Supply Chain and Operations Decision Making

www.scodm.reapress.com

Sup. Chain Oper. Decis. Mak. Vol. 3, No. 2 (2026) 144–155.

Paper Type: Original Article

Designing Resilient Biomass Energy Supply Networks under Environmental Disruptions and Market Uncertainty: An Integrated Data-Driven Clustering, MCDM, and Bi-Level Optimization Framework

Amirhossein Amou Jafari[1,*], Kiana Salehi[2]

[1] Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran; Jafaryamir558@gmail.com.
[2] Department of Industrial Engineering, University of Tehran, Tehran, Iran; Kiana.salehi@ut.ac.ir.

Citation:

Received: 02 January 2026 *Revised: 14 March 2026* *Accepted: 17 May 2026*	Amou Jafari, A., & Salehi, K. (2026). Designing resilient biomass energy supply networks under environmental disruptions and market uncertainty: An integrated data-driven clustering, MCDM, and Bi-level optimization framework. *Supply Chain and Operations Decision Making*, *3*(2), 144-155.

Abstract

Given the growing energy demand, limitations of fossil resources, and the urgent need to mitigate environmental impacts, developing renewable energy supply networks, especially those based on biomass, has gained significant importance. Accordingly, this study proposes a hybrid multi-phase decision-making framework for the optimal design of a resilient biomass energy supply network under environmental and market uncertainties. In the first phase, spatial data from 120 agricultural sites across eight provinces of Iran were clustered into homogeneous groups using four data-driven algorithms, namely Constrained K-Means (CKM), the Gaussian Mixture Model (GMM), Spectral Clustering, and Affinity Propagation (AP). Subsequently, Bayesian Best-Worst and Best-Worst PROMETHEE methods were employed to weight and rank decision criteria, facilitating the identification of optimal hub locations. The resulting clusters and hub scores were then incorporated as input parameters for the network design problem: in the first phase, hub location and cluster-to-hub allocation were determined; in the second phase, a bi-level, multi-period mathematical model was developed to optimize network flows and intra-cluster routing under uncertainty. Numerical results reveal that employing Spectral Clustering led to a 28% reduction in total supply chain costs compared to the uncluttered baseline, with other algorithms achieving reductions between 10% and 19%. This research, grounded in real-world data collected from eight provinces, demonstrates the effectiveness of data-driven and multi-criteria approaches in enhancing the sustainability, cost-efficiency, and resilience of biomass energy supply networks.

Keywords: Biomass supply chain, Data-driven clustering, Multi-criteria decision-making, Network optimization, Hub location, Uncertainty modeling.

1 | Introduction

In recent decades, the rapid growth of the global population, increasing energy demand, climate change, and heavy dependence on fossil fuels have posed significant challenges to energy systems and the environment. Among the various renewable energy alternatives, biomass has emerged as a promising renewable energy

Corresponding Author: Jafaryamir558@gmail.com
https://doi.org/10.48313/scodm.v3i2.58

source, as it can not only reduce dependence on fossil fuels but also contribute to improved waste management, rural employment generation, and mitigation of environmental pollution. However, the design of biomass-based energy supply chain networks has become a complex and multifaceted challenge due to the spatial dispersion of biomass resources, environmental and economic uncertainties, and infrastructural complexities, particularly under crisis conditions.

In this research, an integrated modeling framework is developed for the design of a biomass energy supply chain network, employing a novel approach aimed at mitigating risks and enhancing system resilience. In the first stage, spatial data on agricultural lands are analyzed and partitioned into distinct clusters using Constrained K-Means (CKM), Gaussian Mixture Model (GMM), Affinity Propagation (AP), and Spectral Clustering (SC).

This clustering process not only facilitates the logical aggregation of biomass resources but also provides a foundation for cost reduction, transportation efficiency, and infrastructure optimization. In the next stage, a combination of two Multi-Criteria Decision-Making (MCDM) methods is employed to determine the optimal locations for hub establishment. First, the weights of evaluation criteria are obtained using the Bayesian Best-Worst Method (B-BWM), and subsequently, candidate locations are ranked through the Best-Worst PROMETHEE approach. This methodology enables the incorporation of uncertainty in decision-makers' preferences while improving the robustness, accuracy, and defensibility of location-selection decisions. The clustering and ranking outcomes are then directly integrated into the mathematical network design model, which is formulated as a bi-level, two-phase, stochastic, and multi-period optimization model. In the first phase, strategic decisions regarding hub location and cluster-to-hub allocation are determined. In the second phase, the complete biomass energy network is designed by considering material and energy flows, storage operations, transportation activities, operational costs, uncertainty scenarios, and intra-cluster routing decisions. The proposed model seeks to provide a reliable and resilient framework for the optimal utilization of biomass resources under global crises and uncertain operating conditions. In addition to ensuring economic viability, the model explicitly incorporates environmental and social sustainability considerations into the design of the energy supply chain network.

2 | Literature Review

A significant body of literature has addressed the strategic design of biomass supply chain networks. Abandansari et al. [1] proposed a multi-objective optimization model for a global biomass-based electricity supply chain that simultaneously considers economic, environmental, and resilience criteria. Similarly, Nunes et al. [2] developed a multi-stage biomass supply chain design model covering all stages from biomass collection to distribution under uncertainty while accounting for economic and environmental requirements. Murele et al. [3] formulated a mathematical model for facility location, resource allocation, and material flow management across multiple network levels. Furthermore, Petridis et al. [4] employed mixed-integer linear programming to design a sustainable biomass supply chain considering economic, social, and environmental dimensions.

