Volume 16, Issue 3, 162507 (1-18)

International Journal of Industrial Chemistry (IJIC)

https://doi.org/10.57647/j.ijic.2025.1603.07

A Multi Objective Data-Driven Chemical-Aware Distribution Network Design Model Under Uncertainty (Case Problem: Bonakchi)

Mohammad Amirahmadi[1], Hamid Esmaeeli[1*],
Kia Parsa[1], Amin Mostafaee[2]

[1]*Department of Industrial Engineering, NT.C., Islamic Azad University, Tehran, Iran*
[2]*Department of Mathematics, NT.C., Islamic Azad University, Tehran, Iran*

[*]**Corresponding author:** hamidesmaeeli@iau.ac.ir

Research Article

Received:
10 March 2025

Revised:
22 August 2025

Accepted:
11 September 2025

Published in Issue:
30 September 2025

Abstract:

In this research, we aim to address the gaps in past studies by presenting a data-driven network design model for the distribution of edible oils, a critical sector with unique chemical and logistical challenges. This model begins with a thorough analysis of customer demand and the specific requirements of edible oils (e.g., perishability, storage conditions, and transportation constraints) using data mining and machine learning tools. Based on the insights gained from analyzing customer behavior, the demand amounts across different geographical areas and their changing patterns will be identified and used as inputs for the network design model. For this purpose, the KNN method will be employed for data classification and analysis, and customer demand will be estimated for network design using new dimensions. Subsequently, taking into account real-world constraints and obstacles, a new mathematical model will be developed with environmental considerations. It is worth noting that during the modeling phase, in addition to optimizing the number, location, and capacity of facilities and flow in the network, the optimization of fleet type and its composition will also be addressed. Finally, to solve the model, the multi-objective nature of the problem will first be addressed using the e-constraint method. Then, considering the model's dimensions and complexities, an appropriate solution method will be proposed, and the model will be solved using real data from the Bonakchi company, extracting management insights.

Keywords: Chemical-Aware Distribution Design, Data-mining, E-constraint method, Uncertainty, Demand management

Cite this article: Amirhmadi, M., Esmaeeli, H., Parsa, K., Mostafaee, A. (2025). A multi objective data-driven chemical-aware distribution network design model under uncertainty (case problem: Bonakchi), *International Journal of Industrial Chemistry*, 16(3). https://doi.org/10.57647/j.ijic.2025.1603.07

1. Introduction

In recent years, researchers and scientific institutions have increasingly focused on the design of distribution networks. Recently, reputable sources in supply chain management and distribution systems have addressed the design of these networks, as they play a significant role in delivering products, services, and overall responsiveness to customer needs [1]. Various factors influence the design of distribution networks and systems, including demand variations, demand density, transportation costs, distribution structure, and uncertainty [2].

The design of distribution networks becomes even more critical when handling chemically sensitive products, such as edible oils, which require specialized logistics due to their perishability, temperature sensitivity, and susceptibility to oxidation. Recent advancements in food science, such as the encapsulation of oils in nano capsules (e.g., using polycaprolactone or triblock copolymers [3], further complicate supply chain demands. These innovations, while enhancing product shelf-life, introduce new constraints—such as controlled storage conditions and fragile transportation requirements—that must be addressed in network design [4]. For instance, nano-capsule-based oils may require stricter temperature controls or reduced transit times to maintain structural integrity. This study bridges this gap by proposing a data-driven distribution model that explicitly accounts for the chemical and physical properties of edible oils, ensuring both operational efficiency and product quality [5].

The necessity of creating a distribution system using data mining tools is of great importance. With the increasing volume and diversity of data in today's world, utilizing data mining tools to analyze data and extract patterns and useful information has become essential [6]. By establishing a distribution system that incorporates data mining tools, significant improvements can be made in system performance and optimization of distribution processes. By analyzing demand data, buyer patterns, customer behaviors, and other related parameters, more accurate planning of goods and services distribution can be achieved, which enhances strategic decision-making processes in distribution [7]. Specifically, the use of data mining tools in distribution systems can lead to the optimization of warehouses, route scheduling, inventory management, demand forecasting, and greater resource efficiency. These tools can analyze precise data and employ various algorithms to enhance the efficiency and performance of distribution systems, thereby increasing customer satisfaction and reducing costs in the distribution chain [8]. The necessity of creating an integrated and responsive distribution system to demand data has been well understood, especially in light of the pandemic context and the disruptions that have occurred in distribution and supply systems. Furthermore, with the decline of vertical integration in supply chains and their globalization, unexpected events in any part of the chain can have widespread impacts [9].

In the literature on the design of distribution systems, adequate attention has not been given to the important discussion of demand patterns and the optimal location of facilities and equipment based on demand patterns and density. Considering these factors not only increases the system's flexibility but also reduces logistical and transportation costs, improving the distribution process. The optimization of transportation fleets in a distribution system is extremely important, especially when there are many products and customers [10]. Optimizing the transportation fleet can help achieve the following goals:

- Cost Reduction: By optimizing the transportation fleet, costs related to fuel, labor, maintenance, and repairs can be reduced, leading to increased efficiency.
- Reduced Delivery Time: By optimizing routes and transport scheduling, the delivery time of goods to customers can be shortened, thereby improving service levels.
- Improved Inventory Management: With the optimization of the transportation fleet, inventory levels can be optimized, reducing the need for large and additional warehouses.
- Reduced Environmental Pollution: By optimizing the transportation fleet and utilizing sustainable transport methods, environmental pollution can be decreased, contributing to the preservation of the environment.

Therefore, using modern technologies such as Geographic Information Systems (GIS), fleet type optimization, fleet capacity utilization optimization, route optimization, transport automation, and artificial intelligence can be very effective in optimizing transport fleets and improving distribution system performance [11]. In most network design models, the focus has been on the strategic location of facilities based on minimizing supply chain costs, but important factors such as demand patterns, density, and changes in customer behavior have been overlooked.

Thus, these factors need to be taken into account in the design of distribution systems to create a responsive system capable of providing better services to customers. For this reason, concurrently considering demand patterns and optimizing fleet composition can significantly bridge gaps in the literature and help design and deliver a system that is highly responsive and capable of serving more customers. The goal of this issue is to optimize performance by presenting a new design model. The use of location concepts, communication optimization, and transport fleet optimization in the proposed model will be effective and contribute to the creation of the final model. The objectives of the research can be summarized as follows:

- Providing an optimization model that considers demand variations, customer behaviors, and existing uncertainties.
- Designing and configuring a data-driven distribution system.
- Optimizing the transport fleet in the proposed distribution system.

The proposed model represents a multi-product, multi-period supply chain network. The envisioned network includes various centers and different layers. In this network, the flow of goods or items is initiated by customer orders in different cities. The flow formation starts from the beginning of the network and the points of receiving goods, continuing to the end of the network and the points of delivering goods. This research aims to address the gaps and new challenges highlighted in the existing foundational models in the literature and to model the issue effectively.

The article is organized as follows: Section 2 reviews the current literature on network design problems. In the Section 3, the proposed model is presented, followed by the Section 4, where the mathematical model is solved and determined various options. To illustrate the model application, a real case problem is implemented in Section 5. Finally, the results and outline directions for future research is discussed in Section 6.

2. Literature Review

To understand the literature on the subject and to conduct a thorough review of previous research, it is essential to briefly introduce and present the most significant and relevant studies conducted in this field. This section will address those studies.

Perl and Sirisoponsilp [12] conducted the first research on the integration of mixed decision-making models for location and transportation. Unlike the initial article, its validation was performed through numerical testing. Another notable contribution in this area is the work by Jayaraman [2], who examined the equilibrium between location, inventory, and transportation decisions in distribution network design. Barahona and Jensen [13] addressed the three-level location problem with deterministic demand. In this problem, the capacities of distribution centers and factories are unlimited, and the order quantities are infinite; additionally, the product flow is multi-product. Gabor and van Ommeren [14] studied a location problem in a distribution network with inventory review policies in 2006. The objective function included costs such as ordering, operations inside facilities, and transportation. Melo, Nickel [15] also conducted a study in 2006 focusing on the performance of a dynamic location model that encompasses various aspects, including the structure of the distribution network, inventory management, dynamic facility reassignment, and future study opportunities. While prior studies have extensively explored generic distribution networks, few address the logistical complexities of chemically sensitive products like edible oils. For example, research on nanocapsule-encapsulated oils (e.g.,[16]) highlights the need for precise environmental controls during storage and transit to prevent degradation. Later on, Snyder, Daskin [17] examined a distribution network with a continuous inventory review policy, where facility capacities are unlimited, and demand is uncertain with indefinite inventory levels. This study aimed at minimizing distribution network costs. Then, Tanonkou, Benyoucef [18] and Tang, Yang [19] investigated facility capacities and order quantities. [20] examined a supply chain that includes distribution centers and customers with uncertain demand. The goal of this chain was to minimize total costs. In this issue, distribution center capacity is unlimited, while order quantity is limited, and inventory policy is evaluated.

The article by Bogataj, Grubbström [21] presents a continuous inventory review policy, and other assumptions and costs are similar to those in the work of Chen, Li [22]. However, the main difference between these two articles lies in the method of calculating transportation costs. Jha, Somani [23] addressed the location problem without considering inventory policies and with limited capacity facilities. In this case, the demand is deterministic, and the focus is on multi-product flow, with the goal of minimizing location and transportation costs. The article by Berman, Krass [24] discussed decision-making regarding the locations of distribution centers, retailer allocation, inventory levels, and the time intervals for inventory reviews in a two-level supply chain. This issue was modeled using mixed-integer nonlinear programming and solved using the Lagrangian relaxation method. The article by Diabat, Battaïa [25] considered location, order allocation, and inventory decisions with the aim of minimizing total inventory, transportation, ordering, and facility costs. This issue involves a manufacturer, distribution centers, and retailers, determining the locations of the distribution centers. The flow is single-product, and demand is deterministic. Ahmadi-Javid and Hoseinpour [26] presented a location problem aimed at maximizing the revenue of a distribution network in a supply chain with price-sensitive demand and multi-product flow. This issue was modeled using mixed-integer nonlinear programming and solved using Lagrangian relaxation for both limited and unlimited capacity scenarios.

