

Robust Multi-Product Inventory Allocation and Replenishment for Online Retailers under Demand Uncertainty

Majid Toulabi¹ , Vajhollah Ghorbanizadeh² , Mehrnaz Naderi³ , Ahmad Taherianfar⁴ , Maryam Piroujan⁵ 

1. Corresponding author, Assistant Professor, Department of Public Administration (MIS), Allameh Tabataba'i University, Tehran, Iran, Email: Majid.toulabi@gmail.com
2. Professor, Department of Industrial Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran, Email: ghorbanizadeh@atu.ac.ir
3. MsC, Department of Financial Management, North Tehran Azad University, Iran, Email: Mhrnzndri@gmail.com
4. MsC, Department of Philosophy (Ph.D.) Management Information Systems, Allameh Tabataba'i University, Tehran, Iran, Email: Ahmadtaherianfar@yahoo.com
5. MsC, Faculty of Electrical and Computer Engineering, Technical and Vocational University, Semnan, Iran, Email: maryampiroujan@gmail.com

Article Info	ABSTRACT
Article type: Research Article Article history: Received 15 May 2026 Received in revised form 16 June 2026 Accepted 17 September 2026 Published online 1 October 2026 Keywords: inventory management, inventory allocation, demand uncertainty, online retailing, metaheuristic algorithms	The rapid growth of e-commerce and increasing demand fluctuations have made inventory management and replenishment decisions among the most critical challenges faced by online retailers. Under such conditions, overlooking demand uncertainty may lead to inventory shortages, increased holding costs, reduced service levels, and decreased customer satisfaction. In this study, a mathematical programming model based on a robust optimization approach is proposed for the multi-product inventory allocation and replenishment problem under demand uncertainty. The proposed model aims to minimize the total system costs, including holding, shortage, and replenishment costs, while maintaining an acceptable service level within the online retail network. The model is implemented using GAMS software and solved through the CPLEX optimizer, and its performance is evaluated across different problem scales. The computational results indicate that although the exact solution approach encounters computational limitations for large-scale instances, the application of metaheuristic methods is capable of generating high-quality solutions within reasonable computational times. The findings demonstrate the effectiveness of the proposed model in improving inventory management decisions and enhancing the resilience of online retail systems under uncertain demand conditions.

Cite this article: Toulabi, M. & et al, (2026)., Robust Multi-Product Inventory Allocation and Replenishment for Online Retailers under Demand Uncertainty. *Journal of Engineering Management and Soft Computing*, 12 (4). 133-150.

DOI: <https://doi.org/10.22091/jemsc.2026.15997.1382>

© The Author(s) retains the copyright.

Publisher: University of Qom



Doi: <https://doi.org/10.22091/jemsc.2026.15997.1382>

1) Introduction

The rapid growth of e-commerce and the wide variety of products have made inventory management one of the major challenges in supply chain operations. Online retailers must allocate multiple products among distribution centers and periodically replenish their inventories, while demand uncertainty caused by market fluctuations and unpredictable customer behavior (Mansouri et al., 2023) complicates these decisions. Ignoring such uncertainty may lead to stock shortages, increased holding costs, delayed order fulfillment, and reduced service levels. Moreover, making simultaneous decisions regarding inventory allocation, replenishment timing, and order quantities transforms the problem into a complex multi-dimensional optimization problem, where traditional deterministic approaches often fail in the dynamic environment of online retailing.

In the field of goods distribution, the vehicle routing problem (VRP) is one of the most fundamental optimization problems. It was first introduced by Dantzig and Ramser in 1959 (Darehmiraki, 2012). VRP involves determining optimal routes for a fleet of vehicles that serve customers located at different geographical points with the objective of minimizing total operational costs (Hemmati et al., 2025; Nedjati et al., 2017). Over time, various extensions of VRP have been developed due to additional practical constraints (Dolatnezhad samarin et al., 2012; Saberi & Mokhtari, 2025), and the problem has been widely applied in areas such as goods distribution, waste collection, school transportation, postal services, and logistics systems (Bertsimas et al., 2020; Kallaj et al., 2021; Niavand et al., 2024; Rezaei Nedjati et al., 2017; Tirkolaei et al., 2022). VRP problems are classified as NP-hard problems (Laporte & Nobert, 1987; Lenstra & Kan, 1981), and they are typically solved using exact methods such as branch-and-bound as well as metaheuristic algorithms, including genetic algorithms, tabu search, and simulated annealing.

Considering these issues, there is a strong need for a robust decision-making framework that incorporates multi-product demand uncertainty while determining optimal inventory allocation and replenishment decisions. The primary objective is to minimize inventory-related costs while maintaining an acceptable level of customer service under different demand scenarios (Krishnamoorthy & Sanaei, 2026). In this study, a robust mathematical programming model is developed that simultaneously considers inventory decisions and routing costs of heterogeneous vehicles, time window constraints, and pickup and delivery operations. The proposed model is formulated in the GAMS environment and solved using the CPLEX solver. Computational results indicate that the exact mathematical model is not efficient for large-scale instances; therefore, a metaheuristic approach is employed to solve larger instances and obtain the required solutions.

The remainder of the paper is organized as follows. Section 2 presents a review of the literature. Section 3 introduces the proposed research methodology, including the mathematical modeling. Section 4 presents the research results, and finally, Section 5 provides the overall conclusion along with suggestions for future research.

2) Literature Review

The VRP is one of the important problems in supply chain management in which a number of vehicles located at one or more depots (warehouses or nodes) must visit a set of customers and provide services to them, where each customer has a specific demand (Kumari et al., 2023). This problem seeks to employ mathematical models and route optimization in such a way that the traveled distance, total