Another stream of research has focused on uncertainty, resilience, and risk management in biomass supply chains. Allegranzi et al. [5] proposed a bi-objective mathematical model to minimize both total costs and environmental emissions under uncertain conditions. Likewise, Fattahi et al. [6] adopted a two-stage stochastic programming framework to plan biomass-based electricity supply chains while accounting for seasonal biomass availability and social impacts. More recently, a resilient agricultural biomass supply chain network design framework was developed using a data-driven robust optimization approach capable of addressing uncertain disruptions and risk-averse decision-making environments [7]. In addition, a multi-objective fuzzy optimization model for biomass-to-electricity supply chain design was introduced to simultaneously minimize system costs and carbon emissions under multiple sources of uncertainty [8]. These studies demonstrate the growing importance of incorporating uncertainty and resilience considerations into biomass network design.

Several studies have also investigated operational efficiency and network configuration improvements. Grigoroudis et al. [9] evaluated biomass network nodes using the RDEA algorithm to identify inefficient components and improve overall network performance. Ghani et al. [10] examined biomass network optimization under different incentive policies and resource allocation strategies. Rasekh et al. [11] addressed sustainable network design in hybrid biomass supply chains by simultaneously considering economic, environmental, and technical objectives. Furthermore, a recent study integrated fixed and portable preprocessing depots into biomass supply chain network design and demonstrated significant reductions in logistics costs through improved preprocessing strategies [12]. Another recent contribution developed a multi-period biomass power generation supply chain model that explicitly incorporates inventory dynamics and seasonal biomass supply fluctuations [13]. In parallel, increasing attention has been given to the application of advanced analytical and intelligent techniques in biomass supply chain management. Azad et al. [14] employed a MCDM framework to evaluate renewable energy alternatives, including biomass and solar energy systems.

Despite notable advances in biomass supply chain network design, several important research gaps remain. Most existing studies focus on facility location, cost minimization, emission reduction, or resilience enhancement separately, while limited attention has been paid to the integrated consideration of spatial clustering, hub location selection, and vehicle routing decisions within a unified framework. Furthermore, although machine learning and data-driven approaches have recently gained increasing attention, their integration with MCDM techniques and mathematical optimization models remains relatively limited. In addition, the combined effects of spatial complexity, market volatility, environmental disruptions, and operational uncertainty are often overlooked in existing biomass supply chain models. These limitations highlight the need for a comprehensive decision-support framework capable of simultaneously addressing strategic, tactical, and operational decisions in biomass energy supply chain design.

3 | Problem Statement

To address the research gaps identified in the literature, this study develops an integrated decision-making framework for biomass energy supply chain network design. As illustrated in *Fig. 1*, the proposed framework combines machine learning-based clustering techniques, MCDM methods, and a bi-level optimization model. First, agricultural lands are grouped into spatial clusters according to their geographical characteristics. Next, candidate hub locations are evaluated and ranked using MCDM approaches. Finally, the clustering and ranking results are incorporated into the optimization model, which determines hub locations, biomass routing plans, and energy network configurations under uncertainty. The proposed framework aims to improve both the efficiency and resilience of biomass supply chain operations while considering economic, environmental, and operational factors.

This section consists of three main parts. Section 3.1 presents the clustering methodology used to aggregate agricultural lands. Section 3.2 describes the MCDM framework adopted for hub evaluation and ranking. Section 3.3 develops the bi-level optimization model for biomass energy supply chain network design under uncertainty.

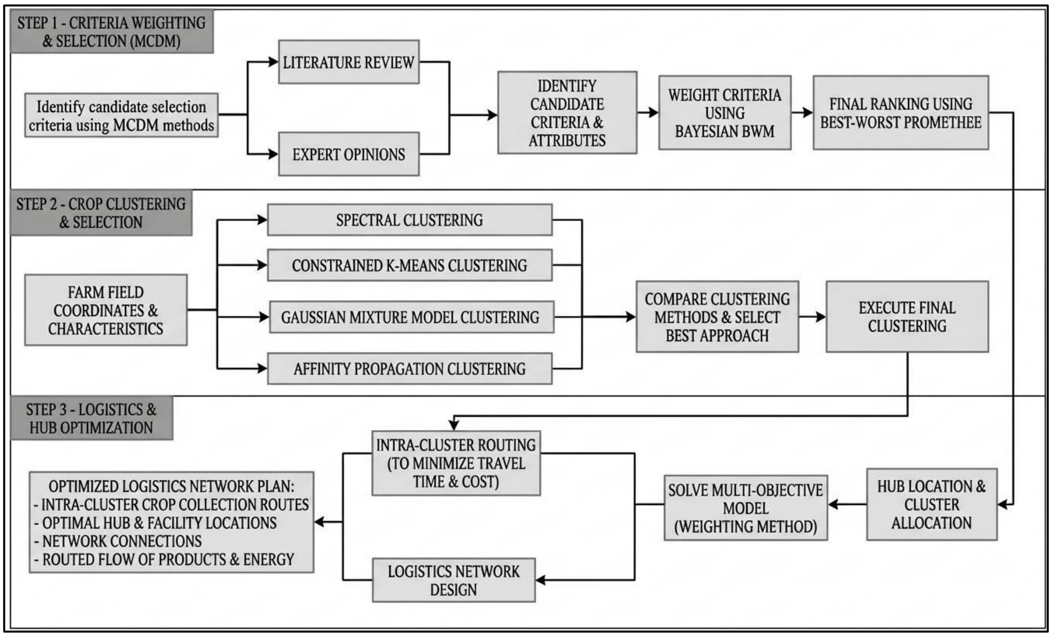

Fig. 1. Approach examined in the present article.