Gzara, Nematollahi [27] examined a comprehensive decision-making model involving two sets of decisions. The first set includes strategic decisions regarding the location of distribution centers and their allocation to customers, while the second set pertains to tactical and operational decisions about inventory levels at each distribution center. These two sets of decisions are used to address the problem simultaneously, contrasting with past approaches where they were treated separately. The primary goal of this model is to minimize total costs, including opening distribution centers, transportation, and inventory holding costs. The article by Shahabi, Unnikrishnan [28] studied a three-level inventory location problem concerning the relationship between demand. Decisions examined include the location of factories and wholesalers, the allocation of wholesalers to factories, and retailers to wholesalers, with an evaluation of safety stock levels and inventory held at each wholesaler.

Regarding urban rapid service systems, vehicle capacities and hub capacity have been examined Wu, Qureshi [29]. To reduce costs, routing and location-inventory problems in a three-level supply chain for perishable products in China have been analyzed through a mixed-integer linear programming (MILP) model; however, service levels were not considered in this model [30].

Furthermore, for waste management in a two-level supply chain, Caramia and Pizzari [31] addressed the location-allocation problem. Subsequently, Fahimi and colleagues utilized a mixed-integer nonlinear programming (MINLP) model to study location-allocation facility problems in a supply chain to reduce costs for perishable products. Other research has addressed location-allocation issues in various fields, including disaster response, healthcare, food supply chains, and facilities location planning. For example, Fahmy, Zaki [32] proposed a model for designing emergency medical service networks including multiple facility locations and multi-transportation modes. Additionally, Hasani Goodarzi, Zegordi [33] explored the location of transit warehouses through a MINLP model in an automotive company in the Middle East. Behnamian, Fatemi Ghomi [34] studied the location problem regarding the allocation and routing of trucks at multiple terminals with the goal of cost reduction; however, their two-stage model did not consider time windows for delivering goods to customers. Simultaneously, Theophilus, Dulebenets [35] addressed truck planning in a transit warehouse related to a cold chain to minimize costs, yet did not account for temporary storage capacity limitations in their proposed model.

Vincent, Aloina [36] examined the simultaneous collection and delivery of goods to minimize total travel costs and penalties imposed on drivers, utilizing a mixed-integer linear programming model. In conjunction with this, to reduce costs of transferring goods between distribution centers, the location-allocation and routing problem considering simultaneous collection and delivery of products has been studied [37]. Moreover, the issue of green location and routing, accounting for budget constraints and simultaneous collection and delivery of products, aims to reduce the loss of time and total network costs [38].

Additionally, Shi, Lin [39] investigated the simultaneous collection and delivery of medical supplies via drones during the COVID-19 pandemic to expedite delivery times and reduce patient contact. Concurrently, Hosseini-Motlagh, Farahmand [40] researched vehicle routing to cover traffic and optimize speed on each route while considering simultaneous collection and delivery of products. Lastly, a precise algorithm for the problem of optimizing vehicles for simultaneous collection and delivery is proposed by Che and Zhang [41], Parilina, Yao [42] where customer demands for collection are assumed to be random.

In this research, an effort has been made to draw insights from the existing literature, particularly the studies by Jayaraman [43], Yang, Ng [44], and Ağralı, Geunes [45], which were among the first to focus on distribution network design. The main characteristics and key issues of distribution network design have been understood, and then based on the research conducted by Selim and Ozkarahan [46], Maihami, Kannan [47] the changes in customer behavior and demand pattern analysis were taken into account. Initially, the application of data mining tools for demand analysis is implemented, and then, considering the demand situation, the design of the distribution system is examined using various distribution approaches based on the type of demand.

This approach is distinguished from previous research due to the existing articles and reviewed literature, bringing the proposed models in distribution system design closer to real-world conditions. Upon reviewing the existing literature on distribution system design, it becomes evident that the optimization of fleet composition and quantity alongside data-driven network design has either not been adequately addressed or has received limited attention. In most conducted studies, it has been assumed that distribution systems operate with either deterministic or stochastic demand, and analysis of the demand pattern has not been performed. In cases where uncertainty has been considered, the focus has predominantly been on fleet optimization, with less attention given to the simultaneous combination of both aspects. However, if this issue can be examined simultaneously, and a data-driven network is properly designed, it could have significant impacts on meeting service levels and enhancing customer satisfaction.

In this research, a detailed study and analysis of customer demand using data mining tools will be conducted first. Then, to respond to customer demands, a mathematical programming model for the design of the distribution system will be developed, incorporating fleet optimization. Subsequently, precise mathematical tools will be employed to solve the model. The developed model will be implemented on real data from the online distribution system of Bankchi, and the results will be reported to derive suitable managerial insights. In summary, the innovations of the present research include:

- Proposing an optimized distribution network design tailored to the problem's constraints.
- Utilizing a data-driven approach for analyzing demand in the distribution network.
- Developing an appropriate approach to address the uncertainties present in the network.
- Providing an optimal model for the number and type of transport fleet for the designed network.

Table 1. Brief data on the most relevant papers in literature

Author(s)	Customer behavior	Product		Time period		Network design decisions				Fleet optimization		Objective		Uncertainty				Data-driven	Non-data-driven
		Multi	Single	Multi	Single	Location	Number of facilities	Capacity planning	Time of establishment	Mode	Number	Multi	Single	Stochastic	Fuzzy	Mixed uncertainty	Deterministic		
Melo, Nickel [15]		*			*	*		*					*				*		*
Snyder, Daskin [17]		*		*		*							*	*					*
Tang, Yang [19]		*		*		*	*						*				*		*
Chen, Li [22]			*		*	*							*	*					*
Jha, Somani [23]		*			*			*					*				*		
Berman, Krass [24]			*		*	*							*				*		*
Shahabi, Unnikrishnan [28]			*		*	*							*	*					*
Diabat, Battaïa [25]			*		*	*							*				*		*
Ahmadi-Javid and Hoseinpour [26]		*	*			*							*	*					*
Kartal, Hasgul [37]			*		*								*				*		*
Behnamian, Fatemi Ghomi [34]			*		*								*				*		*
Hasani Goodarzi, Zegordi [33]			*		*	*							*				*		*
Theophilus, Dulebenets [35]			*	*						*	*		*				*		*
Shi, Lin [39]			*		*	*							*				*	*	
Hosseini-Motlagh, Farahmand [40]		*			*					*	*		*	*				*	
Wu, Qureshi [29]			*		*	*							*				*		*
Caramia and Pizzari [31]			*		*	*							*			*			*
Song and Wu [30]		*		*		*	*			*							*		*
Nasiri, Mousavi [38]		*		*		*				*			*				*		*
Vincent, Aloina [36]			*		*					*	*		*				*		*
Che and Zhang [41]			*		*					*	*		*	*					*
Gao, Lu [48]			*	*		*							*	*				*	
Ziari [49]	*	*		*									*	*				*	
This paper	*	*		*		*	*	*	*	*	*	*			*			*	

3. Problem Definition

In this project, the design of a supply and distribution network for a company is examined and presented. This network is responsible for receiving various items from suppliers and delivering them to consumers. The products within this network include small items such as mobile phones and large items such as refrigerators. Given the nature of these items and the conditions for storage and transportation, certain sections of the network, such as processing centers, operate separately for each of these two groups, while other parts of the chain (receiving and distribution sections) are managed collectively. Fig. 1 illustrates the structure of the proposed distribution network, which includes suppliers of items, processing and packaging centers, sorting centers, fixed distribution centers, mobile distribution centers, and customers.

It is noteworthy that these products are delivered to fixed distribution centers and mobile distribution centers through various methods, such as the company's transportation fleets or vendors' fleets, and then forwarded to processing centers. After processing and transforming the items into parcels, they are sent to sorting centers and subsequently delivered to fixed and mobile distribution centers or postal offices for customer delivery. Distribution centers are also responsible for ensuring that parcels reach the customers. The objective of this design problem is to establish a network that includes the location and capacity planning of facilities, determining the nature of their connections, and specifying the flow between facilities, with the goal of minimizing costs while maintaining service level time.