travel time, the number of transportation vehicles, delay penalties, and ultimately, the transportation cost function are minimized, thereby maximizing customer satisfaction. For example, Neira et al (2020) proposed an integer programming model for the Multi-Trip Vehicle Routing Problem with Time Windows considering service-dependent loading times and travel duration (MTVRPTW-SDLT). The first model presented in this study considers the return of vehicles to the depot. An exact solution approach was used to solve the problem. Huang et al. (2021) introduced a model for the Multi-Trip Vehicle Routing Problem with Time Windows (MTVRPTW). In the proposed model, vehicles unload goods collected from customers at a depot with limited unloading capacity. To solve the problem, a branch-price-and-cut (PBC) algorithm was proposed. Rezaei Kallaj et al. (2021) presented a multi-objective model for the Multiple Vehicle Routing Problem (MO-MVRP) under emergency conditions for blood delivery to injured individuals. In the proposed model, vehicle arrival time and the amount of collected blood were considered as the objective functions. The proposed model was solved using the exact CPLEX solver. Wang et al. (2022) introduced a Collaborative Multi-Depot Vehicle Routing Problem with Dynamic Customer Demands and Time Windows (CMVRPDCDTW) while considering Resource Sharing (RS) and dynamic customer requirements. For this purpose, a bi-objective optimization model was developed to optimize vehicle routes while minimizing total operational costs and the number of vehicles. To solve the proposed model, a hybrid algorithm consisting of an improved k-medoids clustering algorithm and multi-objective particle swarm optimization was presented to obtain near-optimal solutions. Hasanpour Jesri et al. (2022) formulated a Multi-Trip Open Vehicle Routing Problem (MTOVRP). For this purpose, an integer programming model was developed to minimize the total purchasing costs. To solve the model, a decomposition-based algorithm was proposed that divides the problem into two stages. In the first stage, tactical decisions regarding supplier selection and the type of collaboration are made. In the second stage, the visiting sequence of each vehicle is determined. Nozari et al. (2022) presented a Multi-Depot Vehicle Routing Problem (MDVRP) under uncertainty. The primary objective of the proposed model is the location of depots and production centers along with vehicle routing for distributing medical supplies to hospitals. To solve the model, a robust fuzzy approach was used with uncertain parameters such as demand, transportation, and distribution costs. The effect of uncertainty was investigated using the neutrosophic fuzzy programming method. Jiao et al. (2022) proposed a Multi-Stage Vehicle Routing algorithm based on Task Grouping (MSVR-TG) for the rescue vehicle routing problem with energy constraints during natural disasters. To solve the problem, they developed a heuristic algorithm based on k-means clustering and a genetic algorithm. Karimi and Nawrin Ferdous (2022) studied a Vehicle Routing Problem with Time Windows in which each truck has variable capacity. In this model, each vehicle can carry one or more trailers. Customers are visited within specified time windows, and vehicles can perform multiple trips. A mathematical programming model was presented and solved using a branch-and-price algorithm. Mardaneh et al. (2022) considered a periodic ship supply scheduling problem in which weekly demands at several offshore facilities are satisfied by a heterogeneous fleet of vessels. The objective of the proposed scheduling model is to determine the optimal fleet composition and weekly vessel trips in order to minimize total voyage and fixed costs. To solve the periodic ship supply scheduling problem, a branch-and-price algorithm was proposed. Simulation results demonstrated that the proposed

approach can generate high-quality solutions for large-scale industrial problem instances. He et al. (2023) introduced a new rural public transportation system capable of integrating passenger and freight services. This approach can reduce logistics costs and generate additional revenue for transportation operators. Therefore, the objective of minimizing costs for both users and transportation operators was formulated using a mixed-integer linear programming model capable of bus routing and trip scheduling for multiple rural locations. The developed model was solved for small numerical examples using the Gurobi solver, while a tabu search metaheuristic was employed for large-scale instances. The research results show how the number of rural electric buses can be optimally designed to reduce purchasing and maintenance costs.

Although the vehicle routing problem (VRP) has been extensively studied in supply chain management, most existing studies primarily focus on routing optimization, transportation scheduling, and fleet utilization. Numerous works have proposed advanced models to address different VRP variants, such as multi-trip routing with time windows, multi-depot routing, emergency logistics routing, and routing under uncertainty. For instance, Neira et al. (2020) and Huang et al. (2021) investigated multi-trip vehicle routing problems with time windows using exact optimization methods, while Rezaei Kallaj et al. (2021) developed a multi-objective model for vehicle routing in emergency blood delivery systems. Similarly, Wang et al. (2022) addressed collaborative multi-depot vehicle routing with dynamic customer demands, and Hasanpour Jesri et al. (2022) proposed a model for multi-trip open vehicle routing considering supplier selection and collaboration decisions. Furthermore, some studies have incorporated uncertainty and real-world constraints into routing models. For example, Nozari et al. (2022) examined a multi-depot vehicle routing problem under uncertainty using a robust fuzzy approach, and Jiao et al. (2022) proposed a heuristic algorithm for rescue vehicle routing with energy constraints during disasters.

Other studies have also explored vehicle capacity variability, periodic scheduling, and integrated passenger–freight transportation systems to improve logistics efficiency (He et al., 2023; Karimi & Nawrin Ferdous, 2022; Mardaneh et al., 2022). While these studies provide valuable insights into transportation planning and routing optimization, their primary focus remains on vehicle movement and logistics scheduling. In contrast, some studies have emphasized the role of knowledge management in enhancing supply chain agility under uncertainty (Aghaei et al., 2017; Sabatosheikh et al., 2021), which is closely aligned with the need for robust inventory decisions in online retailing. However, limited attention has been paid to the integration of inventory allocation and replenishment decisions with demand uncertainty in the context of online retailing, particularly when multiple products are involved. Moreover, studies in industrial contexts have demonstrated that the interplay between change management, knowledge management, and organizational agility is essential for coping with environmental uncertainties (Husseini et al., 2016; Sabatosheikh et al., 2021; Toulabi & Taherianfar, 2026; Roshdieh, 2024).

According to the above, most existing VRP-based studies do not explicitly consider how inventory should be allocated across facilities or how replenishment policies should be designed to remain robust against uncertain and fluctuating customer demand. In addition, the rapidly evolving nature of e-commerce requires inventory systems that can simultaneously manage multi-product inventory decisions, demand uncertainty, and service level requirements in a coordinated framework.

Therefore, there remains a significant research gap in developing robust optimization models that jointly address multi-product inventory allocation and replenishment decisions for online retailers under uncertain demand conditions. Addressing this gap can provide more reliable and cost-efficient decision support for online retail supply chains, where inventory management and fulfillment responsiveness are both critical for maintaining competitive advantage and high customer satisfaction. Table 1 presents a categorized summary of the literature review (Aghaei et al., 2017).

Table 1. Summary of the Literature Review

References	Objective	Research method		Computational logic			
		Mathematical	Review	Exact	analytical	Heuristic	Meta heuristic
(Darehmiraki, 2012)	Cost minimization in vehicle routing	•		•			
(Dolatnezhad samarin et al., 2012)	Development of VRP variants with practical constraints		•	•	•		
(Rezaei Kallaj et al., 2021)	Improving routing efficiency and service level	•		•			
(Nedjati et al., 2017)	Minimization of total routing cost	•		•			
(Bertsimas et al., 2020)	Distribution and logistics efficiency	•				•	•
(Sheikhzadeh et al., 2014)	Information system support for service operations	•		•			
(Babaei & Rajabi-Bahaabadi, 2019)	Efficient transportation network planning	•		•			
(Koch et al., 2018)	Distribution planning with time windows	•				•	•
(Tirkolaei et al., 2022)	Application of VRP in service and logistics systems	•				•	•
(Lenstra & Kan, 1981)	Complexity analysis of VRP	•				•	•
(Laporte & Nobert, 1987)	Complexity analysis of VRP	•				•	•
Current research		•		•		•	•

3) Methodology

In this section of the study, the details related to the mathematical modeling of the problem and its solution method are presented. For this purpose, the sets, parameters, variables, constraints, and objective functions are first introduced. In this research, a mathematical programming model is developed for a vehicle routing problem with the capability of simultaneous pickup and delivery. The model considers a scenario in which a vehicle commences its route from the hub, delivering orders that must be distributed to the corresponding customers via distribution centers. Concurrently, it

collects orders from these distribution centers for transmission back to the hub, where they undergo processing and sorting. Figure 1 illustrates the structure of simultaneous pickup and delivery of products.