3.1 | Agricultural Land Clustering

In biomass energy supply chain network design, one of the fundamental steps is the effective aggregation of spatially dispersed agricultural areas to define the network's input nodes. Since farm-related spatial data are typically characterized by high dispersion, irregular geographical boundaries, and heterogeneous production capacities, clustering techniques serve as an essential preprocessing tool for reducing problem complexity and improving network coherence. Accordingly, four clustering algorithms, CKM, GMM, AP, and SC, are employed in this study to identify optimal spatial groupings of farms and to facilitate a comprehensive evaluation in terms of clustering quality, operational feasibility, and transportation efficiency.

The first method, CKM, extends the traditional K-Means algorithm by incorporating additional constraints on cluster size and composition. The algorithm seeks to allocate observations to (K) clusters in such a way that the total within-cluster sum of squared distances is minimized while satisfying predefined constraints. The optimization problem is expressed in *Eq. (1)*.

$$\min \sum_{i=1}^{K} \sum_{x_j \in C_i} \left\| x_j - \mu_i \right\|^2 \text{ subject to constraint on } |C_i|, \quad (1)$$

where x_i denotes the input data points, μ_i represents the centroid of cluster i, and C_i is the set of observations assigned to cluster i. This approach is particularly useful when collection-center capacities are limited and explicit control over cluster sizes is required.

In contrast, the GMM adopts a probabilistic perspective by representing the data as a weighted combination of multivariate Gaussian distributions. The probability density function is defined as

$$p(x) = \sum_{i=1}^{K} \pi_i N(x|\mu_i, \sum\nolimits_i), \quad (2)$$

where π_i, and $\sum_i$ denote the mixture weight, mean vector, and covariance matrix of component i, respectively. Parameter estimation is performed using the Expectation-Maximization (EM) algorithm, which iteratively maximizes the log-likelihood function:

$$\max \log \prod_{j=1}^{N} \sum_{i=1}^{K} \pi_i N(x|\mu_i, \sum_i). \quad (3)$$

Compared with K-Means, GMM provides greater flexibility in modeling overlapping clusters and datasets with ambiguous cluster boundaries.

The third technique, AP, differs from conventional clustering methods in that it does not require the number of clusters to be specified a priori. Instead, clusters are formed through an iterative message-passing process between data points. Two types of messages are exchanged: responsibility, which reflects the suitability of one point to serve as the exemplar of another, and availability, which indicates the degree to which a point is willing to act as an exemplar. The clustering objective is to maximize the total similarity between observations and their selected exemplars:

$$\max \sum_i s(i, c_i), \quad (4)$$

where $s(i, c_i)$ denotes the similarity between point i and its assigned exemplar c_i. Through this iterative process, the algorithm is also capable of automatically determining the appropriate number of clusters.

The fourth approach, SC, utilizes spectral graph theory to identify complex and nonlinear structures within the data. Initially, a similarity matrix W is constructed among all observations. Subsequently, the normalized graph Laplacian is computed as

$$L_{sym} = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}. \quad (5)$$

where D is the diagonal degree matrix. After computing the eigenvectors of L_{sym}, the data are projected into a lower-dimensional space and clustered using K-Means. SC is particularly effective for identifying complex and non-convex cluster structures. In this study, the four clustering algorithms are employed to generate alternative spatial configurations of biomass resources rather than merely compare algorithm performance. The resulting diversity in cluster structures enables a comprehensive evaluation of transportation efficiency and network performance, providing a robust basis for selecting the most suitable configuration for biomass supply chain network design.

3.2 | Hub Location Evaluation Using MCDM

To identify the most suitable locations for biomass hubs, a two-stage MCDM framework is employed. First, criterion weights are determined using the B-BWM, which accounts for uncertainty in decision-makers' judgments by modeling preferences as probability distributions rather than fixed values. This approach provides more reliable and robust criterion weights while enabling improved sensitivity analysis.

Subsequently, candidate hub locations are evaluated and ranked using the Best-Worst PROMETHEE method. By combining the robust weighting capability of B-BWM with the ranking effectiveness of PROMETHEE, this hybrid approach improves decision transparency, reduces subjective bias, and enhances the reliability of the final ranking results. The obtained hub scores are then incorporated into the optimization model as an input for hub location decisions. Due to space limitations, the detailed computational procedures are omitted, and interested readers are referred to [15], [16].

3.3 | Bi-Level Biomass Supply Chain Network Design

In this section, the mathematical model for biomass energy network design is presented within a two-phase framework: the first phase is dedicated to hub location and cluster allocation, while the second phase models the complete network structure by accounting for product and energy flows and uncertainties. The complete list of symbols and indices is provided in *Table 1*.

Table 1. Notation and Indexing.