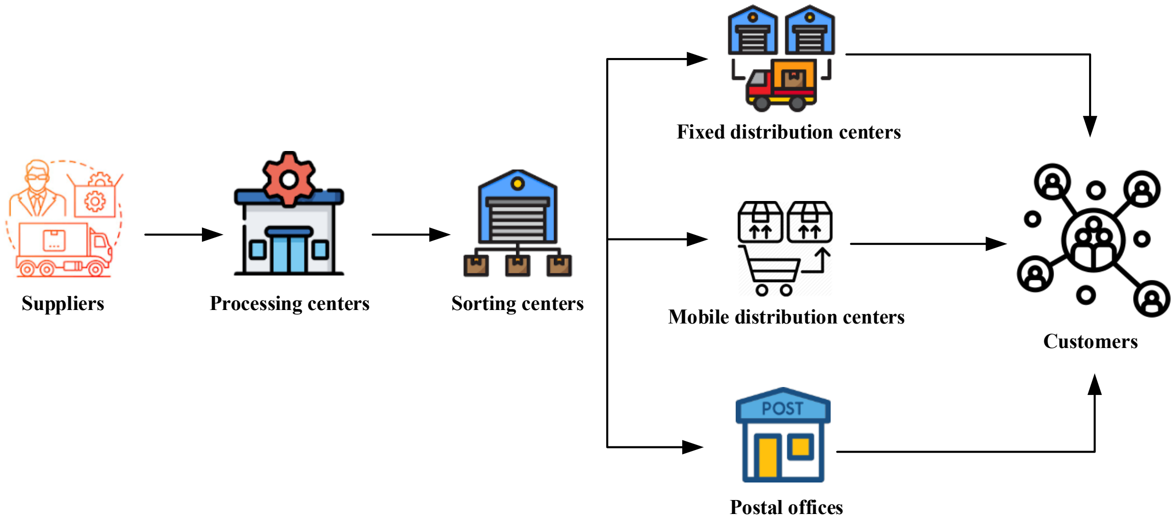

Figure. 1. Proposed distribution network structure

The model presented in this project is a multi-product, multi-period supply chain network. The envisioned network comprises fixed distribution centers, mobile distribution centers, processing centers, and sorting centers. The flow of products or items in the network is initiated by customer orders from various cities. The flow formation continues from the beginning of the network and points of receiving goods to the end of the network and points of delivery. In this network, vendors in different cities can supply the required items for each order to the mobile centers, which are responsible for receiving and delivering within the network, or to processing and sorting centers. In this network, it is essential to implement measures to adequately respond to customer demands; therefore, demand patterns and changes in trends and customer clustering must be defined to effectively plan the distribution network using the obtained patterns. Consequently, data mining tools will be used to explore demand, uncover its patterns, and analyze customer behavior. Additionally, it is important to note that environmental considerations and pollution reduction are as significant as economic goals for the company, necessitating the minimization of carbon dioxide emissions resulting from the transportation of products between different levels of the network. Moreover, within the proposed network, alongside important network decisions, such as:

- The location of each facility
- The annual capacity of each facility
- The active/inactive status of each facility in different years
- The timing of capacity increases or decreases for each facility
- The connections between facilities in the network

It is also necessary to optimize the composition of the fleet and the appropriate number of each type of fleet in the last-mile connections. Therefore, by considering customer behavior and discovering demand patterns, various customer demand clusters will be considered as inputs for optimizing the distribution network design. Then, using a bi-objective mathematical model, network design will be optimized with both economic and environmental objectives while simultaneously optimizing the fleet.

Regarding the uncertainty present in some parameters, this will be explained in subsequent sections. The key assumptions related to network design, which stem from policies and priorities of the problem, environmental constraints, and operational requirements, are as follows:

- The planning horizon for network design is considered limited.
- The location of facilities occurs discretely at the city level.
- Decision-making regarding the merging of facilities introduced in a city will be based on factors such as cost and facility capacity.
- The network design is conducted integrally for both small and large item distribution.
- The capacity of all facilities is determinable and defined and selected in increments.
- The selection and establishment of facilities are conducted sequentially to perform the four main functions of "receiving," "processing," "sorting," and "distributing."
- A maximum allowable time distance between pairs of different facilities (to meet service level times for customers) is defined. Therefore, allocating one facility to another will be contingent upon their placement within the defined permissible time distance.
- A facility can be allocated to one or multiple other facilities.
- Shipping to customers for large or small items is done in a single shipment.
- The use of different fleets alters the time between facilities; thus, the selection of the fleet and optimal patterns between facilities is aimed at minimizing costs while ensuring adequate responsiveness.
- Fleet capacities vary based on the type of item or parcel.
- Facilities such as distribution centers incur construction costs, while contractor facilities do not have startup costs depending on the type of contract.

3.1. Mathematical model

The model includes the following sets, parameters, and decision variables.

Sets

i	Set of facilities, $i = \{1, \ldots I\}$
j	Set of capacity levels, $j = \{1, \ldots, J\}$
k	Set of city nodes, $k = \{1, \ldots, K\}$
o	Set of products, $o = \{1, \ldots, O\}$
l	Set of vehicles, $l = \{1, \ldots, L\}$
t	Set of time periods, $t = \{1, \ldots, T\}$

Parameters

$C_{kk't}^{ii'}$	Transportation costs per trailer from facility *i* in city *k* to facility *i'* in city *k'* in the time period *t*
A_{it}^{o}	Transportation costs per parcel of product type *o* from facility *i* to customer in time period *t*
$D_{kk't}^{o}$	Demand of product *o* in from city *k* to city k' in the time period *t*
M_{ijkt}	Mortgage cost of the facility *i* with capacity level *j* in city *k* in the time period *t*
R_{ijkt}	Rent cost of the of the facility *i* with capacity level *j* in city *k* in the time period *t*
F_{ijt}	Cost of equipping capacity level *j* of the processing facilitation type *i* in time period *t*
ca_{ij}	Capacity level *j* of facility *i*
$cap_l{}^o$	Capacity of vehicle *l* based on product *o*
$\tilde{T}_{ii'}$	Maximum transfer time between facility *i* and *i'*
$d_{lkk'}$	Time distance from origin city *k* to destination city *k'* when transporting with vehicle type *l*
$I_{lkk't}$	Transporting cost from origin city *k* to destination city *k'* when transporting with vehicle type *l* in the time period *t*
Em_l^o	Emission rate of vehicle type *l* per unit of parcel type *o*
O_{lt}	Carbon neutralizing costs of vehicle type *l* in the time period *t*
H_{ik}	1, If it is possible to establish facility *i* in the city *k;* 0 otherwise.
SL	Service level
i_1	Inflation rate
i_2	Interest rate
i_3	Depreciation rate
ε	Epsilon constraint coefficient
M	Big M

Decision Variable

X_{ijkt}^{o}	1, If facility *i* with capacity level *j* in the city *k* is opened in time period *t*; 0 otherwise.
$Z_{kk't}^{oii'}$	1, If facility *i* in the city *k* is assigned to facility *i'* in the city *k'* in time period *t*; 0 otherwise.
$H_{kk't}^{oii'}$	Quantity of product type *o* shipped from facility *i* in the city *k* to facility *i'* in the city *k'* at time period t.
$G_{kk't}^{oi}$	Quantity of product type *o* shipped from facility *i* in the city *k* to customer in the city *k'* at time period t.
$S_{lkk't}$	Required number of vehicle type *l* for transporting from the city *k* to the city *k'* at time period t.

Now, the mathematical model can be developed as follows:

$$
\begin{aligned}
Z_1 \quad &= min \sum_{t} \frac{1}{(1+i_1)^t} [\sum_{i}\sum_{j}\sum_{k} \big(M_{ijkt} \times i_2 \\
&+ R_{ijkt} + F_{ijt} \times i_3\big). X_{ijkt} \\
&+ \sum_{o}\sum_{i}\sum_{k}\sum_{k'}\sum_{t} A_{it}^{o} G_{kk't}^{oi} \\
&+ \sum_{o}\sum_{i}\sum_{i\prime}\sum_{k}\sum_{k'}\sum_{l}\sum_{t} C_{kk't} \times \frac{H_{kk't}^{oii\prime}}{cap_l{}^{o}} \qquad (1) \\
&+ \sum_{l}\sum_{k}\sum_{k'}\sum_{t} S_{lkk't}. I_{lkk't} \\
&+ \sum_{l}\sum_{k}\sum_{k'}\sum_{t} S_{lkk't}. O_{lt}]
\end{aligned}
$$

$$Z_2 \quad = min \sum_{o}\sum_{i}\sum_{k}\sum_{k\prime}\sum_{l}\sum_{t} Em_l^o \,. H_{kk't}^{oii'} \qquad (2)$$

s.t.

$$\sum_{j} X_{ijkt} \le H_{ik} \qquad (\forall i,k,t) \qquad (3)$$

$$d_{lkk'}. Z_{kk't}^{oii\prime} \le \tilde{T}_{ii\prime}. \sum_{j} X_{ijkt} \qquad (\forall o,l,i,i',k,k',t) \qquad (4)$$

$$\sum_{k} Z_{kk't}^{oii\prime} \le 1 \qquad (\forall o,i,i',k',t) \qquad (5)$$

$$H_{kk't}^{oii\prime} \le M. Z_{kk't}^{oii\prime} \qquad (\forall o,i,i',k,k',t) \qquad (6)$$

$$\sum_{i,k} G_{kk't}^{oi} = \sum_{k} D_{kk't}^{o} \qquad (\forall k',o,t) \qquad (7)$$

$$\sum_{i,k} H_{kk't}^{oii\prime} = \sum_{k''} G_{k'k''t}^{oi\prime} \qquad (\forall i',k',o,t) \qquad (8)$$

$$\sum_{i\prime,k'} H_{kk't}^{oii\prime} \le \sum_{j} ca_{ij} X_{ijkt} \qquad (\forall o,i,k,t) \qquad (9)$$

$$\sum_{o}\sum_{i}\sum_{k}\sum_{k'} G_{kk't}^{oi} \ge SL. \sum_{o}\sum_{k}\sum_{k'} D_{kk't}^{o} \quad (\forall t) \qquad (10)$$

$$\sum_{l} cap_l{}^{o} S_{lkk't} > H_{kk't}^{oii'} \qquad \forall(i,i',k,k',o,t) \qquad (11)$$

$$S_{lkk't} \in Z \qquad \forall(l,k,k',t) \qquad (12)$$

$$X_{ijkt}, Z_{kk't}^{oii'} \in \{0,1\} \qquad (\forall o,i,i',j,k,k',t) \qquad (13)$$

$$H_{kk't}^{oii'}, G_{kk't}^{oi} \in R^{+} \qquad (\forall o,i,i',k,k',t) \qquad (14)$$

The objective function (1) refers to minimizing the costs of establishment, equipment, transportation, distribution and neutralizing carbon emissions through various methods, as well as risk costs. The objective function (2) refers to minimizing emitted carbons in the transportation process. Constraints (3) ensure that facilities are established in cities where construction is feasible. Constraints (4) reflect the maintenance of coverage radius requirements for facility allocation. Constraints (5) illustrate how facilities are allocated to one another while maintaining allocation requirements. Constraints (6) indicate that flow can only take place in the network if allocation has been made beforehand. Constraints (7) represent the coverage constraints for urban demand. Constraints (8) show the flow balance in the proposed distribution network. Constraints (9) indicate the capacity of the facility that must be established. Constraints (10) ensure that the service level of the system is maintained. Constraints (11) show how the number of vehicles is determining based on the transported parcels and vehicles' capacities. Constraints (12)-(14) define the domain of the decision variables.

