Advances in Industrial Engineering, June 2024, 58(1): 85-102
DOI: 10.22059/aie.2024.363992.1879
RESEARCH PAPER

Solving MDVRP Using Two-Step Clustering: A Case Study of Pharmaceutical Distribution in Tehran

Abolfazl Bagheri Tofighi[1], Faezeh Gorji[1], Alireza Deravi[1], Erfan HassanNayebi[2*]

[1] *M.Sc. Student, Department of Industrial Engineering & Management Systems, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.*
[2] *Assistant Professor, Industrial Engineering Department, Sharif University of Technology, Tehran, Iran.*

Received: 18 Auguste 2023, Revised: 08 March 2024, Accepted: 31 March 2024

Abstract

The healthcare sector has recently encountered significant challenges, including limited funding and intense competition. These issues have adversely impacted hospital supply chains, resulting in budget cuts, staffing shortages, and logistical difficulties. This study introduces a novel two-step clustering approach to address the multi-depot vehicle routing problem (MDVRP) in healthcare logistics, specifically focusing on optimizing the delivery of pharmaceutical supplies to hospital pharmacies in Tehran. The method begins with the K-means algorithm to identify optimal distribution centers in the first step. In the second step, K-means clustering, incorporating vehicle capacity and demand values, is applied to each distribution center to allocate demand points for each vehicle. The vehicle routes are then determined by solving the traveling salesman problem. By optimizing the number of distribution centers using the silhouette score, which resulted in a score of 0.3567 for four centers, the study shows that deploying five vehicles from four strategically located centers can meet the needs of Tehran hospitals with a total travel distance of 119.68 km. A comparative analysis with two alternative methods reveals that the proposed approach offers a 14% improvement in minimizing the total travel distance. This method not only helps identify optimal locations for new distribution centers but also develops efficient routing plans for pharmaceutical distribution, ultimately reducing costs and improving service quality within healthcare logistics.

Keywords:
Drugs Delivery, Two-Step Clustering, Hospital Logistics, MDVRP, K-Means

Introduction

Supply chain management is a crucial component for the success of businesses, especially within the healthcare sector in recent years [1]. The healthcare industry encounters diverse challenges, including limited human resources, budget constraints, and the lack of essential resources like equipment and medications. Efficient procurement processes play a pivotal role in meeting the demands of hospital pharmacies while effectively managing budget allocations. However, the healthcare supply chain faces key issues such as quality control, insufficient transport conditions, and delays in drug delivery [2]. These challenges significantly impact access to essential drugs and services in low- and middle-income countries [3].

The healthcare supply chain contains diverse types, including pharmacies, blood banks, and

[*] Corresponding author: (Erfan HassanNayebi)
Email: hassannayebi@sharif.edu

the patient safety supply chain [4]. Among these, hospital pharmacies play a crucial role in offering medications at optimal costs while guaranteeing a continuous supply of drugs and equipment throughout the healthcare facility - from operating rooms to patient beds [2].

Considering the critical role of supply chain management within the healthcare sector, particularly the pivotal function of hospital pharmacies in providing high-quality care services, this article focuses on improving drug distribution in Tehran by employing efficient routing solutions. The main objective of this article is to address the enhancement of vehicle routing for drug distribution, with the aim of increasing the efficiency of the healthcare supply chain through the implementation of practical strategies.

The primary focus of this article revolves around a specific type of Vehicle Routing Problem (VRP) that holds significant implications for logistics and supply chain management [5]. More specifically, when it comes to distributing pharmaceuticals among hospital pharmacies in Tehran, the key challenge lies in determining the most efficient routes for the drug-carrying vehicles to ensure prompt delivery to the pharmacies. Essentially, the objective is to minimize the distance traveled by the vehicles. Given the presence of various drug distribution centers, the scenario presents itself as a Multi-Depot Vehicle Routing Problem (MDVRP) [6], where a fleet of vehicles caters to customers from different depots before returning to the same depot.

One of the significant challenges posed by MDVRPs revolves around effectively determining the optimal locations for the distribution centers. This paper introduces an innovative two-step clustering method that considers vehicle capacity, geographical location, and demand to strategically pinpoint the distribution center locations and streamline drug distribution routes. In the proposed approach, distribution centers are initially identified using the K-means algorithm. Subsequently, the K-means clustering algorithm (incorporating vehicle capacity constraints) is deployed for each center to allocate demand points that each vehicle must serve. Ultimately, the routing for each vehicle is fine-tuned utilizing the Traveling Salesman Problem (TSP) model.

The proposed methodology offers a comprehensive solution for determining optimal distribution center locations and optimizing drug distribution routes. This innovative approach not only enhances the efficiency of the distribution network but also maximizes cost-effectiveness and minimizes delivery time or distance. Furthermore, by incorporating distinct approaches for evaluating performance, the methodology provides flexibility and adaptability to account for a range of scenarios, thereby ensuring its effectiveness across different operational environments.

The structure of this article is as follows. Section 2, delves into a comprehensive review of previous researches in the domain of pharmaceutical logistics. Section 3, elaborates on the research methodology employed for this study. Section 4 investigates the numerical results and facilitates a discussion surrounding them. Finally, the conclusions of this work and recommendations for future studies are presented in Section 5.