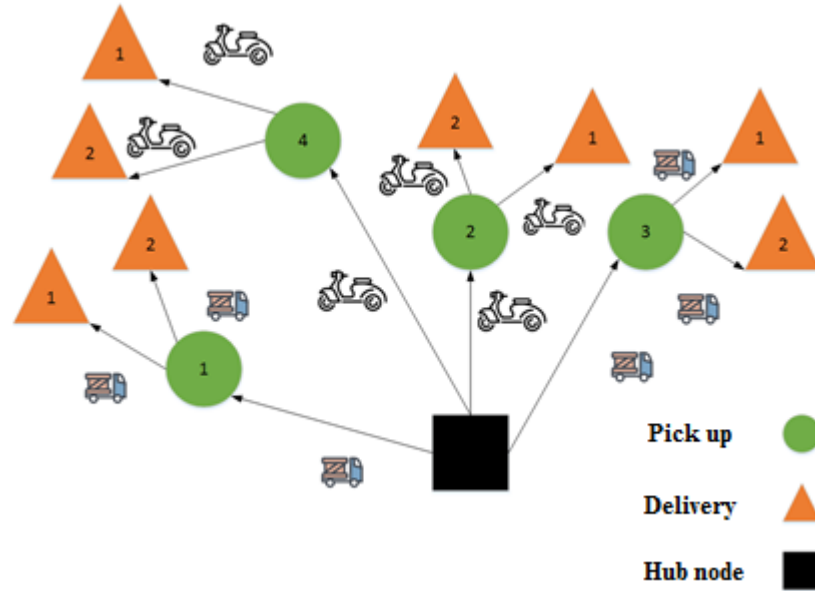


Figure 1. Illustration of Simultaneous Pickup and Delivery Using Heterogeneous Vehicles

In general, the structure considered for the model is based on the following assumptions:

- Each node representing a distribution center (DC) is visited only once. During this visit, both product pickup and product delivery are performed simultaneously by a single vehicle (Case 1).
- It is not necessary for a DC node to be visited only once for both pickup and delivery operations. In other words, the delivery operation may be performed by one vehicle and the pickup operation by another vehicle. Depending on vehicle capacities, delivery and part of the pickup operation may be performed by one vehicle, while the remaining pickup quantity is handled by another vehicle. In such cases, the situation corresponds to sending the packages of a single order to multiple destinations. In other words, in this method, the packages of an order are divided into two or more parts and sent to different addresses. This approach is used to reduce transportation costs and improve customer experience (Case 2).
- All DCs are visited by vehicles for delivery operations; then, pickup operations are carried out from them in the same order in which the DCs were visited (Case 3).
- The total quantity of orders collected from DCs is at most equal to the total quantity of orders delivered to the DCs.
- The total quantity of orders collected from DCs is greater than the total quantity of orders delivered to the DCs.

The aforementioned assumptions give rise to three modeling cases, which will be described in the following sections. However, before presenting the modeling approach, it is necessary to introduce the notation used throughout this research, including the sets, parameters, and variables.

Sets

I : The set of distribution center nodes $DC \ I = 1, \dots, i$

V : The set of all nodes (both distribution centers and the hub) $V = 1, \dots, v$

A : The set of arcs $A = 1, \dots, a$

K : The set of vehicles $K = 1, \dots, k$

Parameters

$[e_i, l_i]$: The time window for a vehicle to be present at node DC_i

Qv^k : The volumetric capacity of the vehicle k

Qw^k : The weight capacity of the vehicle k

α^k : The performance coefficient of the vehicle driver in terms of familiarity with routes

c_{ij} : The distance between nodes i and j

d_i : The number of orders that must be collected from DC_i

h_i^k : The capacity of DC_i

Decision Variables

The decision variables in this research are classified into two types: binary and continuous.

- **Binary Variable**

x_{ij}^k : it will be 1 if vehicle goes from node i to node j ; otherwise, it is 0.

- **Continues Variables**

z_{pi}^k : The quantity of products p collected at node i by vehicle k

z_{di}^k : The quantity of products that can be delivered to node i by vehicle k

s_i^k : The arrival time of vehicle k when entering node i

L_i^k : The upper bound on the load of vehicle k upon entering node i

h_i^k : The amount of loading (delivery) of the vehicle k during its visit to node i

Mathematical Modeling

$$\min_F = \sum_{(i,j) \in A} \sum_k \alpha^k c_{ij} x_{ij}^k + \sum_k \sum_j \text{Fixed cost}^k x_{1j}^k \quad (1)$$

St.

$$\sum_k \sum_j x_{ij}^k = 1 \quad \forall i \in I \quad (2)$$

$$L_i^k + h_i^k + Mx_{ij}^k \leq M + L_j^k \quad \forall k \quad (3)$$

$$h_i^k = p_i - d_i \quad \forall k \quad (4)$$

$$h_1^k = 0 \quad \forall k \quad (5)$$

$$L_i^k \leq Qv^k \quad \forall k \quad (6)$$

$$L_i^k \leq Qw^k \quad \forall k \quad (7)$$

$$\sum_j x_{1j}^k \leq 1 \quad \forall k \quad (8)$$

$$\sum_i x_{ih}^k - \sum_j x_{hj}^k = 0 \quad \forall h \in I, \forall k \quad (9)$$

$$u_i^k - u_j^k + N * x_{ij}^k \leq N - 1 \quad \forall i, j : i \neq j, i, j \geq 2 \quad (10)$$

$$e_i \sum_i x_{ih}^k \leq s_i^k \leq l_i \sum_i x_{ih}^k \quad \forall h \in I, \forall k \quad (11)$$

$$x_{ij}^k \in \{0, 1\} \quad \forall i, j, k \quad (12)$$

Objective function (1) minimizes the total traveled distance as well as the cost of using vehicles. Constraint (2) states that each DC node is visited once and only once for pickup operations. Constraint (3) adjusts the load amount of each vehicle when moving from node i to node j . Constraint (4), according to the definition of h , represents the loading amount of each vehicle upon entering node i . Constraints (3) and (4) also represent the volumetric and weight capacity restrictions of the vehicles. In fact, according to the model assumption that pickup and delivery operations for each node are performed simultaneously and by a single vehicle, the vehicle initially departs from the hub with full capacity to distribute orders. On the other hand, the quantity of collected loads must not exceed the vehicle capacity. Constraint (5) indicates that the hub node (node 1) has neither pickup nor delivery operations. Constraints (6) and (7) correspond to the volumetric and weight capacity limitations of the vehicles, respectively. Constraint (8) ensures that each vehicle, if used, can leave the hub only once (it should be noted that node 1 corresponds to the hub node in the network). Equation (9) represents the flow conservation constraint, meaning that once a vehicle enters a node, it must also leave that node. Constraint (10) prevents the formation of sub-tours. Constraint (11) guarantees that vehicle k arrives within the acceptable time window. Finally, constraint (12) specifies the sign restrictions of the decision variables.