	Description
Sets	
$p = 1,2,\dots,P$	Types of collected products
$f = 1,2,\dots,F$	Agricultural lands
$r = 1,2,\dots,R$	Reactors
$h = 1,2,\dots,H$	Hubs
$c = 1,2,\dots,C$	Condensers
$c' = 1,2,\dots,C'$	Secondary condensers
$t' = 1,2,\dots,T'$	Transformers
$d = 1,2,\dots,D$	Demand points
$t = 1,2,\dots,T$	Time periods
$s = 1,2,\dots,S$	Scenarios
$\eta = 1,2,\dots,\eta$	Available clusters
Phase 2 Parameters	
α	Natural spoilage percentage of residues
M	Big number (M)
PS_s	Probability of scenario s occurring
Cap_r	Capacity of reactor r
Cap_h	Capacity of hub h
$Z'_{\eta ht}$	1 if cluster η is connected to hub h at time t'; 0 otherwise.
$\widetilde{I_{dt}}$	Maximum demand of point d at time t
Q	Capacity of the intra-cluster collection vehicle
q^s_{pft}	Quantity of product p on land f in period t under scenario s
EP^s_{rc}	1 if pipeline between r and c is active under scenario s; 0 otherwise.
$EP^s_{cc'}$	1 if pipeline between c and c' is active under scenario s; 0 otherwise.
dis_{ij}	Distance between land i and land j
Ct_{phr}	Transport cost of product p between hub h and reactor r
$Ct_{o\in\{rc,cc',c't',t'd\}}$	Transport cost between two layers o
$Cp_{o\in\{rc,cc'\}}$	Pipeline cost between two layers o
$F_{e\in\{h,r,c,t'\}}$	Construction cost of facility e
$price_d$	Electricity selling price to the demand point d
hc^{main}_{ts}	Main warehouse holding cost in period t under scenario s
hc^{adj}_{ts}	Auxiliary warehouse holding cost in period t under scenario s
hc^{w}_{ts}	Waste holding cost in period t under scenario s
hc^{I}_{ts}	Inventory holding cost in period t under scenario s
$time^s_{ij\eta t}$	Travel time between farms i and j in cluster η in period t under scenario s
Phase 2: Decision Variables	
$y_{e\in\{h,r,c,t'\}t}$	1 if facility e is established in period t; 0 otherwise.
P^s_{phrt}	Quantity of product p transported from hub h to reactor r in period t under scenario s.
U^s_{rct}	Energy transferred from reactor r to condenser c in period t under scenario s
$U^s_{cc't}$	Energy transferred from condenser c to secondary condenser c' in period t under scenario s
$U^s_{c't't}$	Energy transferred from secondary condenser c' to transformer t' in period t under scenario s
$EL^s_{t'dt}$	Electricity transferred from transformer t' to demand point d in period t under scenario s
φ^s_{rct}	1 if pipeline between reactor r and condenser c is installed in period t under scenario s; 0 otherwise.
$\varphi^s_{cc't}$	1 if pipeline between condenser c and secondary condenser c' is installed in period t under scenario s; 0 otherwise.
ϕ^{main}_{phts}	Main warehouse inventory of product p at hub h in period t under scenario s
ϕ^{adj}_{phts}	Auxiliary warehouse inventory of product p at hub h in period t under scenario s
I^s_{pht}	Inventory of product p at hub h in period t under scenario s
W^s_{pht}	Waste of product p at hub h in period t under scenario s

Table 1. continued.

	Description
Phase 2: Decision Variables	
u_{it}^{s}	Visit the position of land i in period t under scenario s
x_{ijt}^{s}	1 if route goes from farm i to farm j in period t under scenario s; 0 otherwise.
Phase 1 Parameters	
π_h	Score of hub h
$dis_{\eta h}$	Distance between cluster η and hub h
K	Maximum number of hubs to be established.
η	Maximum number of clusters that can be connected to one hub.
Phase 1 Decision Variables	
ω_{ht}	1 if hub h is established in period t; 0 otherwise.
$Z_{\eta h}^{t}$	1 if cluster η is assigned to hub h in period t; 0 otherwise.

$$\mathrm{MaxZ1} = \sum_{H,T} \pi_h \times \omega_{ht}. \quad (6)$$

$$\mathrm{MinZ2} = \sum_{\eta,H,T} dis_{\eta h} Z_{\eta h}^{t}. \quad (7)$$

$$\sum_{H} Z_{\eta h}^{t} = 1, \qquad \forall \eta, T. \quad (8)$$

$$\sum_{\eta} Z_{\eta h}^{t} \leq \eta\,, \qquad \forall H, T. \quad (9)$$

$$\sum_{H} \omega_{ht} \leq K, \qquad \forall T. \quad (10)$$

$$Z_{\eta h}^{t} \leq \omega_{\eta t}, \qquad \forall \eta, H, T. \quad (11)$$

$$Z_{\eta h}^{t}, \omega_{\eta t} \in \{0,1\}. \quad (12)$$

Eq. (6) maximizes total hub scores across planning periods, while *Eq. (7)* minimizes total distance between clusters and their assigned hubs. *Constraints (8)-(11)* ensure each cluster is assigned to exactly one active hub per period, limit cluster-to-hub assignments, and control the number of active hubs. The bi-objective model is solved using the weighted-sum method.