Literature review

In recent years, the optimization of drug supply chains has emerged as a critical imperative. Ensuring the efficient distribution of pharmaceutical products while maintaining high quality and safety standards is crucial [7]. Also, on the one hand, pharmaceutical demand in the healthcare field has increased [8]. Despite the surge in demand, healthcare institutions suffer from delays in drug delivery and distribution that plague hospitals due to inadequate infrastructure and fragmented central/regional pharmacies [2]. Several studies have proposed diverse solutions, such as mathematical models and clustering methods, to address drug logistics challenges.

Midaoui et al. [2] introduced a smart logistics approach to optimize drug delivery and distribution among hospitals. They identified optimal locations for new pharmacies using the weighted K-means clustering method and implemented a Genetic Algorithm for efficient drug distribution within each cluster. Oliveira et al. [9] successfully tackled the optimization challenge of the Pfizer vaccine distribution fleet in Portugal by employing a VRP with Time Windows (VRPTW) model that considers vehicle capacity constraints. By dividing regions in Portugal into clusters based on municipal areas, multiple vaccination centers were strategically positioned within each cluster. Additionally, Magalh et al. [10] introduced a dynamic vehicle routing algorithm designed to enhance drug distribution logistics. This algorithm adjusts routes daily based on incoming orders, involving four key phases: clustering, route construction, choice of the route to be performed, and route improvement. Also, Redi et al. [11] used simulated annealing and the nearest neighbor algorithm to address capacitated VRP (CVRP) in pharmaceutical distribution. The focus of this study was on devising vehicle routes with minimal travel time for efficient drug delivery. KOSE et al. [12] introduced an innovative two-step optimization strategy to discover the most efficient distribution routes that reduce transportation time and costs for delivering drugs to 186 pharmacies across 9 regions. In the realm of resource allocation during epidemiological crises, Thomas et al. [13] investigated the dispensation of crucial medical supplies among various cities in India as an MDVRP model. They tackled this challenge using a genetic algorithm, resolving it for two scenarios: optimal warehouse allotment for all demand points or a restricted number of points, ensuring the shortest distance coverage for resource distribution. Bouziyane et al. [8] utilized a multi-objective VRPTW model for drug distribution with vehicle disruption conditions. They aimed to minimize the total distance traveled by all vehicles, customer delay times, and deviations from initial plans. In addition, Harms [14] addresses home delivery services for selected retail pharmacies, employing a single-depot vehicle routing model for order deliveries. Chiou et al. [15] also utilized the VRP model along with a clustering method to tackle the challenge of delivering medical supplies from a distribution center to homes. They emphasized optimizing sales drivers' routes for daily package pickups and deliveries to minimize time and distance traveled. Moreover, Pérez Lechuga [16] presented a three-stage model to streamline the optimal vehicle fleet of a perishable drug distribution firm. This model encompassed cluster analysis of demand points using a modified Dijkstra algorithm, identifying optimal routes within each cluster via a modified version of the traveling salesman model, and determining values and transport characteristics through mixed integer linear programming. Ying Ji [17] proposed a mixed integer linear programming (MIP) model to reduce the total distance traveled by vehicles in home healthcare scenarios for delivering medical supplies from pharmacies to patients, hospitals to patients, and patients to pharmacies. Shao et al. [18] developed an enhanced mathematical model to lower overall expenses and carbon emissions and optimize the pharmaceutical cold chain logistics route. By incorporating the gray wolf algorithm (GWO) with a nonlinear factor and factoring in seasonal and spatial temperature dynamics, they determined the utilization of refrigerant storage to resolve the optimization problem. Also, Ciesla and Mrowczynska [19] found the best route for delivering medicines to pharmacies by employing two methods related to the TSP. The first method involved a branch and bound approach, while the second method utilized an artificial immune system algorithm. Mostafa et al. [20] introduced A two-step approach to solve the CVRP with a heterogeneous fleet, combining a balanced K-means clustering algorithm with cutting techniques to find near-optimal solutions efficiently. The approach enhances quality, vehicle utilization, and computational time in solving benchmark CVRP instances. Alfiyatin et al. [21] combines K-means clustering and genetic algorithms to solve the VRP with time windows (VRPTW) efficiently. The genetic algorithm initialized from K-means results offers a suitable combination with low computation time, proving effective for route optimization. Addressing the CVRP, Moussa [22] proposed a recursive approach combining K-means clustering, Dijkstra's shortest path algorithm, and mathematical operations to approximate optimal routes. The focus on clustering optimization aims to provide improved solutions, acknowledging the complexity of CVRP as an NP-Hard problem. He et al. [23] proposed a new algorithm, balanced K-means for partitioning areas in large-scale VRP, demonstrating competitive results on a dataset with 1882 customers. Initially, all customers were grouped into multiple clusters, followed by a border adjustment technique in the subsequent phase to ensure balanced distribution of customers across the clusters. The exhaustive details of the reviewed studies and the current study are summarized in Table 1.