4. Solution Method

To effectively address the proposed problem, it is essential to recognize that the challenge at hand is a multi-objective optimization scenario. This situation necessitates the transformation of the original multi-objective framework into a single-objective problem, which can be accomplished through various multi-criteria decision-making (MCDM) methods. This transformation simplifies the decision-making process, thereby facilitating a more straightforward analysis and implementation of solutions.

Before this integration occurs, it is crucial to identify and thoroughly analyze customer demand patterns and the geographical distribution of these demands. This step is vital as it provides the foundational data required for effective decision-making. A data-driven approach should be adopted to harness relevant historical and real-time data, uncovering trends, fluctuations, and correlations in customer demands across different regions. By employing sophisticated analytical techniques such as data mining, a detailed understanding of how demand varies spatially and temporally can be developed.

Once a clear understanding of the demand patterns and geographical dispersions is established, the e-constraint method can be applied to effectively tackle the bi-objective problem. This method involves prioritizing one objective while treating the others as constraints, enabling the exploration of trade-offs between conflicting objectives. By systematically adjusting the epsilon values, solutions can be ~~identified~~ that not only satisfy the primary objective but also adhere to acceptable levels of the secondary objectives. Finally, a possibilistic programming approach can be incorporated to address the inherent uncertainty and fuzziness present in the parameters of the model. Many real-world problems are characterized by uncertainty due to unpredictable market conditions, fluctuating customer preferences, or other unforeseen factors. Possibilistic programming allows for a more effective modeling of these uncertainties by considering various scenarios and their likelihoods. This approach aids in formulating solutions that are robust and adaptable to changes, thereby enhancing the overall resilience and reliability of the decision-making framework.

4.1. Clustering with K-means algorithm

In this section, a paradigm shift in demand analysis is explored, moving away from traditional structural modifications and enumeration methods. These conventional approaches typically rely on static optimization techniques and impose restrictive constraints that fail to adequately reflect consumer behavior or fully utilize features in pricing analysis. The proposed methodology is designed to autonomously identify distinct demand segments tailored to individual retailers by leveraging historical purchase data or patterns of demand [50].

An outstanding characteristic of this model is its adaptability, which allows it to respond fluidly to minor fluctuations in demand as well as unforeseen events, such as global crises or economic downturns. Such flexibility is crucial in today's fast-paced and ever-changing market landscape, where consumer behavior can shift dramatically in response to external factors [51]. The ability to adjust pricing strategies quickly ensures that retailers can maintain operational effectiveness, even amidst unpredictable circumstances. Furthermore, the model incorporates a dynamic framework that grants real-time insights into demand trends and optimization tactics. This aspect is particularly noteworthy as it empowers retailers to engage in ongoing adjustments based on temporal shifts in demand [52].

For implementing the K-means algorithm, the following steps are conducted as can be seen in Fig. 2.

Step 1: Deciding how many clusters is needed to divide data into.

Step 2: Randomly selecting K data points from dataset as the initial centroids.

Step 3: Assigning Data Points to Clusters:
- For each data point, calculating the distance to each centroid.
- Assigning each data point to the cluster whose centroid is the closest.

Step 4: Updating Centroids:
- After all data points have been assigned to clusters, recalculating the positions of the centroids by taking the average of all points in each cluster.

Step 5: Repeating Steps 3 and 4.

Step 6: Output the Clusters:
- Once convergence is achieved, the final clusters and their centroids are output

To sum up, this innovative approach revolutionizes the way demand analysis and prediction are developed and implemented. By embracing a data-driven, autonomous model that accurately reflects consumer behavior and adapts to market dynamics, retailers can implement tailored network structure that led to more effective demand segmentation and ultimately enhanced profitability. As the retail landscape continues to evolve, the adoption of such sophisticated methodologies will be critical in maintaining a competitive edge. Retailers who leverage these advanced analytical techniques can not only respond to current market demands but also anticipate future trends, ensuring long-term success in a volatile economic environment and more responsive network structure. By continuously applying the algorithm over various time intervals and making iterative adjustments to reflect changing demand dynamics and consumer behaviors, the model clusters the geographical demands as an input for the distribution network design problem. Developed in Python, the results of this approach are presented in the following section, where they will be compared with traditional methods. This comparative analysis aims to provide valuable insights into the effectiveness and advantages of the proposed methodology.

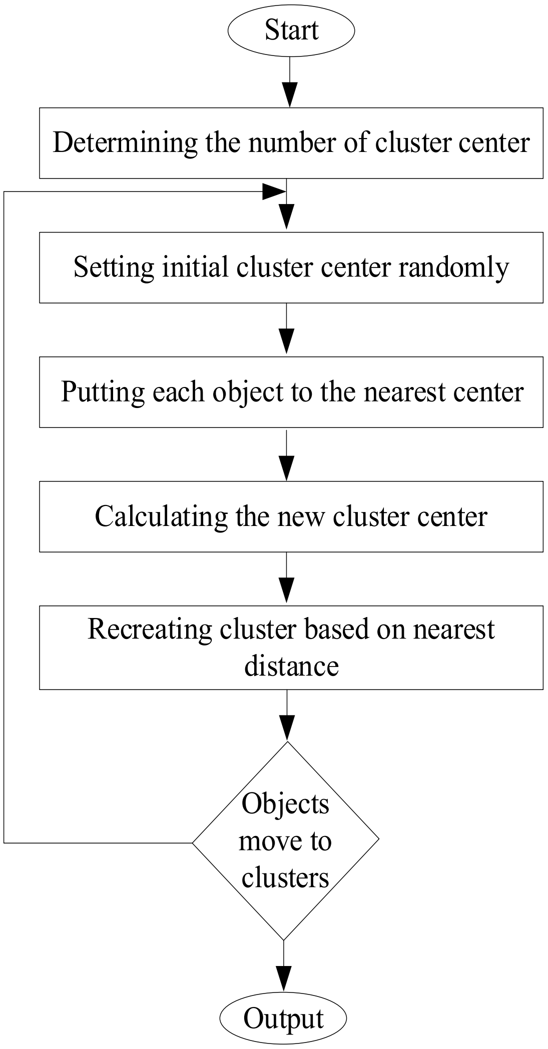

Figure 2. Flowchart of K-means algorithm [53]

4.2. The e-constraint method

Over the past few decades, the field of multi-objective optimization has experienced remarkable advancements, primarily driven by the rise of evolutionary algorithms and a range of innovative techniques. Various strategies, including the weighted sum method, the e-constraint method, scalarizing methods, value function approaches, and goal programming, have been developed to simplify multi-objective problems into single-objective frameworks [54]. Among these strategies, the e-constraint method stands out due to its ability to handle scaling issues in objective functions more effectively than the weighted sum method. This method allows for greater control over the quantity and quality of potential solutions available [55].

In application, the e-constraint method is typically formulated mathematically, allowing practitioners to set specific ranges for the objective functions to ensure that each function remains within defined limits [56]. Essentially, this technique focuses on minimizing one primary objective while placing constraints on others through additional parameters ($\mathcal{E}$). Moreover, the adaptability of the e-constraint method makes it particularly valuable in real-world scenarios where decision makers face complex trade-offs among competing objectives [57]. Its flexibility allows for tailoring the constraints based on varying priorities, providing a systematic approach to achieving optimal outcomes while maintaining the desired balance among multiple objectives. As a result, this method has gained traction across diverse fields, from engineering design to economic modeling, where understanding the nuanced interplay of different objectives is critical [58]. In the proposed model e-constraint method is used as follows. The first objective function is set while the second objective function are constrained using ($\mathcal{E}$).

$$\begin{aligned} & MinZ_1(x) \\ & s.t. \qquad\qquad \forall i \neq k, x \in X \\ & Z_2(x) \leq \varepsilon \end{aligned} \tag{1}$$

It must be noticed that the optimal solution follows the below theorems:

Theorem 1: The Pareto solution is $x^* \in X, \forall k \in \{1,\dots,K\}, i \in \{1,\dots,K\}/\{K\}$for the model (1).

Theorem 2: If $x^* \in X$ specifies a unique solution for some k with $\varepsilon_i = Z_i(x^*), \forall i \in \{1,\dots,K\}/\{K\}$, will the Pareto solution of the model (1)

Theorem 3: For any delivered upper bound vector $\varepsilon = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k\}.x^* \in X$ is the Pareto solution of the model (1).

To effectively implement the e-constraint method within the proposed model, it is essential to define the permissible range of variation for the solutions by setting specific epsilon value. This approach involves minimizing one primary objective function while constraining the other objective function through supplementary constraints. In this model, we focus on two objective functions (Z1, Z2). The primary objective function is maintained as Z1, while Z2 is converted into a constraint with an upper limit denoted by ($\mathcal{E}$). As a result, the bi-objective problem is reformulated into a singular objective model, allowing for a streamlined optimization process.

$$\begin{aligned} Z_1 \quad & = min \sum_t \frac{1}{(1+i_1)^t} [\sum_i \sum_j \sum_k (M_{ijkt} \\ & \times i_2 + R_{ijkt} + F_{ijt} \times i_3). X_{ijkt} \\ & + \sum_o \sum_i \sum_k \sum_{k'} \sum_t A_{it}^o G_{kk't}^{oi} \\ & + \sum_o \sum_i \sum_{i\prime} \sum_k \sum_{k'} \sum_l \sum_t C_{kk't} \\ & \times \frac{H_{kk't}^{oii\prime}}{cap_l{}^o} + \sum_l \sum_k \sum_{k'} \sum_t S_{lkk't}. I_{lkk't} \\ & + \sum_l \sum_k \sum_{k'} \sum_t S_{lkk't}. O_{lt}] \end{aligned} \tag{16}$$

$$\sum_o \sum_i \sum_k \sum_{k\prime} \sum_l \sum_t Em_l^o . H_{kk't}^{oii'} \leq \varepsilon \tag{17}$$

Constraints (3)-(14).