$$\begin{aligned}\mathrm{MinZ'} = \sum_{S} PS_s \Bigg(& \sum_{P,H,R,T} Ct_{phr} \times P_{phrt}^{s} + \sum_{R,C,T} Ct_{rc} \times U_{rct}^{s} + \sum_{C,C',T} Ct_{cc'} \times U_{cc't}^{s} \\ & + \sum_{C',T',T} Ct_{c't'} \times U_{c't't}^{s} + \sum_{T',D,T} Ct_{t'd} \times EL_{t'dt}^{s} + \alpha \left(\sum_{T,e\in\{h,r,c,t'\}} F_e \times y_{et} \right) \\ & + \alpha \left(\sum_{R,C,T} \varphi_{rct}^{s} \times Cp_{rc} + \sum_{C,C',T} \varphi_{cc't}^{s} \times Cp_{cc'} \right) + \sum_{P,H,T} hc_{ts}^{main} \times \phi_{phts}^{main} \\ & + \sum_{P,H,T} hc_{ts}^{adj} \times \phi_{phts}^{adj} + \sum_{P,H,T} hc_{ts}^{w} \times W_{pht}^{s} + \sum_{P,H,T} hc_{ts}^{I} \times I_{pht}^{s} \\ & + \sum_{i,j,T,\eta} x_{ijt}^{s} \times dis_{ij\eta} \times time_{ij\eta t}^{s} - \sum_{T',D,T} EL_{t'dt}^{s} \times price_d \Bigg), \end{aligned} \quad (13)$$

$$\phi_{phts}^{main} + \phi_{phts}^{adj} = \sum_{\eta=1}^{\eta} Z'_{\eta ht} \times \sum_{f\in F_\eta} q_{pft}^{s}, \quad \forall H, P, T, S, \quad (14)$$

$$\sum_{P} \phi_{phts}^{main} \leq Cap_h + M \times y_{ht}, \qquad \forall H, T, S, \quad (15)$$

$$\sum_{P} \phi^{adj}_{phts} \leq Cap_h \times y_{ht}, \quad \forall H, T, S, \tag{16}$$

$$\sum_{P} \phi^{adj}_{phts} \geq \sum_{P} \phi^{main}_{phts} - Cap_h, \quad \forall H, T, S, \tag{17}$$

$$I_{ph(t-1)s} + \phi^{main}_{phts} + \phi^{adj}_{phts} = \sum_{R} P^{S}_{phrt} + I_{phts} + w_{phts}, \quad \forall P, H, T, S, \tag{18}$$

$$P^{S}_{phrt} \leq M \times \varphi^{s}_{hrt}, \quad \forall H, R, P, T, S, \tag{19}$$

$$\sum_{P,H} P^{S}_{phrt} \leq cap_r, \quad \forall R, T, S, \tag{20}$$

$$\sum_{H} P^{S}_{phrt}(1-\alpha) \leq \sum_{C} U^{s}_{rct}, \; \forall P, R, T, , S, \tag{21}$$

$$\sum_{R} U^{s}_{rct}\ (1-\alpha) \leq \sum_{C'} U^{s}_{cc't}, \forall C, T, S, \tag{22}$$

$$\sum_{C'} U^{s}_{cc't}(1-\alpha) \leq \sum_{C'} U^{s}_{c't't} \; \forall C, T', T, S, \tag{23}$$

$$\sum_{C'} U^{s}_{c't't}(1-\alpha) \leq \sum_{D} EL^{s}_{t'dt}, \; \forall T', T, S, \tag{24}$$

$$U^{s}_{rct} \leq M \times EP^{S}_{rc} \times \varphi^{s}_{rct}, \; \forall R, C, T, S, \tag{25}$$

$$U^{s}_{cc't} \leq M \times y_{ct}, \quad \forall C, T', T, S, \tag{26}$$

$$U^{s}_{cc't} \leq M \times EP^{S}_{cc'} \times \varphi^{s}_{cc't}, \quad \forall C, T', T, S, \tag{27}$$

$$U^{s}_{rct} \leq M \times y_{rt}, \quad \forall R, C, T, S, \tag{28}$$

$$EL^{s}_{t'dt} \leq M \times y_{t't}, \quad \forall T', D, T, S, \tag{29}$$

$$\sum_{T'} EL^{s}_{t'dt} \leq \tilde{I}_{dt} \times M, \quad \forall D, T, S, \tag{30}$$

$$u^{s}_{it} + u^{s}_{jt} \geq \sum_{p} q_{pjts} + Q\left(1 - x^{s}_{ijt}\right), \quad \forall (i.j) \in F_{\eta}, \eta, T, S, i \neq j, \tag{31}$$

$$\sum_{i \in F_{\eta}} x^{s}_{ijt} = \sum_{i \in F_{\eta}} x^{s}_{jit}, \quad \forall j \in F_{\eta}, \eta, T, S, \tag{32}$$

$$\sum_{i} x^{s}_{ijt} = 1, \forall j \in F_{\eta}, \eta, T, S, \tag{33}$$

$$\sum_{j \in F_{\eta}} x^{s}_{1jt} = 1, \forall \eta, T, S, \tag{34}$$

$$\sum_{P} q^{s}_{pit} \leq u^{s}_{it}, \forall \eta, i \in F_{\eta}, T, S, \tag{35}$$

$$\sum_{i \in F_{\eta}} u^{s}_{it} \leq Q, \quad \forall \eta, T, S, \tag{36}$$

$$P^{s}_{phrt}, U^{s}_{rct}, U^{s}_{cc't}, U^{s}_{c't't}, EL^{s}_{t'dt}, \phi^{main}_{phts}, \phi^{adj}_{phts}, I^{s}_{pht}, W^{s}_{pht} \geq 0, \tag{37}$$

$$y_{et}, \varphi^{s}_{rct}, \varphi^{s}_{cc't}, x^{s}_{ijt}, \omega_{ht}, \quad Z^{t}_{\eta h} \in \{0.1\}. \tag{38}$$

Eq. (13) minimizes total supply chain costs, including transportation, facility establishment, inventory holding, and intra-cluster routing, net of electricity revenue. *Constraints (14)-(18)* govern biomass flow from farms to hubs, warehouse activation priorities, and inventory balance. *Constraints (19)-(24)* regulate biomass-to-energy conversion flows across reactors, condensers, and transformers. *Constraints (25)-(30)* ensure transfers occur only through active facilities and links, with electricity supplied only to genuine demand regions. Finally, *Constraints (31)-(36)* manage intra-cluster vehicle routing by eliminating sub-tours, enforcing farm visits, ensuring vehicle return, and imposing capacity limits. *Constraints (37)-(38)* define variable domains.