3-1) Robust Mathematical Modeling

As mentioned, the presented model minimizes the total traveled distance as well as the cost of using vehicles. Given that the demand is fluctuating or, in other words, accompanied by uncertainty, the values of the objective functions and the problem's optimal variables can significantly differ from those obtained from the current model. Therefore, to reduce this difference, this section aims to make the model robust against uncertain parameters. In a robust optimization model, two types of variables exist: design variables and control variables. Design variables are determined before the realization of probabilistic parameters and cannot be adjusted afterward. Control variables, however, become subject to adjustment once a specific realization of the probabilistic parameters occurs. Furthermore, the constraints of the robust model comprise both structural constraints and control constraints. Consistent with the robust optimization framework of Mulvey et al. (1995), structural constraints contain no uncertain parameters or variables, whereas control constraints incorporate such elements. First, consider the linear optimization model from (13) to (16):

$$\min z = C^T X + d^T Y_s \quad (13)$$

subject to:

$$AX = b \quad (14)$$

$$B_s X + C_s Y_s = e_s \quad \forall s \quad (15)$$

$$X, Y_s \geq 0 \quad (16)$$

Where X is the vector of design variables and Y is the vector of control variables. S denotes the set of available scenarios, and the probability of occurrence for each scenario is P_s ; therefore, $\sum_{s=1}^S P_s = 1$. Constraints (14) and (16) represent the structural and control constraints of the problem, respectively.

δ_s are deviation variables and measure the amount of infeasibility of the control constraints under scenario S . The robust mathematical model, including mathematical equations (17) to (20), is expressed as follows:

$$\min \sigma(x, y_1, y_2, \dots, y_s) + \omega p(\delta_1, \delta_2, \dots, \delta_s) \quad (17)$$

subject to:

$$AX = b \quad (18)$$

$$B_s X + C_s Y_s + \delta s = e_s \quad \forall s \quad (19)$$

$$X, Y_s \geq 0 \quad \forall s \quad (20)$$

Where the expression $\sigma(x, y_1, y_2, \dots, y_s)$ measure the robustness of the solution and the expression $p(\delta_1, \delta_2, \dots, \delta_s)$ measure the robustness of the model, which can be defined by various functions. For example, $p(\delta_1, \delta_2, \dots, \delta_s)$ is usually considered equal to $\sum_{s \in S} P_s \delta_s$. Let $\sigma(x, y_1, y_2, \dots, y_s)$ be defined as the sum of the expected value and λ times the variance of the objective function, as expressed in equation (21).

$$\sigma(x, y_1, y_2, \dots, y_s) = \sum_{s \in S} P_s Z_s + \lambda \sum_{s \in S} P_s (Z_s - \sum_{s' \in S} P_{s'} Z_{s'})^2 \quad (21)$$

Mulvey et al. (1995) suggested that by substituting $\sigma(x, y_1, y_2, \dots, y_s)$ into the objective function, a second-degree nonlinear programming model is obtained. The expression $\sum_{s \in S} P_s (Z_s - \sum_{s' \in S} P_{s'} Z_{s'})^2$, which can increase the problem's solution time, may be replaced by an absolute-value expression. Therefore, the robustness of the solution is modified as shown in equation (22).

$$\sigma(x, y_1, y_2, \dots, y_s) = \sum_{s \in S} P_s Z_s + \lambda \sum_{s \in S} P_s |Z_s - \min \sum_{s' \in S} P_{s'} Z_{s'}| \quad (22)$$

The objective function is nonlinear; however, minimizing $\sigma(x, y_1, y_2, \dots, y_s)$ is equivalent to solving the linear programming model formulated in equations (23) to (25), presented below:

$$\min \sum_{s \in S} P_s Z_s + \lambda \sum_{s \in S} P_s (Z_s - \min \sum_{s' \in S} P_{s'} Z_{s'} + 2\theta_s) \quad (23)$$

subject to:

$$Z_s - \min \sum_{s' \in S} P_{s'} Z_{s'} + \theta_s \geq 0 \quad \forall s \quad (24)$$

$$\theta_s \geq 0 \quad \forall s \quad (25)$$

As a result, the optimization model will be transformed into the following robust model, presented in equation (26):

$$\min \sum_{s \in S} P_s Z_s + \lambda \sum_{s \in S} P_s (Z_s - \sum_{s' \in S} P_{s'} Z_{s'} + 2\theta_s) + \sum_{s \in S} P_s \delta_s \quad (26)$$

subject to:

Equations (2) to (12)

To better illustrate how demand uncertainty is incorporated in the proposed model, consider a simple example. Suppose an online retailer sells a product whose daily demand is not known exactly but is estimated to vary between 80 and 120 units due to fluctuations in customer behavior. Rather than assuming a fixed demand value, the robust model considers this range of possible demand values and determines inventory allocation and routing decisions that remain feasible and cost-effective under all possible realizations within this uncertainty range. In this way, the proposed robust formulation protects the system against demand variability and reduces the risk of stockouts or excessive inventory levels.

3-2) Solving Approach

In this section of the study, the solution approach for the problem is presented. Generally, metaheuristic algorithms, such as tabu search and simulated annealing, have been widely used for solving the VRP (He et al., 2023). Most metaheuristic algorithms applied to large-scale VRP problems are based on modeling the problem using network flow formulations and implementing the tabu search metaheuristic algorithm. It should also be noted that tuning the initial parameters is one of the major challenges of metaheuristic algorithms, which in daily routing operations may lead to issues such as becoming trapped in local optima and increased computational time. Most of the proposed solution approaches for very large-scale VRP problems are therefore based on heuristic methods. Two of the most efficient heuristic algorithms proposed in this field are the Hybrid Genetic Algorithm and the Iterated Local Search Algorithm, both introduced in 2013. More recently, new approaches based on pruning techniques have been developed that attempt to provide suitable solutions within an acceptable computational time by reducing the number of search operations. In general, most pruning techniques avoid storing the full $n \times n$ distance matrix—where n denotes the number of customers and each entry represents the distance between a pair of nodes—and instead retain only the distances between each node and its nearest neighbors. In this study, the performance of a pruning-based algorithm, called A-G-S, is compared with other heuristic approaches.