In a comprehensive review of the research literature on drug supply chain management, the primary focus has been on various types of VRP models, such as MDVRP. These studies have utilized different strategies to optimize distribution routes and designate central warehouse locations. However, most of these studies have traditionally treated the location and quantity of central warehouses as fixed assumptions. In studies where the assumption of fixed central warehouse locations is relaxed, researchers turn to common clustering methods. This study introduces an innovative two-step clustering method that considers vehicle capacity, geographical location, and demand intricacies. The method determines distribution center locations and optimizes drug distribution routing efficiently. The uniqueness of this method lies in its dual approach to performance evaluation, emphasizing comprehensive assessment of demand allocation, clustering strategies, and routing efficiency. This pioneering method offers a potentially more effective solution for MDVRP compared to existing methods in the literature, offering a fresh perspective on addressing this critical logistics challenge.

Problem statement and formulation

This study presents a novel procedure for effectively addressing the single-objective MDVRP, a problem that is known to be NP-hard [24]. The steps of the proposed approach aimed at enhancing the delivery of pharmaceutical supplies to hospital pharmacies are depicted in Fig. 1. The mathematical formulation of MDVRP, encompassing the inputs, decision variables, and constraints and the details of the proposed approach are described in the following subsections.

MDVRP assumption

The MDVRP entails vehicles providing services to hospitals from multiple drug distribution centers and involves determining which hospitals are assigned to which center. Within our problem, each center has a specific number of vehicles available and a capacity limit, referred to as Q. The following considerations are taken into account:

- Every vehicle starts its route from a centralized pharmacy (CS) and serves a specific group of hospitals. It is mandatory for the vehicle to return to the starting CS, and each hospital can only be visited once by a vehicle.
- The total demand for each route must not exceed the vehicle's capacity Q. In other words, the drugs transported by the vehicle cannot surpass its maximum limit, which is determined by the CS and restricted.
- Each hospital is served by exactly one vehicle
- It is not permitted for a CS to serve another CS. The primary objective of the MDVRP in this scenario is to efficiently establish connections between hospitals and their respective CS, minimizing the usage of resources and reducing the overall distance traveled by vehicles across the network.
- The n hospitals are grouped into m clusters.

To begin with, there is a graph $G = (V, E)$ consisting of nodes and arcs/edges connecting each node pair. The node set V is categorized into two subsets: V_h representing hospitals requiring drugs, and V_c representing the drug distribution centers. Each hospital $v_i \in V_h$ has a demand of d_i, a non-negative value. For each arc in E, there is a corresponding cost c_{ij} denoting distance. We assume that the matrix of C follows symmetry, ensuring distances between points satisfy the triangle inequality. V_k denotes all available vehicles, each with a capacity of Q. Thus, a mathematical formulation can be derived based on previous studies [24–26].

Table 1. Summary details of literature review

Authors	Model	Objective function (Min)	Solution methods	Vehicle capacity	Demands	Clustering method
Midaoui et al. [2]	MDVRP	Total distance	GA + one step clustering	Given	HBC[1]	Weighted K-means
Oliveira et al. [9]	VRPTW	Transportation time, Total costs, and CO2 emissions	TSP + one step clustering	Estimated	Given	Not mentioned
Magalh et al. [10]	Dynamic VRP	Delivery time	Not Given	Given	Given	SFC[2]
Redi et al. [11]	CVRP	Travel time	SA & NN[3]	Given	Estimated	-
KOSE et al. [12]	m-TSP	Total distance	TSP	Not Given	Not Given	-
Thomas et al. [13]	MDVRP	Transportation cost, Travel time	GA	Given	Known Demand	-
Bouziyane et al. [8]	VRPTW	Travel time	IMOLS[4]	Given	Dynamic	-
Harms [14]	VRP	Travel time	TSP	Unlimited	Given	-
Chiou et al. [15]	PDDTSPB[5]	Total distance	TSP + one step clustering	Given	Given	Not mentioned
Pérez Lechuga [16]	VRPTW	Travel time	TSP + one step clustering	Based on Normal distribution	Given	Dijkstra modified algorithm
Ying Ji [17]	MIP	Total distance	Exact	Given	Solomon Instances	-
Shao et al. [18]	VRP	Cost and carbon emissions	GWO	Given	Given	-
Ciesla and Mrowczynska [19]	TSP	Total distance	Branch and bound method	Given	Given	-
He et al. [23]	VRP	Cost	Two-stage clustering without optimization algorithm	Given	Given	Balanced K-means
Mostafa et al. [20]	CVRP	Total distance	Clustering + Valid Inequalities	Given	Given	K-means
Alfiyatin et al. [21]	VRPTW	Cost	GA	Given	Given	K-means
Moussa [22]	CVRP	Total distance	Mathematical methods + Dijkstra	Given	Given	K-means
Current study	MDVRP	Total distance	TSP + two step clustering	Estimated	Estimated	K-means

[1]HBC: Hospital Bed Capacity
[2]SFC: Space Filling Curve
[3]SA & NN: Simulated Annealing (SA) & Nearest Neghibor (NN)
[4]IMOLS: Improved Multiobjective Local Search
[5]PDDTSPB: Pickup-Delivery Dynamic Traveling Salesman Problem with Backhauls