4.3. Possibilistic programming approach

Numerous approaches have been explored in previous studies for addressing possibility models that incorporate imprecise coefficients within both the objective function and constraints [59]. Among these, a particularly effective synthesis of computational techniques for resolving fuzzy linear problems emerges from the work of [60]. Their methodology primarily builds upon the foundational model, which adeptly maintains the linear nature of the problem while avoiding any increase in the number of objective functions or the complexity of constraint inequalities. This approach is versatile, demonstrating applicability across a wide range of membership functions, including triangular and trapezoidal forms, as well as other non-linear scenarios. Its robustness stems from strong mathematical principles, such as the expected value range and the expected value of fuzzy numbers [61]. Ultimately, the integration of these methodologies not only enhances the understanding of fuzzy linear programming but also opens new avenues for research, particularly in complex systems that require nuanced decision-making strategies in the face of uncertainty. As the field continues to evolve, the foundational work laid by these researchers plays a crucial role in developing innovative solutions for real-world applications. Hence, implementing interval and expected value of a fuzzy number definition, the deterministic model constraint can be expressed in the following form.

$$d_{lkk'}. Z_{kk't}^{oii\prime} \leq \tilde{T}_{ii\prime}. \sum_j X_{ijkt} \tag{18}$$

$$\begin{aligned} & d_{lkk'}. Z_{kk't}^{oii'} \leq [\alpha(\frac{T_{ii'}^P + T_{ii'}^m}{2}) \\ & +(1-\alpha)(\frac{T_{ii\prime}^P + T_{ii\prime}^m}{2})]. \sum_j X_{ijkt} \end{aligned} \tag{19}$$

Now, the model can be easily solved. Therefore, the single-objective deterministic model is solvable in the following form.

$$\begin{aligned} Z_1 &= min \sum_t \frac{1}{(1+i_1)^t} [\sum_i \sum_j \sum_k \big(M_{ijkt} \times i_2 \\ &+ R_{ijkt} + F_{ijt} \times i_3\big).X_{ijkt} \\ &+ \sum_o \sum_i \sum_k \sum_{k'} \sum_t A_{it}^{o} G_{kk't}^{oi} \\ &+ \sum_o \sum_i \sum_{i\prime} \sum_k \sum_{k'} \sum_l \sum_t C_{kk't} \times \frac{H_{kk't}^{oii\prime}}{cap_l{}^o} \\ &+ \sum_l \sum_k \sum_{k'} \sum_t S_{lkk't}.I_{lkk't} \\ &+ \sum_l \sum_k \sum_{k'} \sum_t S_{lkk't}.O_{lt}] \end{aligned} \tag{20}$$

s.t.

$$\sum_o \sum_i \sum_k \sum_{k\prime} \sum_l \sum_t Em_l^o \,.H_{kk't}^{oii'} \leq \varepsilon \tag{21}$$

$$\sum_j X_{ijkt} \leq H_{ik} \qquad (\forall i,k,t) \tag{22}$$

$$\begin{aligned} & d_{lkk'}.Z_{kk't}^{oii'} \leq \\ & [\alpha(\frac{T_{ii'}^P + T_{ii'}^m}{2}) \qquad (\forall o,l,i,i',k,k',t) \\ & +(1\ -\alpha)(\frac{T_{ii\prime}^P + T_{ii\prime}^m}{2})].\sum_j X_{ijkt} \end{aligned} \tag{23}$$

$$\sum_k Z_{kk't}^{oii\prime} \leq 1 \qquad (\forall o,i,i',k',t) \tag{24}$$

$$H_{kk't}^{oii\prime} \leq M.Z_{kk't}^{oii\prime} \qquad (\forall o,i,i',k,k',t) \tag{25}$$

$$\sum_{i,k} G_{kk't}^{oi} = \sum_k D_{kk't}^o \qquad (\forall k',o,t) \tag{26}$$

$$\sum_{i,k} H_{kk't}^{oii\prime} = \sum_{k''} G_{k'k''t}^{oi\prime} \qquad (\forall i',k',o,t) \tag{27}$$

$$\sum_{i\prime,k'} H_{kk't}^{oii\prime} \leq \sum_j ca_{ij}X_{ijkt} \qquad (\forall o,i,k,t) \tag{28}$$

$$\begin{aligned} & \sum_o \sum_i \sum_k \sum_{k'} G_{kk't}^{oi} \\ & \geq SL.\sum_o \sum_k \sum_{k'} D_{kk't}^o \quad (\forall t) \end{aligned} \tag{29}$$

$$\sum_l cap_l{}^o S_{lkk't} > H_{kk't}^{oii'} \qquad \forall(i,i',k,k',o,t) \tag{30}$$

$$S_{lkk't} \in Z \qquad \forall(l,k,k',t) \tag{31}$$

$$X_{ijkt}, Z_{kk't}^{oii'} \in \{0,1\} \qquad (\forall o,i,i',j,k,k',t) \tag{32}$$

$$H_{kk't}^{oii'}, G_{kk't}^{oi} \in R^+ \qquad (\forall o,i,i',k,k',t) \tag{33}$$

5. A Numerical Case Example

Bonakchi online wholesale is the first specialized platform for the bulk supply of fast-moving and packaged goods, including edible oils, which require careful handling due to their chemical properties (e.g., susceptibility to oxidation, temperature sensitivity, and shelf-life constraints). Bonakchi is the largest online distributor of edible oils and other essential goods in Iran, making it an ideal case for testing our data-driven distribution network model. The unique challenges of edible oils—such as maintaining quality during transit, minimizing exposure to adverse environmental conditions, and ensuring timely delivery to prevent spoilage—are explicitly addressed in our network design.

The edible oils distributed by Bonakchi exhibit specific chemical features, such as varying viscosity, flash points, and storage requirements, which directly influence the design of the distribution network. For instance, oils with higher perishability require shorter transportation times and specialized storage facilities, while others may need temperature-controlled environments. Our model incorporates these factors by optimizing facility locations, fleet composition, and routing to ensure product integrity and minimize waste. The real-world data from Bonakchi's edible oil supply chain validates the practicality of our approach, demonstrating its ability to handle sector-specific challenges.

This distributor has been operating in this sector in Iran for almost 6 years, providing and supplying goods in bulk. To evaluate the applicability of the proposed model and its feasibility, data from Bankchi has been utilized, with any unavailable information randomly generated. The proposed model was solved using the CPLEX solver from the GAMS optimization package. The developed structure includes 10 suppliers, 100 customers, 4 processing center, 4 sorting center, 20 fixed distribution center, 30 mobile distribution center, 50 types of products and 4 time periods.

5.1. Model validation

To thoroughly assess the accuracy and validity of the model and ensure that all components function correctly, a specific parameter will be focused on, and necessary adjustments will be implemented to test the anticipated outcomes. In this analysis, the primary emphasis will be on transportation costs. It is understood that when the cost of transporting products within the system rises, there is a natural tendency for the system to resist increasing the number of vehicles. This behavior occurs primarily to avoid escalating expenses associated with last-mile delivery. Consequently, it is hypothesized that an increase in last-mile transportation costs will lead to a reduction in the number of vehicles utilized. If the model accurately reflects this relationship, it can be deemed to be functioning correctly. To validate this hypothesis, a series of sensitivity analyses specific to the parameter of last-mile transportation costs ($I_{ikk't}$) will be conducted.

Results from these analyses which can be seen in Table 2, indicate that as transportation costs for last-mile delivery increase, a consistent decrease in the total number of vehicles is observed across all examined periods. This finding not only supports the initial expectations but also reinforces the validity of the proposed model. By demonstrating a clear correlation between rising transportation costs and a reduction in vehicle numbers, it can be confidently asserted that the model accurately captures the dynamics of the system under investigation. In conclusion, the findings underscore the robustness of the model, showcasing its ability to respond appropriately to changes in key parameters while maintaining logical consistency in operational outcomes. Further investigation into other influencing factors may provide additional insights and enhance the model's applicability in real-world scenarios.

Table 2. Validation test on last mile transportation costs

Experiments	Last mile transportation cost	Time periods				Total number of vehicles
		1	2	3	4	
1	37.450	290	291	290	294	1165
2	43.067	281	279	283	282	1125
3	49.527	269	270	271	266	1076
4	56.956	254	253	254	255	1016
5	65.500	241	241	240	241	963

5.2. Sensitivity analyses

In this section, the impact of important parameters of the model on its behavior and outputs is examined, aiming to assist managers, decision-makers, and retailers in predicting inputs and potential scenario changes. Key parameters, such as maximization of profit, minimization of CO2 objectives, processing centers' capacity, and service level, will be investigated within the proposed model. Measuring changes and their impact on the model's behavior can provide valuable insights for demonstrating expected values.

5.2.1. *Examining the profit maximization versus CO2 emissions*

In the proposed model, we focus on two critical objectives: maximizing profit and minimizing CO2 emissions. While the primary goal of maximizing profit enables the retailer to cover operational expenses, minimizing CO2 emissions underscores the importance of social responsibility and environmental sustainability. To analyze the trade-offs between these objectives, we first establish the Pareto front for the multi-objective problem. This involves solving the optimization problem solely for profit to identify the best possible solution. Subsequently, we reintroduce the profit objective alongside a CO2 minimization function to refine our optimal solution.

