehicles to each cluster. Considering that, in the present case study, the transportation fleet consists of two types of vehicles—small and large capacity—and that during the clustering phase, the demand size of each cluster was accounted for relative to vehicle capacities, vehicle allocation can thus be performed based on the total demand of the nodes within each cluster.

The proposed logic for vehicle allocation is as follows:

If the total demand of the nodes in a cluster is less than the capacity of a small vehicle, one small vehicle is allocated to that cluster; otherwise, a vehicle with a larger capacity is assigned.

Let qj denote the demand at node j, and Qk represent the capacity of vehicle type k. Suppose the demand points have been grouped into C clusters. Among all vehicles whose capacities exceed the aggregate demand of the cluster in question, the vehicle with the smallest adequate capacity is allocated to that cluster.

Step Three: Solving the TSP

Following the clustering of demand nodes and vehicle allocation for each cluster, the research problem is reduced to solving k symmetric TSPs, where k is the number of clusters. After the first two steps, k clusters exist, each comprising several demand nodes to be serviced by a single vehicle (salesman). A genetic algorithm is employed to solve each TSP. The pseudocode of the Genetic Algorithm utilized for addressing the problem is illustrated in Figure 4.

```
[Initialization]
    [Initialize Parameters] (PopSize, NumGen, Pc, Pm, StopCriteria, ...)
    [Initialize Parameters] Generate PopSize chromosomes randomly.
[Evaluation] Evaluate the fitness of each chromosome.
[New Generation]
Repeat
    [Selection] Select parents based on the selection strategy.
    [Crossover] Produce (PopSize*Pc) of offspring with Crossover.
[Mutation] Produce (PopSize*Pm) of offspring with Mutation.
[Reproduction] Copy the remaining chromosomes based on elitism.
      [Replacing] Place new offspring in the new population.
      [Evaluation] Evaluate the fitness of new chromosomes.
Until StopCriteria is met.
[End] return the best solution in the final population.
```

Figure 4. Pseudocode of the Genetic Algorithm

Results

Case study

The method is examined in a food manufacturing company to evaluate the proposed model and algorithm. The company produces a packaged protein product. This product is distributed through a central warehouse to 2,106 retail outlets in Tehran. The company's transportation fleet consists of two types of vehicles: large and small. Deliveries are made twice weekly, on Sundays and Wednesdays, and the proposed model pertains to each delivery cycle. Customer demand data was collected over six months, and based on Chebyshev's inequality, the confidence interval for the demand at each node was estimated. Following a conservative approach, the model was solved using the upper bound of this confidence interval. Additionally, a heterogeneous fleet list for the company was obtained, comprising 151 vehicles, including trucks and vans. The subsequent section outlines the step-by-step application of the proposed algorithm to the problem.

Step 1: Clustering Demand Nodes

Given the large scale of the problem under consideration, as previously noted, the first step involves clustering the demand nodes to reduce problem complexity by dividing it into smaller subproblems. The constrained clustering algorithm illustrated in Figure 3 was employed to generate these clusters. The inputs to this algorithm are the geographical coordinates (latitude and longitude) of the demand nodes, alongside a constraint on the demand of each cluster. Specifically, this constraint ensures that the total demand within each cluster does not exceed the maximum capacity of a large vehicle. The algorithm was coded and executed using MATLAB software. Upon running the algorithm, 91 clusters of demand nodes were obtained as shown in Figure 5. According to the clustering results, the smallest cluster contains only one demand node, while the most significant cluster comprises 49 demand nodes.