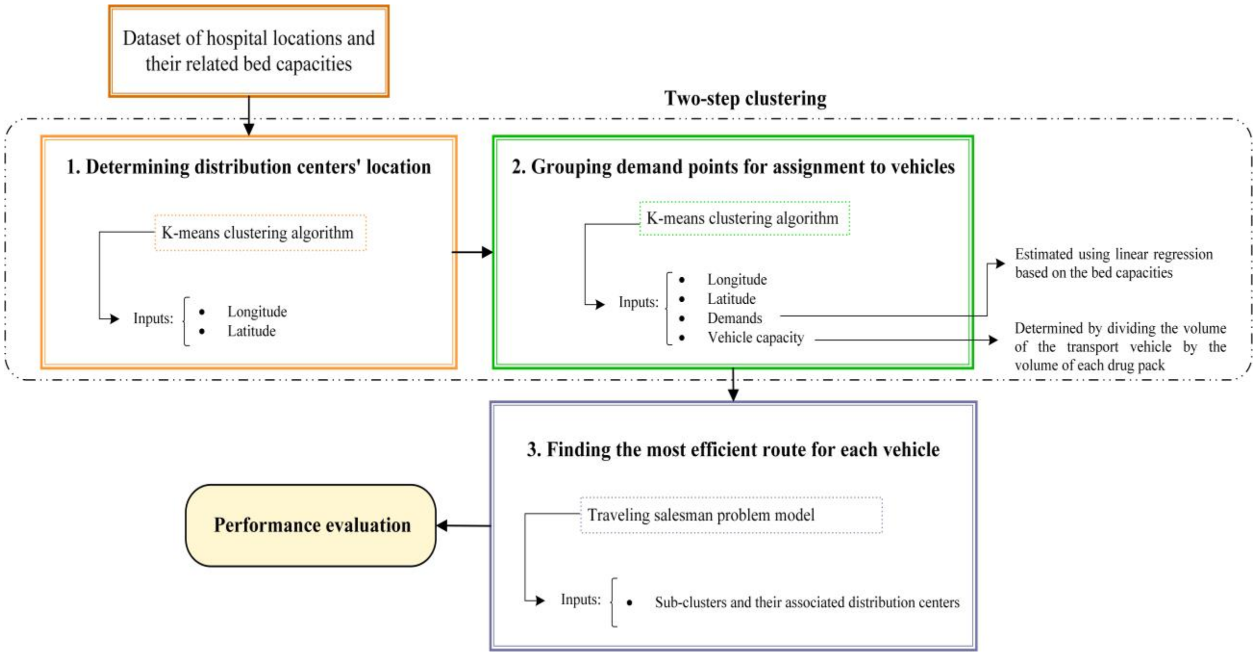

Fig. 1. Overview of the proposed approach

Mathematical model

Sets:

V_h: is the set of all hospitals
V_c: is the set of all drug distribution centers
V_k: is the set of all vehicles

Parameters:

d_i: demand of i^{th} hospital
Q: vehicles' capacity
H: Total number of hospitals
C: Number of distribution centers
c_{ij}: length of the arc (V_i, V_j); $i, j \in V_c \cup V_h$

Decision variables:

m_c: Number of vehicles in C^{th} distribution centers
x_{ijk}: Vehicle k goes directly to vertex j after vertex i or not; The vertex includes hospitals and distribution centers.
y_{ijk}: The amount of drug that vehicle k has when it leaves vertex i to vertex j.
z_i: Auxiliary variables that are needed for sub tour elimination.

$$Min\ z = \sum_{i \epsilon V_h \cup V_c} \sum_{j \epsilon V_h \cup V_c} \sum_{k \in V_k} c_{ij} x_{ijk} \quad (1)$$

$$\sum_{i \epsilon V_h \cup V_c} \sum_{k \epsilon V_k} x_{ijk} = 1 \qquad j \in V_h \quad (2)$$

$$\sum_{j \epsilon V_h \cup V_c} \sum_{k \epsilon V_k} x_{ijk} = 1 \qquad i \in V_h \quad (3)$$

$$\sum_{i \epsilon V_h} x_{ijk} = \sum_{i \epsilon V_h} x_{jik} \qquad j \in V_c, k \ \epsilon\ V_k \quad (4)$$

$$\sum_{i \epsilon V_h \cup V_c} y_{ijk} - \sum_{i \epsilon V_h \cup V_c} y_{jik} = d_j \qquad j \in V_h \cup V_c, k \ \epsilon\ V_k \quad (5)$$

$$d_j x_{ijk} \le y_{ijk} \le (Q - d_i) x_{ijk} \qquad i,j \in V_h \cup V_c\,, k \in V_k \tag{6}$$

$$m_i = \sum_{j \in V_h} x_{ijk} \qquad i \in V_c\,, k \in V_k \tag{7}$$

$$x_{ijk} = 0 \qquad i = j \in V_h \cup V_c\ \,, k \in V_k \tag{8}$$

$$z_i - z_j + (H + C) \sum_{k \in V_k} x_{ijk} \le H + C - 1 \qquad i,j \in V_h \cup V_c \tag{9}$$

$$x_{ijk}, z_i \in \{0,1\}, \qquad y_{jik} \ge 0\ \&\ integer \qquad i,j \in V_h \cup V_c\,, k \in V_k \tag{10}$$

According to Eq. 1, the objective function minimizes the total delivery distance of each vehicle within a CS. Eqs. 2-3 indicate that each hospital must be served only once by a vehicle. Eq. 4 states that all vehicles return to the center they originated from. Eq. 5 determines that the difference between the amount of drug that the vehicle has before and after visiting a hospital is equal to the demand of that hospital. Eq. 6 ensures that the demand of hospitals cannot exceed the vehicle's capacity, and the vehicle goes to a hospital that can serve its need. Eq. 7 states that the number of vehicles in a CS is equal to the number of vehicles left that CS. Eq. 8 mentions that the movement of a vehicle from a hospital or CS to itself must be avoided. Sub tour elimination is provided by Eq. 9. Finally, Eq. 10 shows the values that decision variables can take.