Due to the similarity between the routing algorithm investigated in this research and those employed by parcel delivery companies such as FedEx, determining optimal routes becomes challenging because of the high volume of orders received by FedEx. However, an important consideration is that, according to FedEx, the distances between pickup and delivery points are relatively short. This has led FedEx to divide the entire city into smaller subregions, with each subregion assigned to a specific driver. Such an assignment increases the driver's familiarity with the designated area and its customers, thereby enabling customer service to be provided more quickly and efficiently. Accordingly, inspired by the FedEx approach, Tehran was divided into 16 regions in this study, and the FedEx pickup and delivery algorithm was implemented based on this regional division. Moreover, each customer is assigned to one of these regions according to their geographical

location within Tehran. The positive impact of driver route familiarity on operational productivity has been well-documented in logistics literature (Husseini et al., 2016).

4) Findings and Discussion

In this section, the routing cost associated with delivering customer packages on a given day is first calculated for an online retail company in Tehran. Since the available data do not include the travel time between each pair of points and only their geographical coordinates are provided, the travel time between each pair of locations was estimated using the given longitude and latitude coordinates. Specifically, the travel time between two particular points at the beginning hour of the day was determined, and subsequently, the travel times between all other pairs of points were calculated proportionally.

Considering the large volume of distribution data for the company under study (2,260 records for a single day), optimization software such as CPLEX is unable to find solutions for data sets of this scale. Therefore, a clustering-based heuristic algorithm was developed to generate near-optimal solutions within a reasonable computational time. It is expected that the total cost obtained from this algorithm will be lower than the cost resulting from the company's current routing operations in the distribution sector.

Given the computational complexity of solving VRP problems, instances with approximately 200 to 400 nodes are classified as large-scale problems, and considerable efforts have been made to solve them. However, in the distribution data of the company examined in this study, the number of customer nodes ranges from approximately 2,000 to 4,000. For this reason, the problem is categorized as a very large-scale vehicle routing problem.

It is evident that existing optimization solvers, such as CPLEX, BARON, and similar tools can find optimal solutions for problems with such a large number of nodes. Consequently, various solution approaches have been proposed for these large-scale routing problems. In the following, the results of the most important algorithms used for solving such problems are discussed.

Figure 2 comparatively illustrates the solution times required by three well-known VRP algorithms for solving very large-scale VRP problems. As can be observed, for problem sizes of around 1,000 nodes, both the Iterated Local Search (ILS) algorithm and the Hybrid Genetic algorithm require more than 500 minutes (approximately 8.5 hours) to generate a solution. Furthermore, when the problem size reaches approximately 10,000 nodes, the required computation time for generating a solution extends to several months.

In contrast, the AGS algorithm demonstrates superior performance to the other two methods and is capable of producing solutions for the problem within a significantly shorter computational time. It is evident that existing optimization solvers, such as CPLEX, BARON, and similar tools can find optimal solutions for problems with such a large number of nodes. Consequently, various solution approaches have been proposed for these large-scale routing problems. In the following, the results of the most important algorithms used for solving such problems are discussed. Given the computational complexity of solving VRP problems, instances with approximately 200 to 400 nodes are classified as large-scale problems, and considerable efforts have been made to solve them. However, in the distribution data of the company examined in this study, the number of customer

nodes ranges from approximately 2,000 to 4,000. For this reason, the problem is categorized as a very large-scale vehicle routing problem.

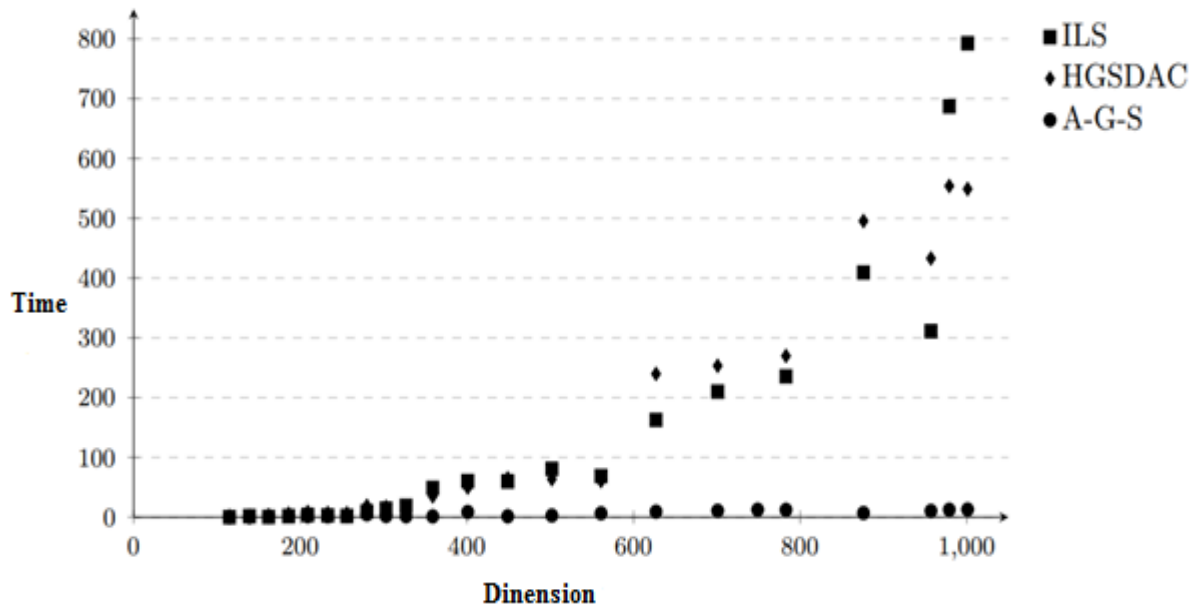


Figure 2. Computational Time

To implement the heuristic algorithm based on the FedEx approach, the city of Tehran is divided into 16 regions, and each customer is assigned to one of these regions according to their geographical location within Tehran. For this purpose, the longitude and latitude coordinates of five primary areas, i.e., Center, North, South, East, and West, are first identified using the map. Then, the coordinates of the intermediate Northwest, Northeast, Southwest, and Southeast are determined by averaging the coordinates of the four main directions (North, South, East, and West) and the center. Each customer node is then assigned to one of these 16 regions based on its longitude and latitude coordinates. Since the number of customers at the beginning of the process was 52,260, the number of customers assigned to each region is presented in Table 2. The sixteen-region division is illustrated in Figure 3.