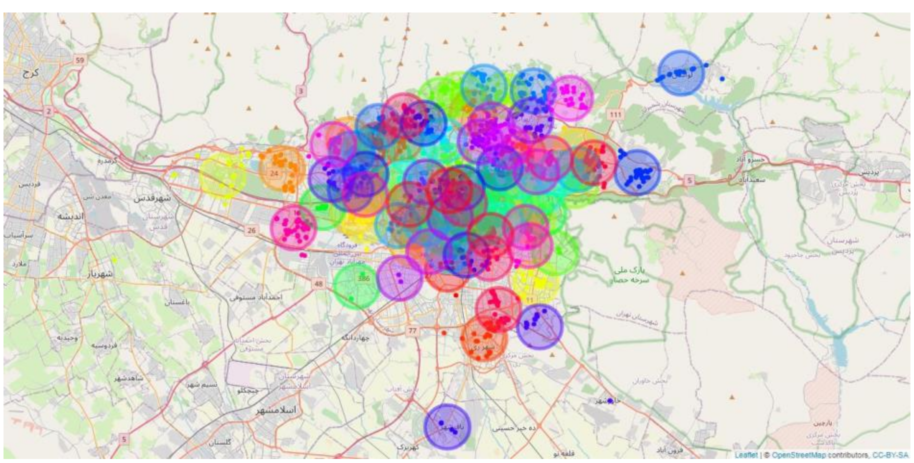


Figure 5. Clustering of Demand Nodes in Tehran

Step 2: Vehicle Allocation to Clusters

According to the designed algorithm, vehicle allocation to the clusters is carried out in the second step. The transportation fleet comprises two types of vehicles: large-capacity vehicles and small-capacity vehicles. To simplify the solution of the large-scale heterogeneous vehicle routing problem with uncertain demand, vehicle allocation in this step is based on the total demand of each cluster. According to the proposed rule, if the total demand of the nodes within a cluster is less than the capacity of a van, one small vehicle is assigned to the cluster; otherwise, one large vehicle is allocated. In total, 26 small vehicles with a capacity of 70 boxes and 65 large vehicles with a capacity of 280 boxes are required for the network.

Step 3: Solving the TSP in Each Cluster

In the third step of the proposed algorithm, the symmetric TSPs are solved for each cluster. Following the first two steps of the algorithm, 91 symmetric TSPs—one for each cluster, each consisting of several demand nodes and one vehicle—are solved. A genetic algorithm is employed to address the TSP in the obtained clusters. In this algorithm, the chromosomes representing the solution for a problem with n cities are n-length strings, where each gene corresponds sequentially to the index of the visited city. An example chromosome is illustrated in Figure 6.

P	Q	…	i	j	…	m
1	2	…	k	$k+1$		n

Figure 6. Chromosome Representation of the Solution in the Genetic Algorithm

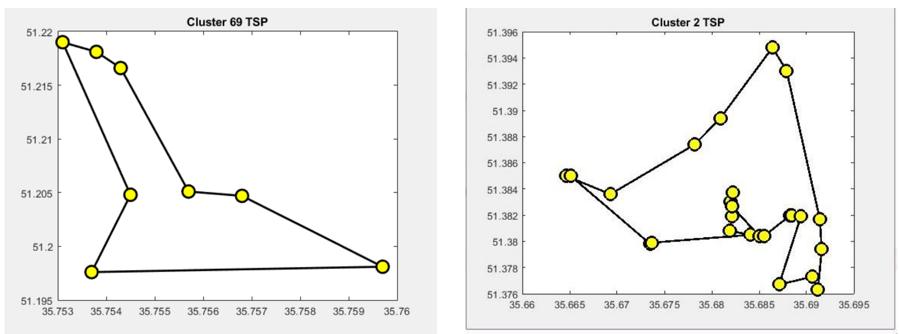
In Figure 6, cities numbered p and q are visited first and second, respectively, while cities i and j are visited at positions k and k + 1, respectively. Essentially, the chromosome represents the sequence of cities visited in order.

For applying the genetic algorithm in solving these problems, the algorithm parameters were set considering Hakimzadeh Abyaneh (2012) as follows:

- Population size: 50
- Crossover rate: 30%
- Crossover operators: Three crossover operators generated offspring, selecting the best-performing operator via a roulette wheel selection mechanism.
- Mutation operators: Three types of mutation operators, including the competency-based operator, the ordered operator, and the partially mapped operator, were employed. The roulette wheel selection was chosen for superior operator selection.