The proposed two-step clustering method

In MDVRP problems, the geographical placement of distribution centers plays a critical role in determining optimal solutions. In real-world scenarios, these centers are not predetermined; indeed, one of the primary objectives of this study is to detect the optimal numbers and locations of these centers. The first clustering step aims to group demand points into sections, with each section center becoming a distribution center where vehicles start their routes and return upon completion. This step leverages geographic data (latitude and longitude) of demand points.

After identifying the distribution centers and their respective coverage areas, the second clustering step comes into play. The objective of this step is to create sub-clusters within each larger cluster. These sub-clusters consist of closely located demand points, ensuring that the total demand within each sub-cluster aligns with the vehicle capacity. Effectively, this step determines which demand points each vehicle should serve. In addition to geographic data, this step considers the demand value at each point and the vehicle’s capacity.

The K-means algorithm is a widely adopted clustering method in various VRP problems [2,20–23,26], and the effectiveness of this method has been proven [2]. Hence, it is implemented in both clustering steps. The pseudo-codes for Algorithms 1 and 2 outline the steps of the proposed two-step clustering method, showcasing how K-means methodology is utilized.

Algorithm 1 Pseudo-code of the K-means clustering algorithm (first step)

Input: Number of clusters (*k*).
Output: Assignment of all points to *k* clusters
Randomly generate *k* clusters and determine the cluster centers, or directly generate *k* random points as cluster centers.
While (convergence criterion is not met or the assignment has not changed) **do**
  for each data point **do**
    Assign the point to the nearest cluster center
  end for
  Recompute the new cluster
end While

Algorithm 2 Pseudo-code of the K-means clustering algorithm considering demands and vehicle capacity (second step)

```
Inputs: Number of clusters (k), vehicle capacity (Q) and demands.
Output: Assignment of all points to k clusters
Randomly generate k clusters and determine the cluster centers, or directly generate k random points as cluster centers.
While (convergence criterion is not met or the assignment has not changed) do
   for each data point do
      Determine the closest cluster center to the current data point that has not been chosen in previous iterations
      if the nearest cluster can serve the point’s demand and stay within capacity limits, then
         Assign the point to the nearest cluster center
         Update the nearest cluster center’s status in serving demands
      end if
   end for
   Recompute the new cluster
end While
```

Determining the optimal route for each vehicle

Once the proposed two-step clustering method has determined the distribution centers and grouped the demand points (sub-clusters), a crucial step remains in solving the MDVRP: selecting the optimal sequence for vehicles to visit these points. The objective is to find a route that starts from a distribution center, visits all demand points exactly once, and then returns to the same distribution center. Consequently, solving the MDVRP entails addressing the TSP for every subcluster, ensuring efficient and effective vehicle routing.

Results and discussion

A case study is carried out focusing on Tehran hospitals to evaluate the efficacy of the proposed framework. The results obtained from this method are compared with two alternative approaches, considering both performance metrics and practical applicability. In the following subsections, key input parameters such as demand values and vehicle capacity are specified, and then a detailed step-by-step implementation of the proposed methodology is provided.

Demand estimation

Demand value is one of the main inputs of VRP models. However, there are instances where demand values may be incomplete or imprecise. In such scenarios, it is necessary to estimate missing demand values using regression techniques that clarify the relationships between demand and other available variables. Previous studies [27–29] have evidenced that simple approaches tend to outperform sophisticated methodologies when dealing with limited or small datasets [29]. Simple techniques show resilience to data constraints, enhancing their practicality and interpretability compared to modern alternatives [29]. Linear regression stands out as a reliable approach particularly suited for small datasets [27].

Linear regression indicates a relationship between the exploratory variable and the dependent variable. If only one input variable is examined, this form of linear regression is referred to as "basic linear regression". The best-fit line in linear regression is calculated using a common slope intercept [30].

The formula structure of linear regression is as follows:

$$Y_i = a + bX_i + \varepsilon \tag{11}$$

where:

- Y_i is the i^{th} dependent variable,
- X_i is the i^{th} independent variable,
- a is the intercept,
- b is the linear regression coefficient, and
- ε is the error term.

To develop the regression model, the first step involves preprocessing the dataset. This process addresses missing values and handles outliers by employing three distinct methods: replacement with mean, median, and mode values. Subsequently, the regression model evaluation was based on MSE, MAE, and R^2 metrics, with their mathematical relationships detailed in Eqs. 12-14, respectively. Upon analyzing Table 2, it is evident that replacing missing and outlier data with mean values yields superior MSE and MAE. Hence, the equation derived from this method is utilized to predict demand values. Eq. 15 can then be employed to estimate the required number of packs based on the number of beds.