4 | Numerical Results

The proposed model was solved under three uncertainty scenarios over a planning horizon of seven time periods. The dataset consisted of spatial information from 120 agricultural lands distributed across eight provinces of Iran, and it was assumed that ten hub facilities would be established for biomass collection and distribution.

In the first stage, eight key criteria were considered for hub location selection. Criterion C1 represents accessibility to the inter-regional transportation network, while C2 reflects the availability of technical infrastructure, including electricity and water supply. Criterion C3 evaluates physical expansion and storage capacity, whereas C4 considers geotechnical stability and site safety. Criterion C5 accounts for land acquisition or leasing costs, and C6 measures proximity to consumption centers and downstream industries. Criterion C7 incorporates local community acceptance, while C8 assesses industrial development status and regional synergy potential. These criteria collectively formed the basis for evaluating candidate hub locations. Subsequently, the performance of the four clustering algorithms CKM, AP, SC, and GMM, was evaluated using four widely adopted clustering quality metrics. As illustrated in *Fig. 2*, SC achieved the best or near-best performance across all evaluation criteria, obtaining a Silhouette coefficient of 0.78, a Davies–Bouldin index of 0.22, a Dunn index of 0.71, and a Calinski–Harabasz score of 310. These results indicate that SC provides superior cluster cohesion, inter-cluster separation, and overall structural quality compared with the alternative methods. CKM was identified as the second-best clustering approach and also demonstrated satisfactory performance.

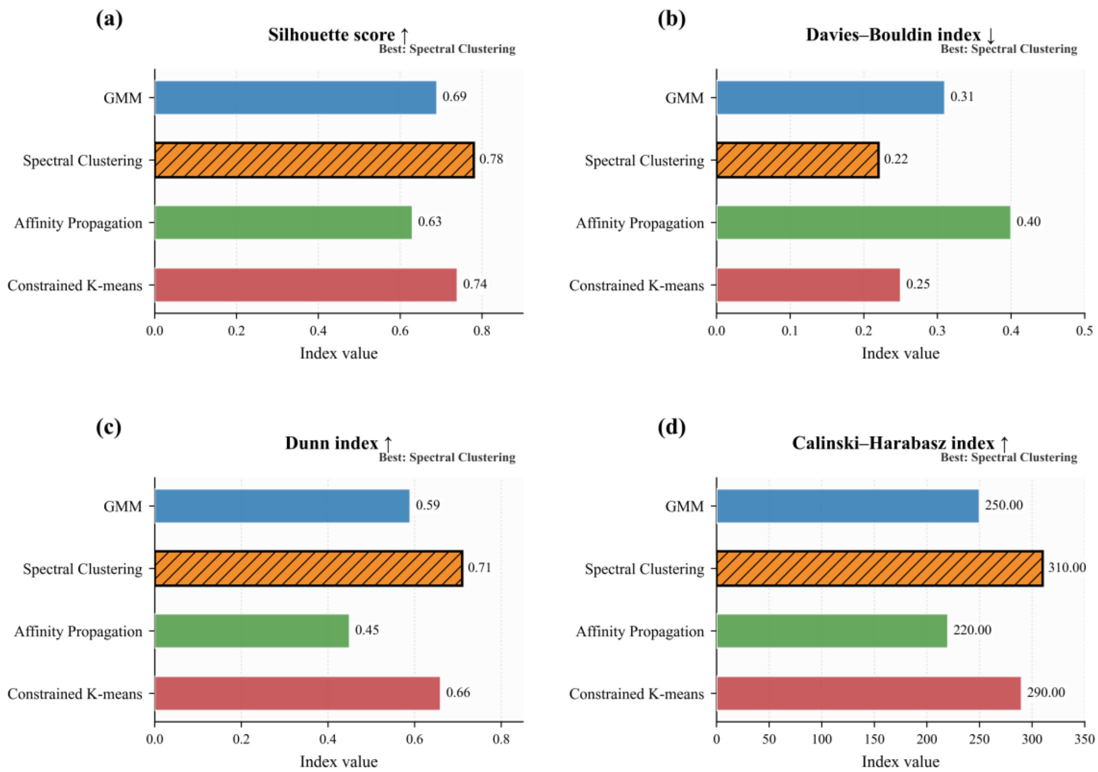

Fig. 2. Numerical comparison of clustering quality obtained by the four clustering algorithms based on standard cluster validation metrics.

The results of the mathematical optimization model are presented in *Fig. 3*. In Phase 1, optimal hub locations were determined by simultaneously minimizing total assignment distances and maximizing hub evaluation scores. The resulting cluster-to-hub allocation structure is illustrated in *Fig. 3*. To demonstrate the routing capability of the proposed framework, Gilan Province was selected as a representative case study, and the corresponding routing and biomass collection results are also reported. Although the same analysis was conducted for all provinces under consideration, only the results for Gilan Province in the third planning period under the pessimistic scenario are presented for brevity.