- Mutation rate: 5%, with 35% of selected solutions undergoing swapping mutation, 20% undergoing inversion mutation, and 45% being replaced.
- Elitism: Apart from the percentage of solutions in each generation subjected to crossover and mutation, the remaining top-performing solutions were carried over to the next generation.

The TSP is solved for each cluster. Then, the optimal tour comprising the visiting sequence of demand nodes is achieved for each cluster. For instance, the optimal tours for clusters 2 and 69 are illustrated in Figures 7(a) and 7(b), respectively. These tours represent the sequence of visits by the vehicle dispatched from the central warehouse to the respective clusters.



(b) (a)

Figure 7. Optimal Tour for Clusters: (a) Cluster 2, (b) Cluster 69

The total distance traveled across all clusters for each order fulfillment cycle is estimated to be 2,645.53 kilometers. Regarding the scale of the problem, presenting the routes for all clusters is impractical; therefore, the distribution of distances traveled per cluster is depicted as a histogram shown in Figure 8. Figure 8 reveals that the frequency distribution of cluster travel distances follows a skewed distribution with a peak frequency around 10 kilometers. In fact, for most clusters, the traveled distance is symmetrically centered around 10 kilometers. However, considering the vast area of metropolitan Tehran, there is one cluster located at a significantly greater distance from the central warehouse, which requires a travel distance of approximately 85 kilometers for delivery.

As can be seen, the demand of the majority of clusters is met within a travel distance of less than 10 kilometers, which indicates a reduction in the overall travel distance within Tehran. This reduction will lead to enhanced distribution network efficiency in light of urban traffic conditions.

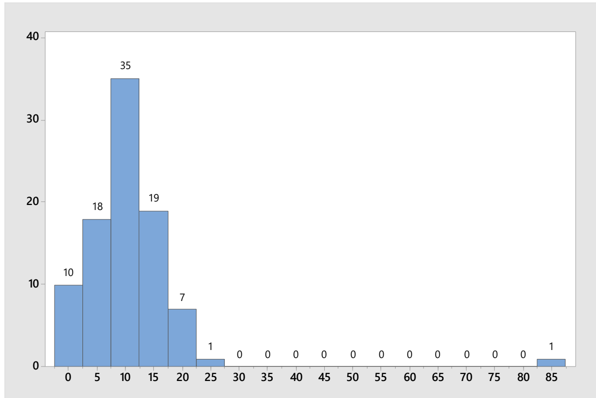


Figure 8. Histogram of Distance Traveled Within Clusters

The execution time of each of the three main stages was measured separately to assess the computational efficiency of the proposed algorithm. The results are categorized as follows:

- The clustering of demand nodes was completed in less than 30 seconds while using the developed constrained clustering algorithm in MATLAB.

- Vehicle assignment to clusters incurred a negligible computational cost and was executed approximately immediately (under 1 second).

- Finally, a symmetric TSP was solved using a genetic algorithm for each of the 91 clusters. Due to variations in cluster sizes (ranging from 1 to 49 nodes), the solution time for each TSP varied between 1 and 8 seconds, respectively. On average, the entire third step was completed in approximately 300 to 400 seconds across multiple runs.

Thus, the total runtime of the proposed algorithm for the case study lasted about 5 to 7 minutes, indicating its feasibility for real-world, large-scale applications.

Discussion and Conclusion

In many manufacturing companies, the distribution network design is assumed to be a critical component of operational and financial performance. It has a significant dual effect on both costs and revenues. This issue becomes even more highlighted in the context of food products, due to their limited shelf life. Route optimization in supply chains has long been recognized as one of the most vital components of green SCM.

A three-stage heuristic algorithm is developed to solve a large-scale supply chain routing problem for a real-world case study. The mathematical model incorporating over 2,000 demand nodes was formulated to minimize transportation and fuel costs. Regarding the problem's scale and complexity, obtaining exact solutions through deterministic methods is not feasible within a reasonable timeframe. For this purpose, a heuristic solution approach was proposed.