Table 2. Analyzing the impact of various missing values and outliers handling methods on linear regression performance metrics

Missing values and outliers handling method	MSE	MAE	R^2
The mean value	79.5325	6.5306	0.8848
The median value	81.8719	6.8015	0.8677
The mode value	83.6919	7.2009	0.9251

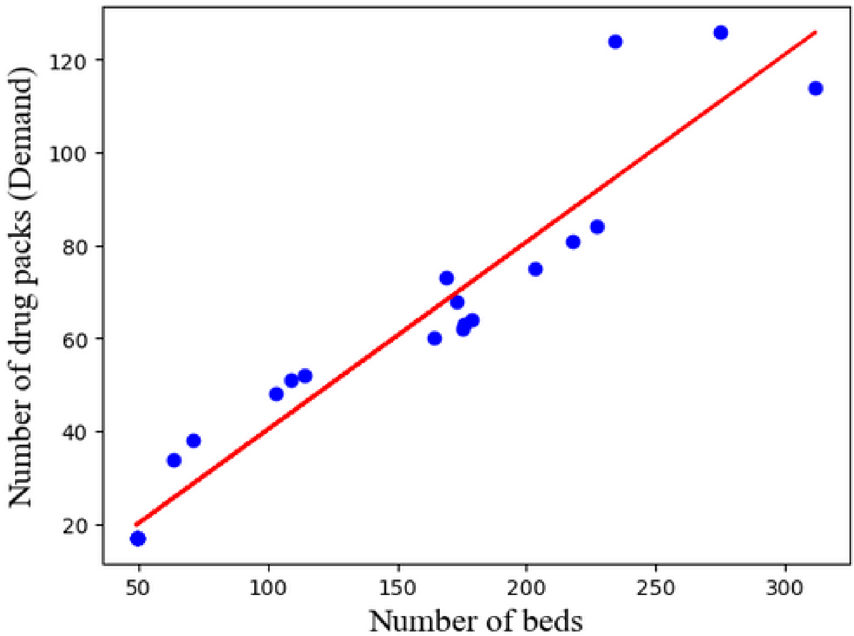

Fig. 2. Relationship between the number of beds and demand value with linear regression analysis

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{12}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{13}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{14}$$

$$D = \lceil 0.40349 \times B + 0.0534 \rceil \tag{15}$$

Here, D represents the demand, and B denotes the number of beds. Given the possibility of decimal demand values, we round the results up to the nearest whole number.

Vehicle capacity

The capacity of the vehicle is another unknown parameter that must be determined so that it can be used in the proposed method. It can be determined by knowing the volume of the vehicle and also the volume of each drug pack. If the length, width, and height of a medicine packs are 30 cm, 25 cm, and 20 cm, respectively, the volume of a single pack is 0.015 m^3. For drug deliveries, a 5-ton truck is typically used, and the dimensions of its cabin are 6 meters (length), 2.5 meters (width), and 3 meters (height). Therefore, the total volume of the truck's cabin is 45 m^3. To determine the capacity of the truck, we divide the cabin volume by the pack volume, which equals 3000 packs.

Experimental results

To implement the proposed framework, hospitals in Tehran were initially grouped according to the 22 municipal districts, similar to previous studies [2,26], as shown in Table A.1. The center of each district was then identified and considered as a single demand point. The average geographic coordinates of hospitals were calculated using a weighted average based on the number of beds in each hospital. Subsequently, the demand for hospitals within each district was aggregated to determine the demand for these centers. Table A.1 represents the number of beds and the demand values for each district.

As per the proposed approach, the first step involves clustering the demand points using K-means. To determine the appropriate number of clusters, the silhouette score is utilized, calculated as shown in Eq. 16:

$$Silhouette\ score = \frac{b-a}{max(a,b)} \tag{16}$$

Here, a represents average intra-cluster distance, while b represents average inter-cluster distance.

K-means clustering is implemented with the number of clusters set ranging from 2 to 16, and the silhouette score value is calculated. As depicted in Fig. 3, the silhouette score is highest for four clusters. Consequently, the first step of the proposed method involves conducting K-means clustering with four clusters.

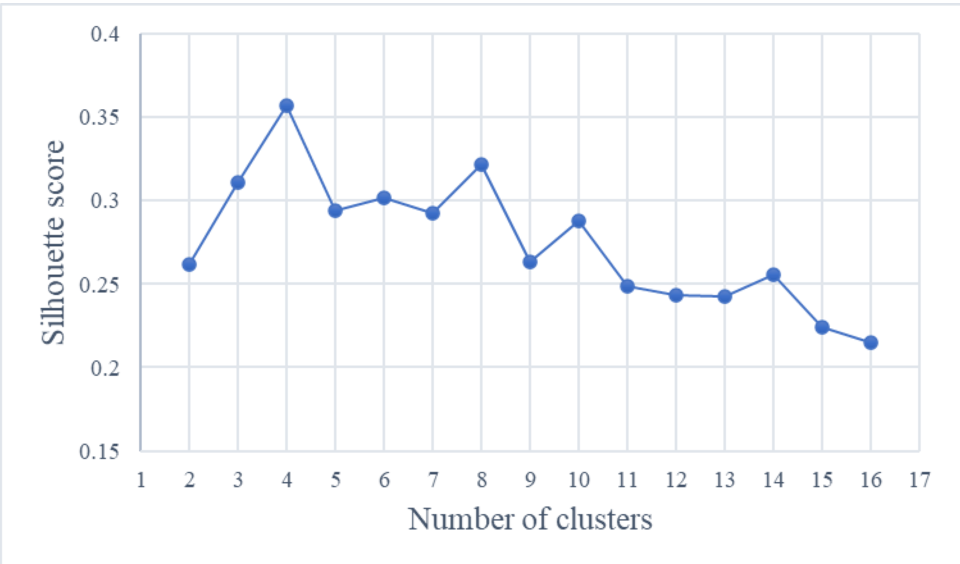

Fig. 3. Silhouette analysis of K-means clustering for optimal cluster number in the proposed method