However, it's crucial to understand that when both objectives are considered simultaneously, the optimal solutions for profit and emissions derived independently are no longer valid. This necessitates a nuanced approach since the Pareto frontier illustrates varying profit levels corresponding to different waste levels. Notably, if the system disregards the impact of CO2, it may achieve substantial profits; however, this scenario is practically infeasible.

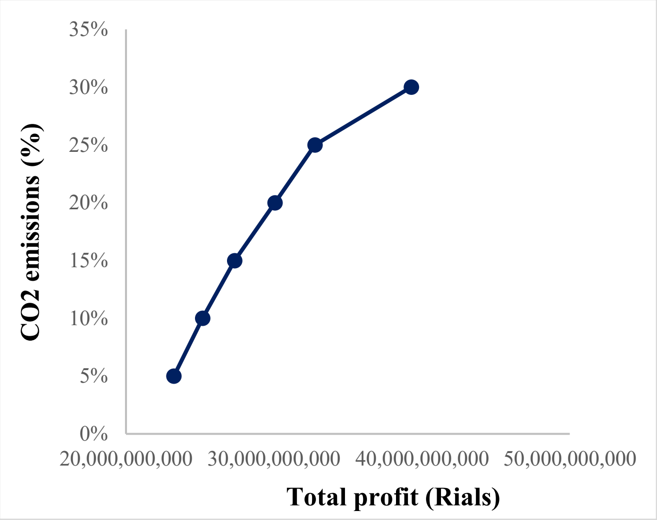

Figure 3. Profit maximization versus CO2 minimization

The analysis reveals that the acceptable emission threshold for maintaining profitability lies between 12-15% CO2 emissions. Exceeding this range significantly, while initially boosting profit, leads to an infeasibility in the problem as the CO2 minimization objective. Conversely, restricting emissions to more than 25% might appear profitable but lacks logical consistency in a real-world context. Thus, the feasible region that balances both profit and environmental responsibility lies within the 12-15% waste emissions range, where retailers can achieve reasonable profits while adhering to sustainable practices. This quantitative analysis emphasizes the intricate trade-offs involved in decision-making, highlighting the necessity of optimizing both profit margins and emission levels to create a viable and responsible distribution strategy.

5.2.2. *Examining the effects of capacity costs*

The limitations of retail capacity due to the availability of products and purchasing from the suppliers significantly impact the performance and improvement of the proposed distribution system. One of the most important analyses in the problem of distribution network design is the examination of the capacities of centers and the corresponding changes due to cost variations. As shown in the Figure 4. Network structure while using basic input parameters, when the problem is solved using basic information, three processing centers are needed at the national level to meet the network's demand, which are to be established in Tehran, Isfahan, and Mashhad. Additionally, the processing center established in Tehran is at capacity level 3, the processing center in Isfahan is at capacity level 2, and the processing center in Mashhad is at capacity level 1.

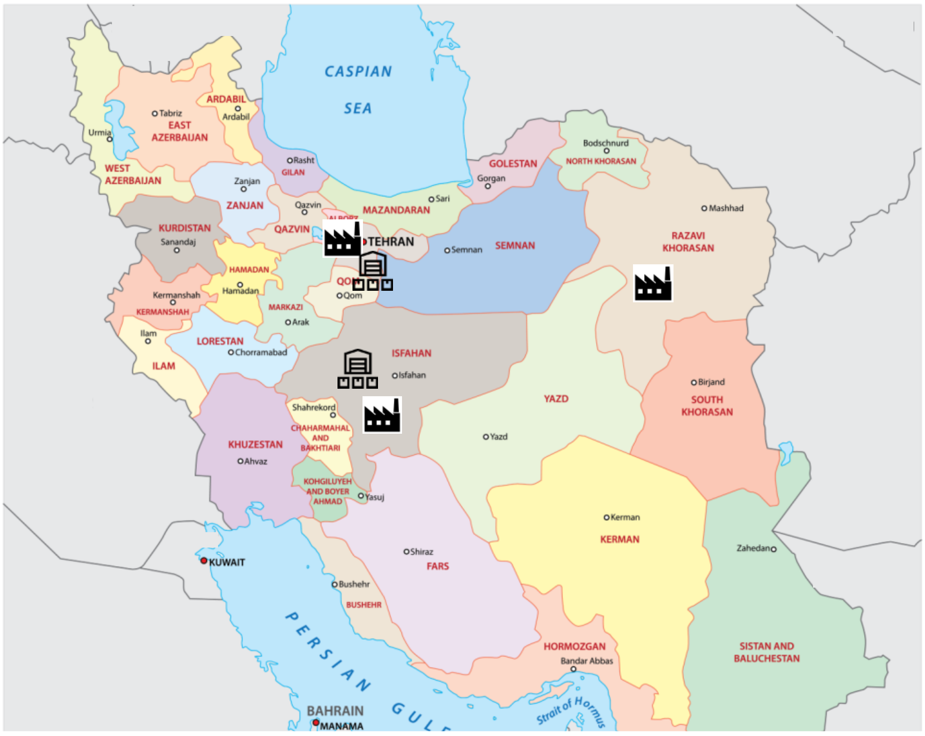

Figure 4. Network structure while using basic input parameters

Figure 5. Network structure increasing set-up and facilitation costs

As can be seen in the figure, two sorting centers have also been established at capacity levels 2 and 1 in Tehran and Isfahan, respectively, which sufficiently cover demand in the planning horizon. One of the most important analyses is the examination of the costs of constructing and equipping these centers.

Therefore, the problem was solved again with a 25% increase in the related costs. As shown in the Fig. 5, this time the number of established centers decreased, and the capacity level of the centers increased. This means that no processing center was built in Mashhad, and instead, the capacity levels of the processing centers in Tehran and Isfahan were each increased by one level. It should be noted that the locations of the sorting centers did not change, but the capacity of both established centers has been increased.

5.2.3. *Examining the service-level effects*

One of the important analyses that can have a significant and noticeable role in logistical processes, as well as in distribution networks and the manner of service delivery, is the sensitivity analysis on the service level parameter. This analysis helps in understanding how variations in service levels can impact overall operational efficiency and associated costs. In this context, one of the critical constraints outlined in the problem is the necessity for the system to meet a specific level of demand consistently. This implies that any changes or fluctuations in service levels can have profound implications on fulfilling customer expectations and maintaining inventory balance. Thus, exploring the sensitivity of this parameter becomes essential for making informed managerial decisions that align with organizational goals.

To provide a comprehensive view, the analysis on the service level parameter has been conducted at three distinct levels: 80%, 85%, and 90%. Each of these levels represents different strategies that the organization might adopt to balance customer satisfaction with operational costs. The outcomes of these analyses not only reflect potential performance under varying conditions but also help in identifying the trade-offs between service levels and resource allocation. The results of this comprehensive analysis are summarized in the Table 3, Table 4 and Table 5, which outlines how each service level affects key performance indicators such as transfer time, establishment, processing and transportation costs. This analysis not only aids in strategic planning but also equips decision-makers with valuable insights to enhance service delivery while optimizing logistical and distribution networks.

Table 3. Analysis on the service-level parameter

SL=80%		Time period			
		1	2	3	4
Mini-sized products	Establishment costs	205.9	386.2	688.3	1055.7
	Facilitation costs	3.4	4.7	6.8	9.1
Mega-sized products	Establishment costs	270.2	391.9	701.9	1339.2
	Facilitation costs	3.1	4.2	6.3	8.8
Middle-mile costs		387.3	576.6	864.2	1204.3
Last-mile costs		485.9	743.1	1206.4	1941.6
Covered demand	Fixed distribution centers	65%	65%	66%	66%
	Mobile distribution centers	30%	30%	30%	30%
	Postal offices	5%	5%	4%	4%

Table 4. Analysis on the service-level parameter

SL=85%		Time period			
		1	2	3	4
Mini-sized products	Establishment costs	205.9	386.2	688.3	1055.7
	Facilitation costs	3.4	4.7	6.8	9.1
Mega-sized products	Establishment costs	270.2	391.9	701.9	1339.2
	Facilitation costs	3.1	4.2	6.3	8.8
Middle-mile costs		396.4	604.1	921.3	1314.6
Last-mile costs		516.2	791.6	1282.2	2311.5
Covered demand	Fixed distribution centers	67%	67%	67%	66%
	Mobile distribution centers	30%	30%	30%	30%
	Postal offices	3%	3%	3%	4%

Table 5. Analysis on the service-level parameter

SL=90%		Time period			
		1	2	3	4
Mini-sized products	Establishment costs	205.9	386.2	688.3	1055.7
	Facilitation costs	3.4	4.7	6.8	9.1
Mega-sized products	Establishment costs	270.2	391.9	701.9	1339.2
	Facilitation costs	3.1	4.2	6.3	8.8
Middle-mile costs		396.4	604.1	921.3	1314.6
Last-mile costs		516.2	791.6	1282.2	2311.5
Covered demand	Fixed distribution centers	67%	67%	67%	66%
	Mobile distribution centers	30%	30%	30%	30%
	Postal offices	3%	3%	3%	4%

5.3. Managerial insights

The analysis of distribution network design highlights crucial strategic considerations for decision-makers in logistics and supply chain management. Understanding how capacity and costs interact allows managers to optimize resource allocation and enhance operational efficiency. Here are some key insights:

1- Capacity Planning: The need to reduce the number of processing centers while increasing their capacity suggests that organizations should prioritize scalability in their distribution strategy. By focusing investment on fewer, higher-capacity centers, companies can streamline their operations, reduce overhead, and enhance service levels to meet demand more effectively.

2- Cost Management: The response to a 25% increase in construction and equipping costs underscores the importance of cost control in facility management. Managers should conduct sensitivity analyses regularly to understand how fluctuations in costs impact overall network design and capacity decisions. Leveraging existing centers more efficiently can mitigate the risk of rising costs.