As outlined, in the first step, to reduce effectively the dimensions of the original problem, 91 clusters were formed. Then, according to the algorithm's output, the company's original fleet—87 large and 64 small vehicles—was reduced to 91 vehicles, specifically 65 large and 26 small ones. These two stages of the algorithm alone reduced fleet size from 151 to 91 vehicles, which can significantly decrease fixed and variable transportation costs.

In the third stage, the problem was divided into multiple symmetric TSPs. The total cost of dispatching vehicles from the central warehouse to cluster centers and performing intra-cluster deliveries was estimated at approximately 459 million IRR per distribution cycle, considering standard vehicle rental rates and fuel consumption. Compared to the current operational cost of 575 million IRR, the estimated figure represents a 25% cost reduction.

A key novelty of this research is introducing a novel distribution system for the food supply chain that utilizes a heterogeneous fleet in a large-scale setting under uncertain demand conditions. It is analyzed via a worst-case scenario approach and accounting for fuel consumption. Moreover, due to the problem's complexity, a hybrid metaheuristic algorithm (a genetic algorithm) was proposed to solve the mathematical model. The model ensures solution feasibility under all demand scenarios by incorporating worst-case analysis.

This study is subject to several limitations, primarily from its assumptions and modeling scope. At first, in the modeling framework, each delivery tour is assumed to be served by a single vehicle, with no provision for inter-cluster vehicle reallocation. Also, the model assumes a specific network comprising one central warehouse and many retail points, regardless of the possibility of expanding the distribution network by adding supplementary warehouses. This model treats key parameters such as inter-node distances as constant, without incorporating traffic-related constraints or their impact on travel times. Although customer demand data are inherently stochastic, the model is solved under a worst-case scenario by assuming maximum demand levels. This approach ensures solution feasibility under all potential realizations and guarantees robustness; however, it emphasizes primarily solution resilience over model flexibility. Here, the problem is formulated and solved for a single period. Incorporating temporal variations and extending the model to a multi-period setting could provide deeper insights and improve practical applicability. Finally, the model could observe explicit objectives or constraints related to customer satisfaction levels or order fulfillment rates. These concepts are increasingly important in modern logistics systems.

Given the aforementioned limitations, several promising directions for future studies are suggested: first, introducing flexibility in vehicle allocation across clusters could significantly improve fleet utilization and operational efficiency. Modeling and solving integrated location-routing problems that determine the optimal number and placement of warehouses in the network can be suggested to cover the network limitations. Designing routing algorithms that explicitly account for urban traffic conditions and their impact on travel time variability between nodes is likely another issue for future research. Also, implementing robust models based on less conservative frameworks, e.g., the Ben-Tal and Nemirovski approach, the Bertsimas and Sim methodology, or data-driven robust optimization techniques, can lead to a more accurate capture of demand uncertainty. Finally, developing multi-objective models that simultaneously consider logistical efficiency, accessibility, and customer satisfaction metrics can enhance the proposed model's adaptability to practical situations.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank all participants of the present study.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

Accorsi, L., & Vigo, D. (2024). A Fast and Scalable Heuristic for the Solution of Large-Scale Capacitated Vehicle Routing Problems. *Transportation Science*, *55*(4), 832–856. https://doi.org/10.1287/trsc.2021.1059

Alesiani, F., Ermis, G., & Gkiotsalitis, K. (2022). Constrained clustering for the capacitated vehicle routing problem (CC-CVRP). *Applied Artificial Intelligence*, *36*(1), 1995658. https://doi.org/10.1080/08839514.2021.1995658

Alzate, P., Londoño, G.A.M., Carrasco, J.M.S., Isaza, G.A., Toro, E.M., & Jaramillo-Garzon, J.A. (2024). Scientific mapping and research perspectives of the vehicle routing problem: An approach from sustainability strategies. *Sustainable Futures*, *8*, 100390. https://doi.org/10.1016/j.sftr.2024.100390.