In the next step, the clustering process is executed by determining the number of sub-clusters for each distribution center based on the demand coverage and vehicle capacity. When calculating this, if a decimal value arises, it is rounded up to ensure a whole number of sub-clusters. For instance, consider the total demand of the second cluster encompassing areas 1 to 4, which amounts to 4071 packs. Dividing this by the vehicle capacity of 3000 packs results in 1.357, indicating the necessity for 2 sub-clusters for this specific cluster. This value serves as an input for Algorithm 2, which facilitates the determination of the optimal route for each sub-cluster by solving the TSP. The results are demonstrated in Fig. 4, showcasing the proposed method's effectiveness in distributing essential medicines to Tehran hospitals. The outcomes of the proposed method are detailed in Table 3. Furthermore, it's worth mentioning that leveraging the longitude and latitude values of geographical locations enables the calculation of distances between them using Eq. 17.

$$Dis(a,b) = sin\,(\frac{\lambda_a - \lambda_b)}{2})^2 + cos\,(\lambda_a) \times cos\,(\lambda_b) \times sin\,(\frac{\varphi_a - \varphi_b)}{2})^2 \quad (17)$$

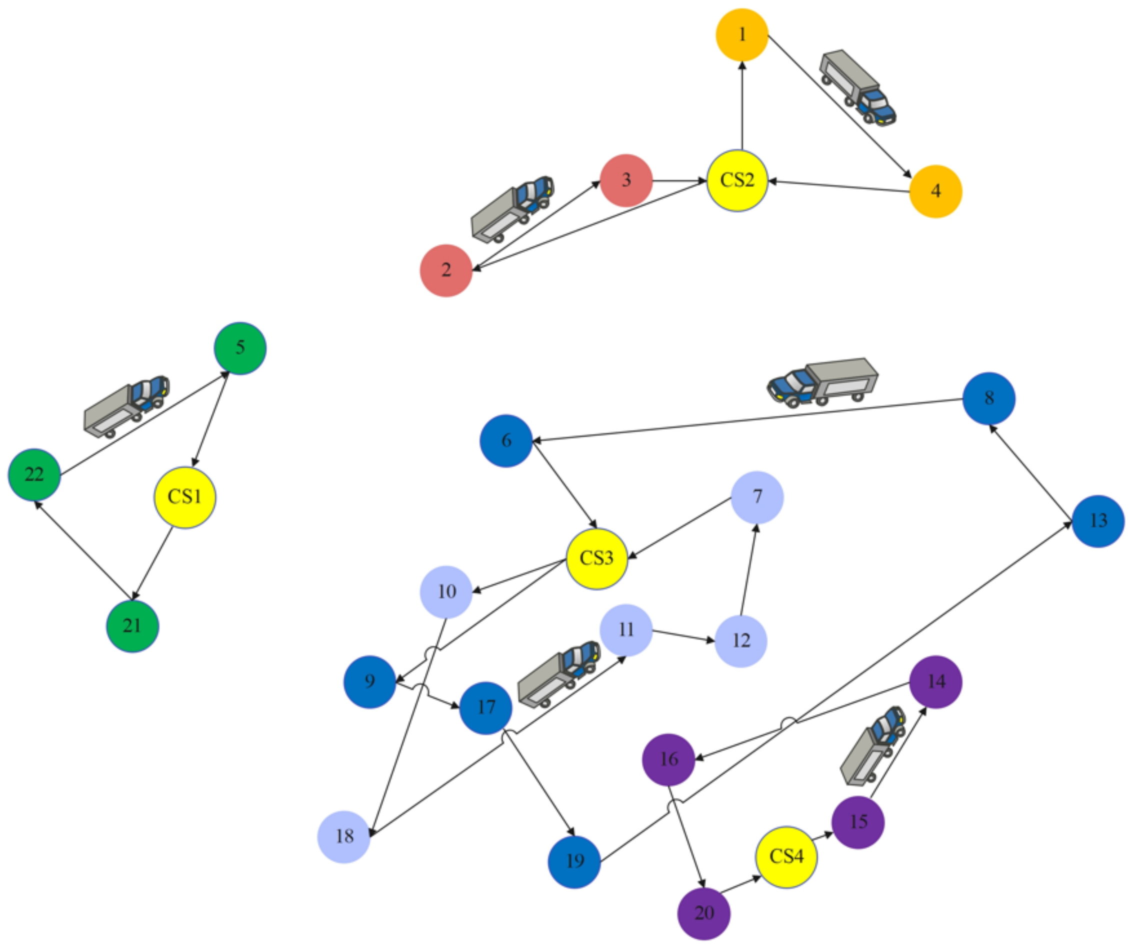

Fig. 4. Optimal vehicle routes for drug distribution among Tehran hospitals based on municipal districts using the proposed method