Fig. 4 compares the total biomass supply chain costs obtained under five different configurations. The results reveal that the highest total cost is associated with the no-clustering scenario, reaching 21,354,851 monetary units. In contrast, SC produced the lowest total cost of 15,326,458 monetary units, representing a cost reduction of approximately 28% relative to the baseline case. CKM also achieved significant cost savings, reducing total costs by approximately 18.9% compared with the no-clustering scenario. These findings highlight the critical role of data-driven clustering strategies in enhancing the operational efficiency of biomass supply chain networks and reducing overall system costs, particularly under conditions characterized by uncertainty and operational complexity.

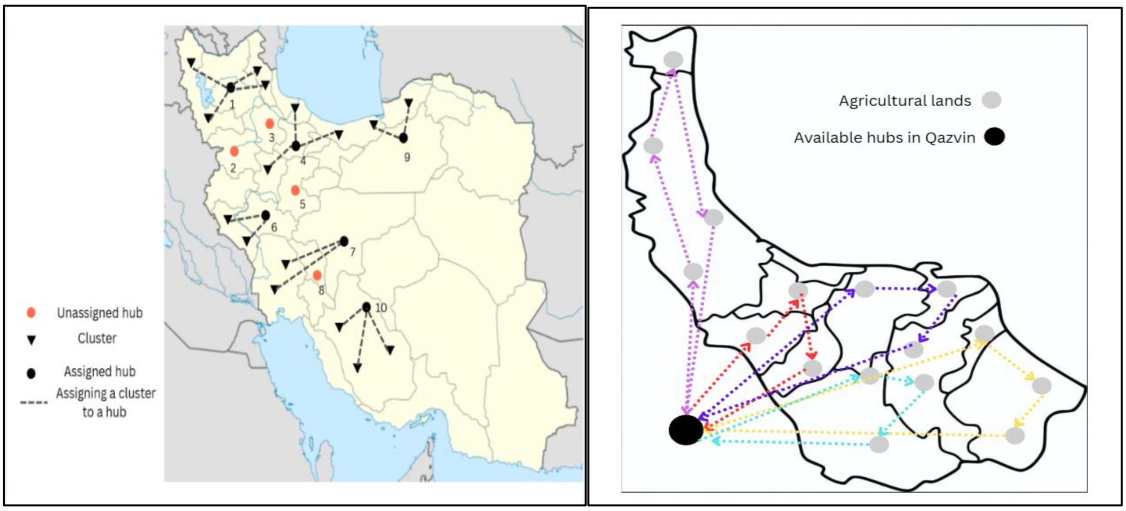

Fig. 3. Cluster-to-hub assignment and vehicle routing results for Gilan Province under the pessimistic scenario.

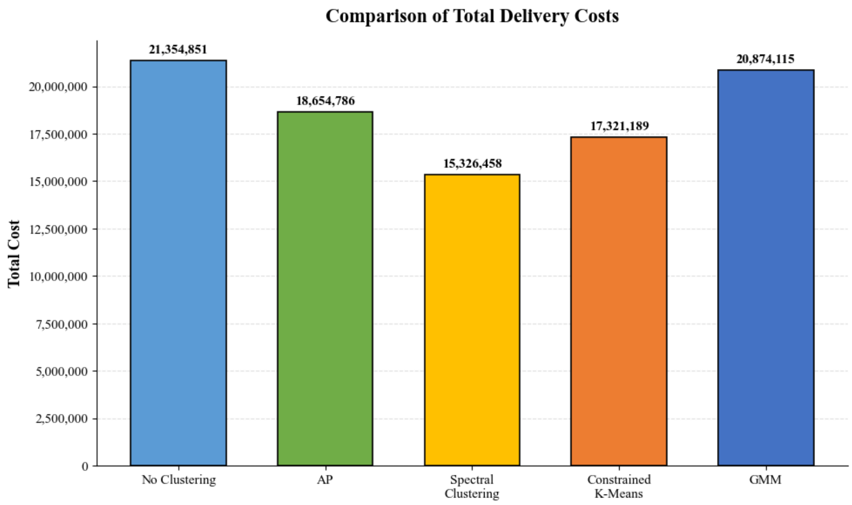

Fig. 4. Comparison of total biomass supply chain costs under the pessimistic scenario in Period 3.

5 | Conclusion

This study proposed a multi-stage decision-making framework for the design of a biomass energy supply chain network under uncertainty. The framework simultaneously addresses hub location selection, resource allocation, and product and energy flow design by integrating data-driven clustering techniques, MCDM methods, and a bi-level optimization model. Through the combination of these methodologies, a comprehensive and practical framework was developed to support the design and management of sustainable bioenergy supply chain networks. The computational results demonstrated that the integration of clustering techniques with optimization modeling can significantly enhance both the economic and operational performance of the network. The proposed approach achieved substantial reductions in total supply chain costs, improved resource allocation efficiency, reduced logistical complexity, and enhanced overall network effectiveness. Furthermore, the findings revealed that the selection of an appropriate clustering method constitutes a strategic decision that can considerably influence network design outcomes, particularly in large-scale problems characterized by high-dimensional and spatially dispersed data.