Andrejić, M., Kilibarda, M., & Pajić, V. (2018). A framework for assessing logistics costs. *International Journal of Logistics Management*, *27*(3), 770-794.

Asgari, A., & Salehi, S. (2023). The role of food industry principles and standards in sustainable development. *7th International Conference on Agricultural Engineering and Environment with a Development Approach*, March 15, Tehran, Iran. (in Persian)

Asghari-Zadeh, E., Jafarnejad, A., Zandieh, M., Joybar, S. (2017). Explanation of Traffic Modeling Pattern in Vehicle Routing Problems Based on Green Transportation Paradigm (Case Study: Zamzam Company). *Industrial Management Journal*, *9*(2), 217–244. https://doi.org/10.22059/imj.2017.228099.1007197 (In Persian).

Baker, B.M., & Ayechew, M.A. (2003). A genetic algorithm for the vehicle routing problem. *Computers & Operations Research*, *30*(5), 787-800. https://doi.org/10.1016/S0305-0548(02)00051-5.

Braekers, K., Ramaekers, K., Van Nieuwenhuyse, I. (2016). The vehicle routing problem: state of the art classification and review. *Computers & Industrial Engineering*, *99*, 300-313. https://doi.org/10.1016/j.cie.2015.12.007.

Campbell, A. M., & Wilson, J. H. (2014). Forty years of periodic vehicle routing. *Networks*, *63*(1), 2–15. https://doi.org/10.1002/net.21527.

Cavaliere, F., Accorsi, L., Laganà, D., Musamanno, R., Vigo, D. (2024). An efficient heuristic for very large-scale vehicle routing problems with simultaneous pickup and delivery. *Transportation Research Part E: Logistics and Transportation Review*, *186*, 103550. https://doi.org/10.1016/j.tre.2024.103550.

Chen, J., Dan, B., & Shi, J. (2020). A variable neighborhood search approach for the multi-compartment vehicle routing problem with time windows considering carbon emission. *Journal of Cleaner Production*, *277*, 123932. https://doi.org/10.1016/j.jclepro.2020.123932.

Eksioglu, B., Vural, A.V., Resiman, A. (2009). The vehicle routing problem: A taxonomic review. *Computers & Industrial Engineering*, *557*(4), 1472-1483. https://doi.org/10.1016/j.cie.2009.05.009.

European Commission report (2023). Field to fork: Global food miles generate nearly 20% of all CO2 emissions from food. Available Online at: https://environment.ec.europa.eu/news/field-fork-global-food-miles-generate-nearly-20-all-co2-emissions-food-2023-01 25_en#:~:text=This%20indicates%20that%20transport%20accounts,GHG)%20released%20during%20their%20production

Ganji, M., Kazemipoor, M., Haji Molana, S.M., & Sajadi, S.M. (2020). Development of Integrated Multi-objective Green Supply Chain Scheduling Model: Production, Distribution and Heterogeneous Vehicle Routing with Customer Time Windows. *Industrial Management Journal*, *12*(1), 47 – 81. https://doi.org/10.22059/imj.2020.294629.1007697 (in Persian).

Gansterer, M., Hartl, R.F. (2018). Collaborative vehicle routing: a survey. *European Journal of Operational Research*, *268*(1), 1–12. https://doi.org/10.1016/j.ejor.2017.10.023.

Giallanza, A., & Puma, G. L. (2020). Fuzzy green vehicle routing problem for designing a three-echelon supply chain. *Journal of Cleaner Production*, *259*, 120774. https://doi.org/10.1016/j.jclepro.2020.120774.

Han, X. (2023). Path Planning Algorithm for the Multiple Depot Vehicle Routing Problem Based on Parallel Clustering. *Scientific Programming*, *1*, 7588595. https://doi.org/10.1155/2023/7588595.

Haripriya, K., Ganesan, V.K. (2022). Solving Large-Scale Vehicle Routing Problems with Hard Time Windows under Travel Time Uncertainty. *IFAC-Papers Online*, *55*(10), 233-238. https://doi.org/10.1016/j.ifacol.2022.09.394.