Table 2. Number of Customers in Each Region at the Beginning of the Process Day

Region number	Number of customers	Region number	Number of customers	Region number	Number of customers	Region number	Number of customers
1	1	5	1	9	1	13	39
2	0	6	49	10	0	14	909
3	2	7	269	11	222	15	2
4	2	8	742	12	19	16	1

Given the computational complexity of solving VRP problems, instances with approximately 200 to 400 nodes are classified as large-scale problems, and considerable efforts have been made to solve them. However, in the distribution data of the company examined in this study, the number of customer nodes ranges from approximately 2,000 to 4,000. For this reason, the problem is categorized as a very large-scale vehicle routing problem.

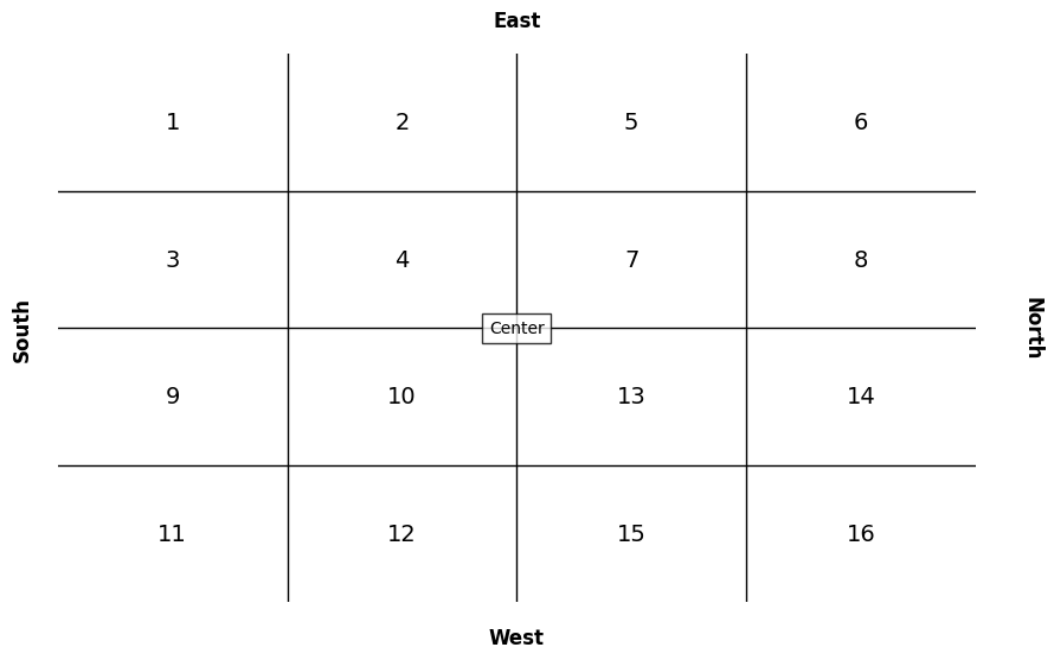


Figure 3. Regional division of Tehran

As can be observed, the highest volume of order distribution for the company under study is concentrated in the northeastern and northwestern areas of Tehran. Due to the NP-hard nature of the problem, the CPLEX solver is capable of generating an optimal solution for approximately 37 nodes, with the solution time for a 37-node problem being around 3 seconds. Therefore, another heuristic algorithm was designed to divide the customers in each region into groups containing a maximum of 37 nodes, such that customers within the same group are geographically close to one another and have relatively short mutual distances. The first approach that naturally comes to mind for this partitioning process is the use of clustering algorithms, such as k-means, which group nearby points into the same cluster. However, the issue with these clustering algorithms is that they do not impose any restriction on the number of points assigned to a cluster. In other words, a cluster may contain more than 37 customer nodes, which cannot be solved by CPLEX due to its computational limitations. Therefore, to address the problem of grouping customers into sets of 37, a fixed-cost formulation based on Constraints (37)–(40) is employed. By considering the distance matrix between every pair of points, this formulation groups nearby customers together while ensuring that the number of customers in each group does not exceed 37. To implement this model in CPLEX, the 2,260 distribution nodes of the company under study were divided into the following time intervals: 9–12, 12–15, 15–18, and 18–21, and the data corresponding to each interval were executed separately in CPLEX. The primary issue encountered in this process was the memory limitation of the CPLEX solver. Therefore, another heuristic approach was developed in this research. First, within each region, the Euclidean distance between every customer node and a selected reference node was calculated, where the hub node was considered as the reference node. The resulting distances were then sorted in ascending order, and the nodes were partitioned sequentially into successive groups of 37 nodes until all nodes had been assigned. It should be noted that 37 nodes represent the maximum problem size that CPLEX can solve optimally within a short computational time. Each of these 37-node groups was solved separately in CPLEX, and the resulting solutions were stored. For regions with fewer than 37 customer nodes, the

nodes were, where possible, merged with nodes from neighboring regions, and the corresponding routing problem was solved using CPLEX. Table 3 presents the results obtained from solving the algorithm. According to this table, the travel cost associated with each of the sixteen considered regions is reported.

Table 3. Travel Cost per Region

Region number	Travel cost	Region number	Travel cost	Region number	Travel cost	Region number	Travel cost
1	0.995	5	0.507	9	0.000	13	0.919
2	0.000	6	0.000	10	0.000	14	7.433
3	0.743	7	1.925	11	2.371	15	0.000
4	1.257	8	7.512	12	0.000	16	0.000
Total travel cost	2.995	Total travel cost	9.944	Total travel cost	2.371	Total travel cost	8.263

The total travel cost, considering the travel cost within each cluster, is equal to 23.573. By taking into account the total number of vehicles, the cost coefficient associated with the traveled distance, and the fixed cost of the vehicles, the total cost is calculated as 302,730 monetary units. After solving the problem, the number of remaining customers who had not yet been visited in each region is illustrated in Figure 4.

East			
Zone 1:1	Zone 2:0	Zone 5:1	Zone 6:14
Zone 3:2	Zone 4:2	Zone 7:10	Zone 8:2
Zone 9:1	Zone 10:0	Zone 13:2	Zone 14:0
Zone 11:12	Zone 12:19	Zone 15:2	Zone 16:1
West			

Center

South **North**

Figure 4. Customers Remained in Each Region

4-2) Robustness Results of the Proposed Model

To evaluate the performance of the robust model, a computational problem of varying dimensions was investigated, and the model was tested. The relevant results are presented in this section. The mathematical model was solved using the CPLEX solver within GAMS 24.9 software on a personal computer. The model was solved using the specified data after parameter descaling, and a sensitivity analysis was conducted on the demand parameter to confirm its performance, the report of which is provided below. To examine the solvability of the model across different dimensions, 6 samples of various sizes were generated, and their solution results are presented in Table 4. The second column, in Table 4, indicates the number of components for each sample. These components are represented

by i , v , k and a , where i is the number of DC, v refers to the nodes, k represents the vehicles, and A is the arcs.