3- Location Optimization: While the initial plan included a processing center in Mashhad, the decision to forgo its construction in favor of expanding capacity in Tehran and Isfahan illustrates the need for ongoing evaluation of site selection. Managers should continuously assess market demands, transportation costs, and operational capabilities before committing to new locations.

4- Future Scalability: By developing higher-capacity centers, businesses position themselves for future growth. This proactive approach ensures that the organization is prepared to scale operations according to market demands without incurring the costs associated with building new facilities.

5- Flexibility and Responsiveness: The ability to adapt to cost changes reflects the importance of flexibility in decision-making. Managers should develop responsive strategies that allow for quick adjustments in distribution plans, fostering resilience in the face of cost fluctuations or unexpected market shifts.

6- Balancing Profit and Sustainability: Managers should recognize the critical trade-off between profitability and CO2 emissions. By strategically aiming for an emission target of 10-15%, organizations can maintain a balance that maximizes profit while adhering to social and environmental responsibilities. This optimal range allows for a sustainable business model that not only meets profit goals but also enhances the company's reputation and compliance with regulatory standards.

7- Segmenting Customers: By analyzing customer segments, businesses can adjust service levels. High-priority clients may require enhanced service, while others may be satisfied with more basic offerings.

8- Balanced Strategy: Striking a balance between service levels and operational constraints is crucial. This strategy optimizes delivery effectiveness while managing costs, supporting sustainable growth.

These managerial insights can guide decision-makers in creating a robust and efficient distribution network that aligns with organizational goals and market dynamics. By focusing on capacity optimization, cost management, and strategic location planning, companies can enhance their competitive edge in the marketplace.

6. Conclusions

The design of distribution networks is a crucial concern across the globe, particularly for perishable and chemically sensitive products like edible oils, where maintaining quality and minimizing waste are paramount. Our study addresses these challenges by proposing a data-driven network design model developed to the unique requirements of Bonakchi's edible oil supply chain. By integrating chemical-specific constraints—such as storage conditions, transportation time limits, and environmental controls—into the optimization framework, our model ensures efficient and sustainable distribution. The results demonstrate the model's effectiveness in balancing cost, service level, and product integrity, offering actionable insights for similar industries handling sensitive goods. To achieve this, integrating decisions such as fleet type optimization, capacity utilization strategies, route planning, transport automation, and machine-learning techniques can significantly enhance fleet efficiency and distribution system performance. Most traditional network design models primarily focus on strategically locating facilities to minimize supply chain costs, often overlooking essential elements such as demand trends, customer density, and shifts in consumer behavior. To develop a more adaptable distribution system capable of better serving customers, it is vital to incorporate these overlooked factors into the design process. By simultaneously evaluating demand patterns and optimizing fleet composition, we can address existing gaps in the literature and create a system that is highly responsive and able to reach more customers effectively. This study aims to enhance performance through the introduction of a novel design model, utilizing location strategies, flow optimization, and transport fleet management to inform the final model.

Ultimately, this research aims to maximize profits, minimize carbon emissions, and facilitate optimal decision-making in configuring distribution networks. A multi-objective design model is proposed, specifically tailored to enhance last-mile delivery operations. This model is formulated as a Mixed-Integer Linear Programming (MILP) problem. To address this model, the K-means clustering technique is utilized to effectively group customers based on varying characteristics. The resulting clusters serve as inputs for the distribution network design, while the e-constraint approach is employed to transform the bi-objective model into a single-objective equivalent. Furthermore, a fuzzy possibilistic method is applied to manage the uncertainties associated with transfer times in the model. This framework is tested on a real-world case involving Bonakchi Company, enabling the extraction of valuable managerial insights.

The case of Bonakchi's edible oils underscores the importance of sector-specific adaptations in distribution network design. For instance, optimizing fleet composition to include temperature-controlled vehicles or prioritizing shorter routes for highly perishable oils can significantly reduce spoilage and enhance customer satisfaction. Our work highlights how data-driven approaches can bridge the gap between generic logistics models and the nuanced demands of specialized products, paving the way for more resilient and responsive supply chains in the food and chemical sectors.

For future research, competition can be added to the existing assumptions when different networks are responsible for fulfilling the customers demand. Including inventory decisions is also an interesting addition to the model. The use of robust and stochastic optimization programming is also recommended for enthusiastic researchers to analyzed the other type of uncertainties in input parameters. Finally, the development of heuristic or meta-heuristic methods, especially during greater disruptions, can be beneficial for solving the model in large-scales.

References

[1] Janjevic M, Merchán D, Winkenbach M. Designing multi-tier, multi-service-level, and multi-modal last-mile distribution networks for omni-channel operations. *European Journal of Operational Research.* 2021;294(3):1059-77. https://doi.org/10.1016/j.ejor.2020.08.043

[2] Li J, Liu Y, Yang G. Two-stage distributionally robust optimization model for a pharmaceutical cold supply chain network design problem. *International Transactions in Operational Research.* 2024;31(5):3459-93. https://doi.org/10.1111/itor.13267

[3] Esmaili H, Bornak R, Shahrokhi A, Jadidi H. Improving production efficiency by implementing DMAIC phases on the Six Sigma method: a case study on the oil industry. *International Journal of Industrial Chemistry*. 2024;15(4):1-26. https://doi.org/10.57647/j.ijic.2024.1504.23

[4] Nugroho G, Santoso SA. Dynamical modeling of substrate and biomass effluents in up-flow anaerobic sludge blanket (UASB) biogas reactor. *International Journal of Industrial Chemistry*. 2019;10(4):311-9. https://doi.org/10.1007/s40090-019-00194-w

[5] Esmaeili A, Sepahvand A, Rostamzadeh R, Joksiene I, Antucheviciene J. Effect of integration of green constructs and traditional constructs of brand on green purchase intention of customers. *E+ M Ekonomie a Management*. 2017;20(3):219-37. http://dx.doi.org/10.15240/tul/001/2017-3-015

[6] Li Y, Han M, Shahidehpour M, Li J, Long C. Data-driven distributionally robust scheduling of community integrated energy systems with uncertain renewable generations considering integrated demand response. *Applied Energy*.2023;335:120749. https://doi.org/10.1016/j.apenergy.2023.120749

[7] Gharavipour M, Sheikh Sajadieh M, Ziari M. Optimizing a hub arc location problem with set-up cost and isolated hub nodes. *Advances in Industrial Engineering*. 2024;58(1):219-36. https://doi.org/10.22059/aie.2024.372107.1888

[8] Ivanov D, Dolgui A. A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Production Planning & Control*. 2021;32(9):775-88. https://doi.org/10.1080/09537287.2020.1768450

[9] Juris JS. Networking futures: The movements against corporate globalization: *Duke University* Press; 2020. https://doi.org/10.2307/j.ctv11sn2g6

[10] Truden C, Hewitt M. The service network design problem with fleet and emissions management. Transportation Research Part C: *Emerging Technologies*.2024;166:104769. https://doi.org/10.1016/j.trc.2024.104769

[11] Wang G, Gunasekaran A, Ngai EW. Distribution network design with big data: Model and analysis. *Annals of Operations Research.* 2018;270: 539-51. https://doi.org/10.1007/s10479-016-2263-8

[12] Perl J, Sirisoponsilp S. Distribution networks: facility location, transportation and inventory. *International Journal of Physical*

Distribution & Materials Management. 1988;18(6):18-26. https://doi.org/10.1108/eb014705

[13] Barahona F, Jensen D. Plant location with minimum inventory. *Mathematical Programming*.1998;83:101-11. http://dx.doi.org/10.1007/BF02680552

[14] Gabor AF, van Ommeren J-KC. Approximation algorithms for facility location problems with a special class of subadditive cost functions. *Theoretical computer science*. 2006;363(3):289-300. https://doi.org/10.1016/j.tcs.2006.04.013

[15] Melo MT, Nickel S, Da Gama FS. Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research*. 2006;33(1):181-208. https://doi.org/10.1016/j.cor.2004.07.005

[16] Esmaeili A, Gholami M. Optimization and preparation of nanocapsules for food applications using two methodologies. *Food Chemistry*. 2015;179: 26-34. https://doi.org/10.1016/j.foodchem.2015.01.115

[17] Snyder LV, Daskin MS, Teo C-P. The stochastic location model with risk pooling. European Journal of Operational Research. 2007;179(3):1221-38. https://doi.org/10.1016/j.ejor.2005.03.076

[18] Tanonkou GA, Benyoucef L, Xie X, editors. Design of multi-commodity distribution network with random demands and supply lead-times. *2007 IEEE International Conference on Automation Science and Engineering*; 2007: IEEE. https://doi.org/10.1109/COASE.2007.4341800

[19] Tang K, Yang C, Yang J, editors. A supply chain network design model for deteriorating items. *2007 International conference on computational intelligence and security* (CIS 2007); 2007: IEEE. https://doi.org/10.1109/CIS.2007.140

[20] Xin-hua L, Jin Q, editors. Joint Decision Problem for Inventory Control and Facility Location. *2007 IEEE International Conference on Automation and Logistics*; 2007: IEEE. https://doi.org/10.1109/ICAL.2007.4339071

[21] Bogataj M, Grubbström RW, Bogataj L. Efficient location of industrial activity cells in a global supply chain. *International journal of production Economics*. 2011;133(1):243-50. https://doi.org/10.1016/j.ijpe.2010.09.035

[22] Chen Q, Li X, Ouyang Y. Joint inventory-location problem under the risk of probabilistic facility disruptions. T*ransportation Research Part B: Methodological*. 2011;45(7):991-1003. https://doi.org/10.1016/j.trb.2011.04.004

[23] Jha A, Somani K, Tiwari MK, Chan FT, Fernandes KJ. Minimizing transportation cost of a joint inventory location model using modified adaptive differential evolution algorithm. *The International Journal of Advanced Manufacturing Technology*. 2012;60: 329-41. DOI10.1007/s00170-011-3595-9