Hou, Q., Yang, J., Su, Y., ang, X., Deng, Y. (2023). Generalize Learned Heuristics to Solve Large-scale Vehicle Routing Problems in Real-time. *The Eleventh International Conference on Learning Representation*, 1-37.

Kazemi, M., Mohamadi Zanjirani, D., & Esmaeilian, M. (2021). The Multi-Objective Locating Model for Cross-Docking Centers and Vehicle Routing Scheduling with Split Demands for Perishable Products. *Industrial Management Journal*, *13*(4), 606–633. https://doi.org/10.22059/imj.2022.333499.1007883 (in Persian).

Koç, C., Laporte, G., and Tükenmez, I. (2020). A review of vehicle routing with simultaneous pickup and delivery. *Computers & Operations Research*, *122*, 104987. https://doi.org/10.1016/j.cor.2020.104987.

Konstantakopoulos, G.D., Gayialis, S.P., & Kechagias, E.P. (2022). Vehicle routing problem and related algorithms for logistics distribution: a literature review and classification. *Operational Research*, *22*, 2033-2062. https://doi.org/10.1007/s12351-020-00600-7.

Li, J., Fan, L., Zhang, J., & Ma, H. (2022). A three-phase framework for large-scale vehicle routing problems with real-world applications. *Computers & Industrial Engineering*, *165*, 107938. https://doi.org/10.1016/j.cie.2022.107938.

Li, M., Jia, N., Lenzen, M., Malik, A., Wei, L., Jin, Y. & Raubenheimer, D. (2022). Global food miles account for nearly 20% of total food-system emissions. *Nature Food*, *3*(6): 445–453. https://doi.org/10.1038/s43016-022-00531-w

Liu, S., Cavalcanti Costa, J.G., Mei, Y., Zhang, M. (2024). GPGLS: Genetic Programming Guided Local Search for Large-Scale Vehicle Routing Problems. In: Affenzeller, M., Winkler, S.M., Konova, A.V., Trautmann, H., Tusar, T., Machado, P., Back, T. (eds.). Parallel Problem Solving from Nature – PPSN XVIII. PPSN 2024. Lecture Notes in Computer Science, Vol 15148, Springer, Cham. https://doi.org/10.1007/978-3-031-70055-2_3.

Mangukiya, R.D., Sklarew, D.M. (2023). Analyzing the three pillars of sustainable development goals at sub-national scales within the USA. *World Development Sustainability*, *2*, 100058. https://doi.org/10.1016/j.wds.2023.100058.

Mensah, J., Casadevall, S.R. 2019. Sustainable development: Meaning, history, principles, pillars, and implications for human action: Literature review. *Cogent Social Sciences*, *5*(1), 1-21. https://doi.org/10.1080/23311886.2019.1653531.

Pandey, P.C., Pandey, M. (2023). Highlighting the role of agriculture and geospatial technology in food security and the Sustainable Development Goals. *Sustainable Development*, *31*(5), 3175–3195. https://doi.org/10.1002/sd.2600.

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science*, *360*(6392), 987-992. https://doi.org/10.1126/science.aaq0216.

Prasanna, S., Verma, P., Bodh, S. (2024). The role of food industries in sustainability transition: a review. *Environment, Development and Sustainability*. https://doi.org/10.1007/s10668-024-04642-1.

Qin, W., Zhuang, Z., Huang, Z., & Huang, H. (2021). A novel reinforcement learning-based hyper-heuristic for the heterogeneous vehicle routing problem. *Computers & Industrial Engineering*, *156*, 107252. https://doi.org/10.1016/j.cie.2021.107252.

Rajaei, M., Moslehi, G., & Reisi-Nafchi, M. (2022). The split heterogeneous vehicle routing problem with three-dimensional loading constraints on a large scale. *European Journal of Operational Research*, *299*(2), 706–721. https://doi.org/10.1016/j.ejor.2021.08.025.