Table 4. Numerical Experiments

No.	Sample size	Optimal value of objective function 2	Solution time (hours)
1	(2, 4, 2, 2)	659.086	02:16:10
2	(2, 8, 2, 2)	687.765	02:24:16
3	(2, 11, 3, 2)	512.624	02:38:27
4	(2, 12, 2, 4)	534.915	01:39:45
5	(2, 15, 2, 3)	712.873	02:40:45
6	(2, 20, 2, 5)	683.901	03:50:34

For validation and performance accuracy, the model was solved in GAMS/CPLEX software using sample data with dimensions of 12, 2, 2, and 6, as the problem was resolved within a short computational time. To test the accuracy of the model, a sensitivity analysis was performed on the cost of vehicle travel distance.

Table 6. Changes in the Model Parameter

Test	v	i	k	a	Objective function value	Total objective function
1	147	138	7	186	534.915	6.492E+9
	147	138	8	186		
2	147	21	9	186	683.901	6.492E+9
	147	186	10	186		
3	21	186	11	186	712.873	6.492E+9
	147	186	11	186		
4	56	167	12	149	812.624	6.492E+9
	39	138	15	186		
5	86	223	16	186	912.023	6.492E+9
	21	138	16	149		

As observed in Table 6, the cost increases with an increase in sample size. However, the total objective function value remained robust. Similarly, as the sample size increased, the total travel distance cost remained constant.

5) Conclusion

This paper proposes a method for the optimal routing of heterogeneous vehicles with time windows, capable of the simultaneous pickup and delivery of products to customers. In general, the proposed model differs significantly from previous studies in several aspects, including the consideration of the driver performance coefficient based on familiarity with routes, the presentation of three types of routing models (simultaneous pickup and delivery routing, open vehicle routing model, and integrated open-closed routing model), and the implementation in a logistics company with a large number of

customers. These features are considered the major innovations of the proposed model compared with prior research. To solve the model, an algorithm was developed using optimization methods based on travel distance and vehicle capacities. The results obtained from implementing the models indicate that the proposed approach performs better in comparison with existing methods. Furthermore, by optimizing vehicle routing, transportation time and costs are reduced, and the probability of delays in product delivery to customers decreases. According to the results obtained from the optimization model, by dispatching 67 different vehicles for multiple pickup and delivery operations and dividing the city of Tehran into 16 regions, it is possible to minimize the costs of the logistics system. Under this approach, total costs are reduced by 38.58% compared with the current situation. Therefore, it can be concluded that the proposed algorithm for heterogeneous vehicle routing with time windows and simultaneous pickup and delivery capability can provide an effective solution for minimizing transportation costs and increasing customer satisfaction. For future research, it is suggested that a multi-depot vehicle routing model be considered for vehicle allocation and dispatching. In addition, stochastic customer demand may be incorporated to formulate a probabilistic vehicle routing problem. The proposed model helps managers make more robust inventory allocation and replenishment decisions under demand uncertainty. It highlights the importance of simultaneously considering inventory decisions, heterogeneous fleet routing, and time window constraints to reduce costs while maintaining service levels. From a practical perspective, the adoption of metaheuristic solution methods provides an effective decision-support tool for large-scale operations in online retail environments. The proposed exact mathematical model is not computationally efficient for large-scale instances, even when solved using standard solvers such as CPLEX, indicating the need for a metaheuristic approach for large problems. In addition, since the model is validated through a specific case study, the generalizability of the results to all practical contexts may be limited. Furthermore, emerging paradigms, such as Industry 5.0, emphasize human-centric business models where value creation extends beyond cost minimization to include social and human factors; incorporating such perspectives could enhance the practical relevance of inventory and routing decisions.

Statements and Declarations

Ethics Statement: The author affirms that this research was carried out in compliance with recognized ethical standards and principles. All information sources have been appropriately cited and credited. Where required, ethical clearance was granted, and informed consent was obtained from all participants. No instances of plagiarism, data fabrication, or falsification were involved in this study.

Use of Artificial Intelligence: The author confirms that AI-based tools were employed exclusively for language polishing and grammatical correction. All AI-generated assistance was thoroughly examined, validated, and revised by the author, who bears full responsibility for the precision, authenticity, and integrity of the final manuscript. AI tools were not utilized to produce or alter research data or to substitute essential scholarly reasoning.

Conflicts of Interest: Any possible conflicts of interest, including financial or personal relationships that might affect the research, are disclosed. The author states that there are no competing interests.

Compliance with Regulations: The author verifies that this study adhered to all relevant institutional, national, and international rules, and that all required permissions for data usage and reproduction were secured.