The results also highlight the importance of adopting data-driven and integrated decision-support tools for the development of sustainable and resilient bioenergy supply chains. As environmental disruptions, resource uncertainties, and market fluctuations continue to increase, the need for flexible and adaptive planning frameworks becomes increasingly critical. The proposed framework provides decision-makers with a systematic tool for improving network performance while accounting for operational and strategic uncertainties. Several directions for future research can be identified. First, the proposed model can be extended to dynamic and real-time decision-making environments by incorporating online data updates and adaptive optimization mechanisms. Second, advanced machine learning techniques may be integrated to improve clustering performance, demand forecasting accuracy, and uncertainty estimation. Third, future studies may explicitly incorporate policy-related considerations, environmental regulations, carbon management strategies, and social equity objectives into the network design process. Such extensions would contribute to the development of more efficient, adaptive, and socially responsible bioenergy supply chain systems capable of supporting long-term sustainability and resilience objectives.

Acknowledgments

The authors sincerely thank all those who contributed to this research through their guidance, expertise, and helpful recommendations.

Funding

No funding was received for this research.

Data Availability

Relevant data supporting the conclusions of this study can be obtained from the corresponding author upon reasonable request.

References

[1] Abandansari, S. A. M., Ghasemi, P., Attar, A., Goodarzian, F., & Ahmadi, S. A. (2025). A hybrid biomass supply chain optimization approach using sequential adaptive fuzzy inference learning: Insights from a cogeneration system in Europe. *Results in engineering*, *26*(2), 105261. https://doi.org/10.1016/j.rineng.2025.105261

[2] Nunes, L. J., Casau, M., Dias, M. F., Matias, J. C. O., & Teixeira, L. C. (2023). Agroforest woody residual biomass-to-energy supply chain analysis: Feasible and sustainable renewable resource exploitation for an alternative to fossil fuels. *Results in engineering*, *17*, 101010. https://doi.org/10.1016/j.rineng.2023.101010

[3] Murele, O. C., Zulkafli, N. I., Kopanos, G., Hart, P., & Hanak, D. P. (2020). Integrating biomass into energy supply chain networks. *Journal of cleaner production*, *248*, 119246. https://doi.org/10.1016/j.jclepro.2019.119246

[4] Petridis, K., Grigoroudis, E., & Arabatzis, G. (2018). A goal programming model for a sustainable biomass supply chain network. *International journal of energy sector management*, *12*(1), 79–102. https://doi.org/10.1108/IJESM-09-2017-0002

[5] Allegranzi, B., Tartari, E., & Pittet, D. (2021). "Seconds save lives—clean your hands": the 5 May 2021 World Health Organization SAVE LIVES: Clean Your Hands campaign. *Antimicrobial resistance & infection control*, *10*(1), 55. https://doi.org/10.1186/s13756-021-00926-7

[6] Fattahi, M., Govindan, K., & Farhadkhani, M. (2021). Sustainable supply chain planning for biomass-based power generation with environmental risk and supply uncertainty considerations: A real-life case study. *International journal of production research*, *59*(10), 3084–3108. https://doi.org/10.1080/00207543.2020.1746427

[7] Zhang, D., Wang, J., Li, S., & Wang, Y. (2026). Resilient agricultural biomass supply chain network design with uncertain disruptions: A data-driven globalised robust optimisation approach. *International journal of production research*, 1–27. https://doi.org/10.1080/00207543.2026.2647006

[8] Nguyen, D. D., Nananukul, N., & Luukka, P. (2026). Multi-objective fuzzy optimization model for a biomass-to-electricity supply chain network design under uncertainty. *Renewable energy focus*, *58*, 100856. https://doi.org/10.1016/j.ref.2026.100856

[9] Grigoroudis, E., Petridis, K., & Arabatzis, G. (2014). RDEA: A recursive DEA based algorithm for the optimal design of biomass supply chain networks. *Renewable energy*, *71*, 113–122. https://doi.org/10.1016/j.renene.2014.05.001

[10] Ghani, N. M. A. M. A., Vogiatzis, C., & Szmerekovsky, J. (2018). Biomass feedstock supply chain network design with biomass conversion incentives. *Energy policy*, *116*, 39-49. https://doi.org/10.1016/j.enpol.2018.01.042

[11] Rasekh, A., Hamidzadeh, F., Sahebi, H., & Pishvaee, M. S. (2023). A sustainable network design of a hybrid biomass supply chain by considering the water–energy–carbon nexus. *Energy science and engineering*, *11*(3), 1107–1132. https://doi.org/10.1002/ese3.1374

[12] Bhatt, G., Upadhyay, A., & Sahoo, K. (2025). Biomass supply chain network design: Integrating fixed and portable preprocessing depots for cost efficiency and sustainability. *Applied energy*, *389*, 125757. https://doi.org/10.1016/j.apenergy.2025.125757

[13] Wang, Z., You, Y., Wang, Z., & Wang, H. (2025). Optimization of the agricultural and forestry biomass power generation supply chain considering multi-period inventory. *Sustainable futures*, *10*, 100831. https://doi.org/10.1016/j.sftr.2025.100831

[14] Azad, A. A. S., Oishi, Z. T., Haque, M. A., Das, P., Udoy, S. A., & Bhuiya, K. M. S. (2024). An integrated framework for assessing renewable-energy supply chains using multicriteria decision-making: A study on Bangladesh. *Clean energy*, *8*(3), 1-19. https://doi.org/10.1093/ce/zkae019

[15] Mohammadi, M., & Rezaei, J. (2020). Bayesian best-worst method: A probabilistic group decision making model. *Omega (United Kingdom)*, *96*, 102075. https://doi.org/10.1016/j.omega.2019.06.001

[16] Ishizaka, A., & Resce, G. (2021). Best-Worst PROMETHEE method for evaluating school performance in the OECD's PISA project. *Socio-economic planning sciences*, *73*, 100799. https://doi.org/10.1016/j.seps.2020.100799