Ruggerio, C.A. (2021). Sustainability and sustainable development: A review of principles and definitions. *Science of the Total Environment*, *786*, 147481. https://doi.org/10.1016/j.scitotenv.2021.147481.

Sar, K., Ghadimi, P. 2023. A systematic literature review of the vehicle routing problem in reverse logistics operations. *Computers & Industrial Engineering*, *177*, 109011. https://doi.org/10.1016/j.cie.2023.109011.

Selvan, T., Panmei, L., Murasing, K.K., Guleria, V., Ramesh, K.R., Bhardwaj, D.R., Thakur, C.L., Kumar, D., Sharma, P., Umedsinh, R.D., Kayalvizhi, D., Deshmukh, H.K. (2023). Circular economy in agriculture: unleashing the potential of integrated organic farming for food security and sustainable development. *Frontiers in Sustainable Food Systems*, *7*, 01–17. https://doi.org/10.3389/fsufs.2023.1170380.

Sevaux, M, Sörensen. K. (2008). Hamiltonian paths in significantly clustered routing problems. Proceedings of the EU/MEeting 2008 workshop on Metaheuristics for Logistics and Vehicle Routing, *EU/ME*, *8*.

Shi, J. (2024). Optimization of frozen goods distribution logistics network based on k-means algorithm and priority classification. *Scientific Reports*, *14*(1), 22477. https://doi.org/10.1038/s41598-024-72723-2.

Sommerauerova, D., Chocholac, J., Polak, M. 2018. Sustainable distribution logistics of retail chains. *Proceedings of the fourth international conference on traffic and transport engineering*, 454-459.

Stavropolo, F. 2022. The Consistent Vehicle Routing Problem with the Heterogeneous Fleet. *Computers & Operations Research*, *140*, 105644. https://doi.org/10.1016/j.cor.2021.105644.

Tao, Y., Lin, C., & Wei, L. (2022). Metaheuristics for a Large-Scale Vehicle Routing Problem of Same-Day Delivery in E-Commerce Logistics System. *Journal of Advanced Transportation*, *2022*, 1-15. https://doi.org/10.1155/2022/8253175.

Torabzadeh, S. A., Nejati, E., Aghsami, A., & Rabbani, M. (2022). A dynamic multi-objective green supply chain network design for perishable products in uncertain environments, the coffee industry case study. *International Journal of Management Science and Engineering Management*, *17*(3), 220–237. https://doi.org/10.1080/17509653.2022.2055672.

Toth, P., & Vigo, D. (Eds.). (2002). The vehicle routing problem. Philadelphia: Society for Industrial and Applied Mathematics.

Toth, P., & Vigo, D. (Eds.). (2014). Vehicle routing: problems, methods, and applications. *Philadelphia: Society for industrial and applied mathematics*.

Valriberas, D. O., Lanzi, E., Klimont, Z., & Van Dingenen, R. (2023). The economic consequences of air pollution policies in Arctic Council countries: A sectoral analysis. *Climate and Clean Air Coalition.*

Ye, H., Wang, J., Liang, H., Cao, Z., Li, Y. Li, F. 2024. GLOP: Learning Global Partition and Local Construction for Solving Large-Scale Routing Problems in Real-Time. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(18), 20284-20292. https://doi.org/10.1609/aaai.v38i18.30009.

Yağmur, E., Kesen, S.E. 2021. Multi-trip heterogeneous vehicle routing problem coordinated with production scheduling: Memetic algorithm and simulated annealing approaches. *Computers & Industrial Engineering*, *161*, 107649. https://doi.org/10.1016/j.cie.2021.107649

Zhang, M., Pratap, S., Zhao, Z., Prajapati, D., & Huang, G. Q. (2021). Forward and reverse logistics vehicle routing problems with time horizons in B2C e-commerce logistics. *International Journal of Production Research*, *59*(20), 6291–6310. https://doi.org/10.1080/00207543.2020.1812749.