References

- Aghaei, M., Dehghani, M. H., Naderi Ziarati, R., & Toulabi, M. (2017). Impact of implementation of knowledge management on supply chain agility. *Iranian Journal of Management Studies*, 10(1), 25-50. <https://doi.org/10.22059/ijms.2017.208152.672111>
- Babaei, M., & Rajabi-Bahaabadi, M. (2019). School bus routing and scheduling with stochastic time-dependent travel times considering on-time arrival reliability. *Computers & Industrial Engineering*, 138, 106125. <https://doi.org/10.1016/j.cie.2019.106125>
- Bertsimas, D., Delarue, A., Eger, W., Hanlon, J., & Martin, S. (2020). Bus routing optimization helps Boston public schools design better policies. *INFORMS Journal on Applied Analytics*, 50(1), 37-49. <https://doi.org/10.1287/inte.2019.1014>
- Darehmiraki, M. (2012). A new innovative algorithm to solve the vehicle routing problem. *Journal of Operational Research in its Applications*, 9(4), 1-7. (In Persian). Magiran/p1125203
- Dolatznezhad samarin, A., Mir Hoseini, S. A., Yousefi khoshbakht, M., (2012). A mixed integer programming model and an efficient method for the fuel supply problem of oil product supply points with a time-dependent time window. *Journal of Operational Research in its Applications*, 14(1), 15-37. (In Persian). Magiran/p1718017
- Hasanpour Jesri, Z. S., Eshghi, K., Rafiee, M., & Van Woensel, T. (2022). The multi-depot traveling purchaser problem with shared resources. *Sustainability*, 14(16), 10190. <https://doi.org/10.3390/su141610190>
- He, D., Ceder, A. A., Zhang, W., Guan, W., & Qi, G. (2023). Optimization of a rural bus service integrated with e-commerce deliveries guided by a new sustainable policy in China. *Transportation Research Part E: Logistics and Transportation Review*, 172, 103069. <https://doi.org/10.1016/j.tre.2023.103069>
- Hemmati, A., Motevalli, S. H., Pourghader Chobar, A., Akhlaghpour, A., & Nazari, L. (2025). Analyzing customer sentiment with AI to improve the smart supply chain. *Journal of Engineering Management and Soft Computing*, 11(1), 306-286. doi: <https://doi.org/10.22091/jemsc.2025.3654.1260>
- Huang, N., Li, J., Zhu, W., & Qin, H. (2021). The multi-trip vehicle routing problem with time windows and unloading queue at depot. *Transportation Research Part E: Logistics and Transportation Review*, 152, 102370. <https://doi.org/10.1016/j.tre.2021.102370>
- Husseini, A. S., Kadkhodayi, S., & Toulabi, M. (2016). Identification and ranking of factors affecting manpower productivity promotion using the ANP technique (Case study: Cement company managers and supervisors). *Quarterly Journal of Productivity Management*, 10(37), 123-145. SID.ir/paper/182046
- Jiao, L., Peng, Z., Xi, L., Guo, M., Ding, S., & Wei, Y. (2022). A multi-stage heuristic algorithm based on task grouping for vehicle routing problem with energy constraint in disasters. *Expert Systems with Applications*, 203, 118740. <https://doi.org/10.1016/j.eswa.2022.118740>
- Karimi, L., & Nawrin Ferdous, C. (2022). Branch and price algorithm for multi-trip vehicle routing with a variable number of wagons and time windows. *Algorithms*, 15(11), 412. <https://doi.org/10.3390/a15110412>
- Koch, H., Bortfeldt, A., & Wäscher, G., (2018). A hybrid algorithm for the vehicle routing problem with backhauls, time windows and three-dimensional loading constraints. *OR Spectrum*, 40(1), 3-49. <https://doi.org/10.1007/s00291-017-0487-1>
- Krishnamoorthy, A., & Sanaei, F. (2026). The retail long game: expanded service or extended return?. *Marketing Intelligence & Planning*, 1-22. <https://doi.org/10.1108/MIP-08-2025-0671>
- Kumari, M., De, P. K., Narang, P., & Shah, N. H. (2023). Integrated optimization of inventory, replenishment, and vehicle routing for a sustainable supply chain utilizing a novel hybrid algorithm with carbon emission regulation. *Expert Systems with Applications*, 220, 119667. <https://doi.org/10.1016/j.eswa.2023.119667>
- Laporte, G., Nobert, Y. (1987). Exact algorithms for the vehicle routing problem. In *North-Holland mathematics studies* (Vol. 132, pp. 147-184). North-Holland. [https://doi.org/10.1016/S0304-0208\(08\)73235-2](https://doi.org/10.1016/S0304-0208(08)73235-2)
- Lenstra, J. K., & Kan, A. R. (1981). Complexity of vehicle routing and scheduling problems. *Networks*, 11(2), 221-227. <https://doi.org/10.1002/net.3230110211>
- Mansouri, S., Samatova, V., Korchiev, N., & Anyanwu, K. (2023, October). DeMaTO: an ontology for modeling transactional behavior in decentralized marketplaces. In *2023 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (pp. 171-180). IEEE. <https://doi.org/10.1109/WI-IAT59888.2023.00029>

- Mardaneh, E., Heydar, M., & Loxton, R. (2022). Branch-and-price for clash-free periodic supply vessel planning problem with split delivery and variable service time. *Available at SSRN 4213089*. <https://doi.org/10.2139/ssrn.4213089>
- Nedjati, A., Izbirak, G., & Arkat, J. (2017). Bi-objective covering tour location routing problem with replenishment at intermediate depots: Formulation and meta-heuristics. *Computers & Industrial Engineering*, 110, 191–206. <https://doi.org/10.1016/j.cie.2017.06.007>
- Neira, D. A., Aguayo, M. M., De la Fuente, R., & Klapp, M. A. (2020). New compact integer programming formulations for the multi-trip vehicle routing problem with time windows. *Computers & Industrial Engineering*, 144, Article 106399. <https://doi.org/10.1016/j.cie.2020.106399>
- Niavand, M., Adibi, M. A., & Pourghader Chobar, A. (2024). Selection of green supplier by multi-moora combination method and two-stage clustering. *Journal of Engineering Management and Soft Computing*, 10(1), 14–49. <https://doi.org/10.22091/jemsc.2024.10977.1181>
- Nozari, H., Tavakkoli-Moghaddam, R., & Gharemani-Nahr, J. (2022). A neutrosophic fuzzy programming method to solve a multi-depot vehicle routing model under uncertainty during the Covid-19 pandemic. *International Journal of Engineering*, 35(2), 360–371. <https://doi.org/10.5829/ije.2022.35.02B.11>
- Rezaei Kallaj, M., Abolghasemian, M., Moradi Pirbalouti, S., Sabk Ara, M., & Pourghader Chobar, A. (2021). Vehicle routing problem in relief supply under a crisis condition considering blood types. *Mathematical Problems in Engineering*, 2021, Article 5590510. <https://doi.org/10.1155/2021/5590510>
- Roshdih, N. (2024). The Effect of Monetary Policy Uncertainty on Stock Market Uncertainty with NARDL Approach. *Research Journal of Finance and Accounting*, 15(10), 1-9. <https://dx.doi.org/10.2139/ssrn.5110149>
- Sabatosheikh, S. M., Hooshmandi, S., & Toulabi, M. (2021). Investigating the effect of export and change management on agility in Khuzestan Steel Company with regard to the mediating role of knowledge management. *Journal of Industrial Management Studies*, 19(62), 87–112. <https://doi.org/10.22054/jims.2020.47053.2393>
- Saberi, M., & Mokhtari, H. (2025). An analysis of multiple shipment models in inventory control systems: A review. *Journal of Engineering Management and Soft Computing*, 10(2), 1-26. <https://doi.org/10.22091/jemsc.2025.11078.1185>
- Sheikhzadeh, M. A., Moradzahipour, M., & Toulabi, M. (2014). Investigation of the role of information systems in tourism marketing as one of the strategies of tourism development. *European Online Journal of Natural and Social Sciences*, 3(3), 708–716. <https://core.ac.uk/pdf/268502396>
- Tirkolaee, E. B., Goli, A., Bakhshi, M., & Mahdavi, I. (2022). A robust multi-trip vehicle routing problem of perishable products with intermediate depots and time windows. *Numerical Algebra, Control and Optimization*, 12(4), 417–433. <https://doi.org/10.3934/naco.2021008>
- Toulabi, M., & Taherianfar, A. (2026). Human-centric business models in Industry 5.0: Redefining economic value creation. In *Industry 5.0's impact on economic innovation*. IGI Global. <https://doi.org/10.4018/979-8-3693-0123-4.ch001>
- Wang, Y., Sun, Y., Guan, X., Fan, J., Xu, M., & Wang, H. (2022). Two-echelon multi-period location routing problem with shared transportation resource. *Knowledge-Based Systems*, 226, Article 107168. <https://doi.org/10.1016/j.knosys.2021.107168>