[24] Berman O, Krass D, Tajbakhsh MM. A coordinated location-inventory model. *European Journal of Operational Research*. 2012;217(3):500-8. https://doi.org/10.1016/j.ejor.2011.09.039

[25] Diabat A, Battaïa O, Nazzal D. An improved Lagrangian relaxation-based heuristic for a joint location-inventory problem. *Computers & Operations Research*. 2015;61: 170-8. https://doi.org/10.1016/j.cor.2014.03.006

[26] Ahmadi-Javid A, Hoseinpour P. Incorporating location, inventory and price decisions into a supply chain distribution network design problem. *Computers & Operations Research*. 2015;56: 110-9. https://doi.org/10.1016/j.cor.2014.07.014

[27] Gzara F, Nematollahi E, Dasci A. Linear location-inventory models for service parts logistics network design. *Computers & Industrial Engineering*. 2014;69: 53-63.https://doi.org/10.1016/j.cie.2013.12.014

[28] Shahabi M, Unnikrishnan A, Jafari-Shirazi E, Boyles SD. A three level location-inventory problem with correlated demand. *Transportation Research Part B: Methodological*. 2014;69: 1-18. https://doi.org/10.1016/j.trb.2014.07.005

[29] Wu Y, Qureshi AG, Yamada T. Adaptive large neighborhood decomposition search algorithm for multi-allocation hub location routing problem. *European Journal of Operational Research.* 2022;302(3):1113-27. https://doi.org/10.1016/j.ejor.2022.02.002

[30] Song L, Wu Z. An integrated approach for optimizing location-inventory and location-inventory-routing problem for perishable products. *International Journal of Transportation Science and Technology*. 2023;12(1):148-72. https://doi.org/10.1016/j.ijtst.2022.02.002

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

[32] Fahmy SA, Zaki AM, Gaber YH. Optimal locations and flow allocations for aggregation hubs in supply chain networks of perishable products. *Socio-Economic Planning Sciences*. 2023;86: 101500. https://doi.org/10.1016/j.seps.2022.101500

[33] Hasani Goodarzi A, Zegordi SH, Alpan G, Nakhai Kamalabadi I, Husseinzadeh Kashan A. Reliable cross-docking location problem under the risk of disruptions. *Operational Research*. 2021;21:1569-612. https://doi.org/10.1007/s12351-020-00583-5

[34] Behnamian J, Fatemi Ghomi SMT, Jolai F, Heidary P. Location-allocation and scheduling of inbound and outbound trucks in multiple cross-dockings considering breakdown trucks. *Journal of Optimization in Industrial Engineering*. 2018;11(1):51-65. DOI:10.22094/JOIE.2017.594.1382

[35] Theophilus O, Dulebenets MA, Pasha J, Lau Y-y, Fathollahi-Fard AM, Mazaheri A. Truck scheduling optimization at a cold-chain cross-docking terminal with product perishability considerations. *Computers & Industrial Engineering*. 2021;156: 107240. https://doi.org/10.1016/j.cie.2021.107240

[36] Vincent FY, Aloina G, Jodiawan P, Gunawan A, Huang T-C. The vehicle routing problem with simultaneous pickup and delivery and occasional drivers. *Expert Systems with Applications*. 2023;214: 119118. https://doi.org/10.1016/j.eswa.2022.119118

[37] Kartal Z, Hasgul S, Ernst AT. Single allocation p-hub median location and routing problem with simultaneous pick-up and delivery. Transportation Research Part E: *Logistics and Transportation Review*. 2017;108: 141-59. https://doi.org/10.1016/j.tre.2017.10.004

[38] Nasiri MM, Mousavi H, Nosrati-Abarghooee S. A green location-inventory-routing optimization model with simultaneous pickup and delivery under disruption risks. *Decision Analytics Journal*. 2023;6: 100161. https://doi.org/10.1016/j.dajour.2023.100161

[39] Shi Y, Lin Y, Li B, Li RYM. A bi-objective optimization model for the medical supplies' simultaneous pickup and delivery with drones. *Computers & Industrial Engineering*. 2022;171: 108389.

https://doi.org/10.1016/j.cie.2022.108389

[40] Hosseini-Motlagh S-M, Farahmand M, Nouri-Harzvili M. A multi-path traffic-covering pollution routing model with simultaneous pickup and delivery. *Computers & Industrial Engineering*. 2022;173: 108644. https://doi.org/10.1016/j.cie.2022.108644

[41] Che Y, Zhang Z. An Integer L-shaped algorithm for vehicle routing problem with simultaneous delivery and stochastic pickup. An Integer L-shaped algorithm for vehicle routing problem with simultaneous delivery and stochastic pickup. *Computers & Operations Research*. 2023;154: 106201. https://doi.org/10.1016/j.cor.2023.106201

[42] Parilina E, Yao F, Zaccour G. Dynamic pricing and advertising in a supply chain with environment-and fairness-concerned consumers. *Central European Journal of Operations Research*. 2025:1-26. https://doi.org/10.1007/s10100-025-00972-1

[43] Jayaraman V. Transportation, facility location and inventory issues in distribution network design: An investigation. *International Journal of Operations & Production Management.* 1998;18(5):471-94. https://doi.org/10.1108/01443579810206299

[44] Yang L, Ng CT, Cheng TE. Evaluating the effects of distribution centres on the performance of vendor-managed inventory systems. *European Journal of Operational Research*. 2010;201(1):112-22. https://doi.org/10.1016/j.ejor.2009.02.006

[45] Ağralı S, Geunes J, Taşkın ZC. A facility location model with safety stock costs: analysis of the cost of single-sourcing requirements. *Journal of Global Optimization*. 2012;54: 551-81. http://dx.doi.org/10.1007/s10898-011-9777-z

[46] Selim H, Ozkarahan I. A supply chain distribution network design model: an interactive fuzzy goal programming-based solution approach. *The international journal of advanced manufacturing technology*. 2008;36:401-18. https://doi.org/10.1007/s00170-006-0842-6

[47] Maihami R, Kannan D, Fattahi M, Bai C. Supply disruptions and the problem of pricing, advertising, and sourcing strategies in a retail duopoly market. *Annals of Operations Research*. 2025;344(1):217-66. https://doi.org/10.1007/s10479-024-06314-9

[48] Gao Y, Lu S, Cheng H, Liu X. Data-driven robust optimization of dual-channel closed-loop supply chain network design considering uncertain demand and carbon cap-and-trade policy. *Computers & Industrial Engineering*. 2024;187: 109811. https://doi.org/10.1016/j.cie.2023.109811

[49] Ziari M. A Data-Driven pricing Model for Distribution Systems Considering Competition. *Advances in Industrial Engineering.* 2024. https://doi.org/10.22059/aie.2024.370998.1885

[50] Liu N, Li C, Chen L, Wang J. Hybrid data-driven and model-based distribution network reconfiguration with lossless model reduction. *IEEE Transactions on Industrial Informatics*. 2021;18(5):2943-54. https://doi.org/10.1109/TII.2021.3103934

[51] Chen X, Wu W, Zhang B, Lin C. Data-driven DG capacity assessment method for active distribution networks. *IEEE Transactions on Power Systems*. 2016;32(5):3946-57. https://doi.org/10.1109/TPWRS.2016.2633299

[52] Shi X, Qiu R, He X, Ling Z, Yang H, Chu L. Early anomaly detection and localisation in distribution network: A data-driven approach. *IET Generation, Transmission & Distribution*. 2020;14(18):3814-25. https://doi.org/10.48550/arXiv.1801.01669

[53] Younus ZS, Mohamad D, Saba T, Alkawaz MH, Rehman A, Al-Rodhaan M, et al. Content-based image retrieval using PSO and k-means clustering algorithm. *Arabian Journal of Geosciences.* 2015;8:6211-24. https://doi.org/10.1007/s12517-014-1584-7

[54] Mavrotas G. Effective implementation of the ε-constraint method in multi-objective mathematical programming problems. *Applied mathematics and computation*. 2009;213(2):455-65. https://doi.org/10.1016/j.amc.2009.03.037

[55] Gunantara N. A review of multi-objective optimization: Methods and its applications. *Cogent Engineering*. 2018;5(1):1502242. https://doi.org/10.1080/23311916.2018.1502242

[56] Deb K, Sindhya K, Hakanen J. Multi-objective optimization. Decision sciences: *CRC Press*; 2016. p. 161-200. https://www.taylorfrancis.com/chapters/edit/10.1201/9781315183176-12/multi-objective-optimization-kalyanmoy-deb-karthik-sindhya-jussi-hakanen

[57] Shah NH, Prajapati NM. Sustainable inventory model for price and credit-dependent demand with carbon emissions cap and trade. *International Journal of Procurement Management*. 2025;22(1):64-80. https://doi.org/10.1504/IJPM.2025.143533

[58] Zhang S, Lee CKM, Wu K, Choy KL. Multi-objective optimization for sustainable supply chain network design considering multiple distribution channels. *Expert Systems with Applications.* 2016;65: 87-99. https://doi.org/10.1016/j.eswa.2016.08.037

[59] Peidro D, Mula J, Jiménez M, del Mar Botella M. A fuzzy linear programming based approach for tactical supply chain planning in an uncertainty environment. *European Journal of Operational Research*. 2010;205(1):65-80. https://doi.org/10.1016/j.ejor.2009.11.031

[60] Wang R-C, Liang T-F. Applying possibilistic linear programming to aggregate production planning. *International journal of production economics*. 2005;98(3):328-41. https://doi.org/10.1016/j.ijpe.2004.09.011

[61] Momenitabar M, Dehdari Ebrahimi Z, Arani M, Mattson J. Robust possibilistic programming to design a closed-loop blood supply chain network considering service-level maximization and lateral resupply. *Annals of Operations Research*. 2023;328(1):859-901. https://doi.org/10.1007/s10479-022-04930-x