Table 3. The results of the proposed method

Cluster	Total delivery distance (km)	Route	Total demand
1	18.79	[CS_1, 21, 22, 5, CS_1]	1361
2	12.67	[CS_2, 2, 3, CS_2]	2543
2	18.95	[CS_2, 1, 4, CS_2]	1528
3	27.64	[CS_3, 10, 18, 11, 12, 7, CS_3]	2586
3	23.5	[CS_3, 9, 17, 19, 13, 8, 6, CS_3]	2929
4	18.13	[CS_4, 15, 14, 16, 20, CS_4]	1540

Comparison analysis results

To better evaluate the performance of the proposed approach, two other methods have also been used to solve the problem of drug distribution among Tehran hospitals. One of these methods is a modified variant of the proposed method, in which the sub-clusters are determined by Algorithm 2. Then, the centers of each sub-cluster are given as input data to Algorithm 1 to determine the distribution centers. In other words, the steps of the proposed two-step clustering are replaced with each other. Another method involves using one-step clustering to identify clusters and converting MDVRP into several VRPs, which are then solved separately [25]. The VRPy Python library is utilized to solve these VRPs.

In the first step of the modified approach, with the total demand across all regions amounting to 12487 packs and considering a vehicle capacity of 3000 packs, the calculation leads to 4.162, indicating a requirement for 5 sub-clusters. Subsequently, K-means clustering is conducted with the cluster count ranging from 2 to 4, and the silhouette score is computed. Fig. 5 illustrates that the silhouette score is maximized when utilizing two clusters. Hence, the subsequent step involves conducting K-means clustering with two clusters. Following the establishment of sub-clusters and distribution centers, the optimal route for each sub-cluster is determined by solving the TSP, with the results displayed in Fig. 6. The outcomes of this methodology concerning the distribution of essential medications for Tehran hospitals are detailed in Table 4.

For the implementation of the second alternative approach, the process begins with the application of weighted K-means clustering, utilizing varying cluster counts ranging from 2 to 16, and subsequently evaluating the silhouette score. The analysis represented in Fig. 7 reveals that the highest silhouette score is achieved with 2 clusters. Consequently, the succeeding stage of this methodology entails carrying out K-means clustering specifically with 2 clusters. Upon the delineation of clusters and distribution centers, the MDVRP is subdivided into two separate VRPs, the solutions of which are illustrated in Fig. 8. The outcomes generated by this approach for the distribution of essential medicines required by Tehran hospitals are succinctly presented in Table 5.

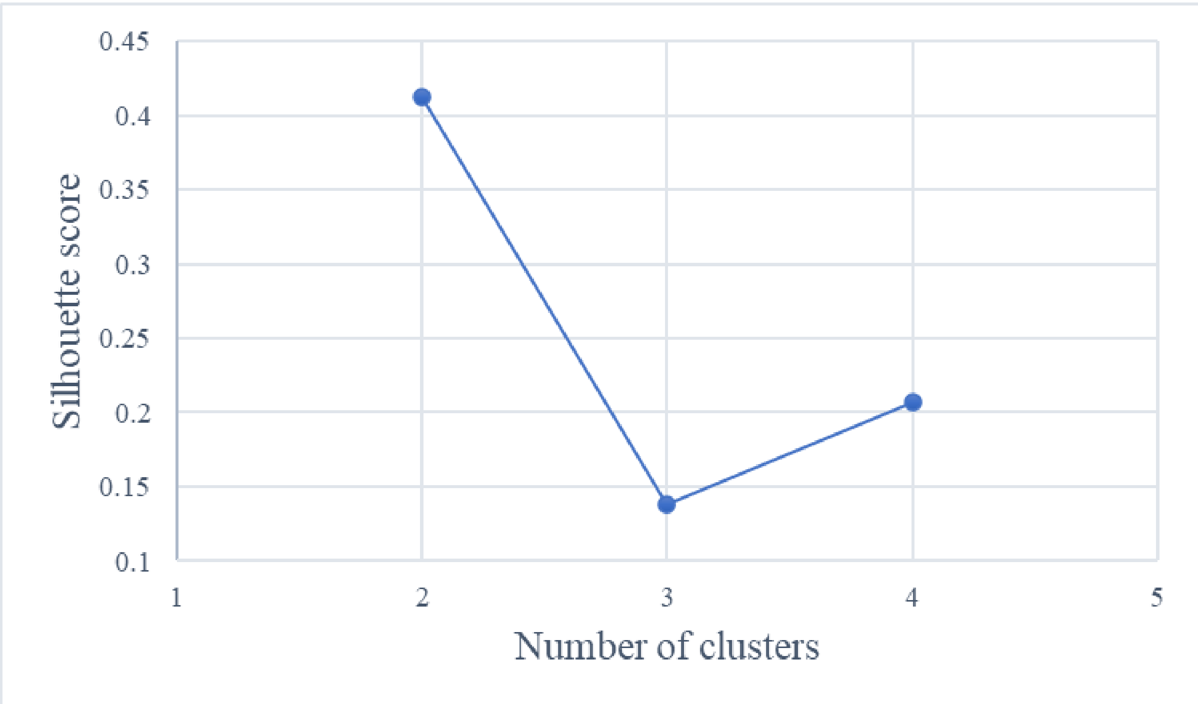

Fig. 5. Silhouette analysis of K-means clustering for optimal cluster number in the modified variant of the proposed method

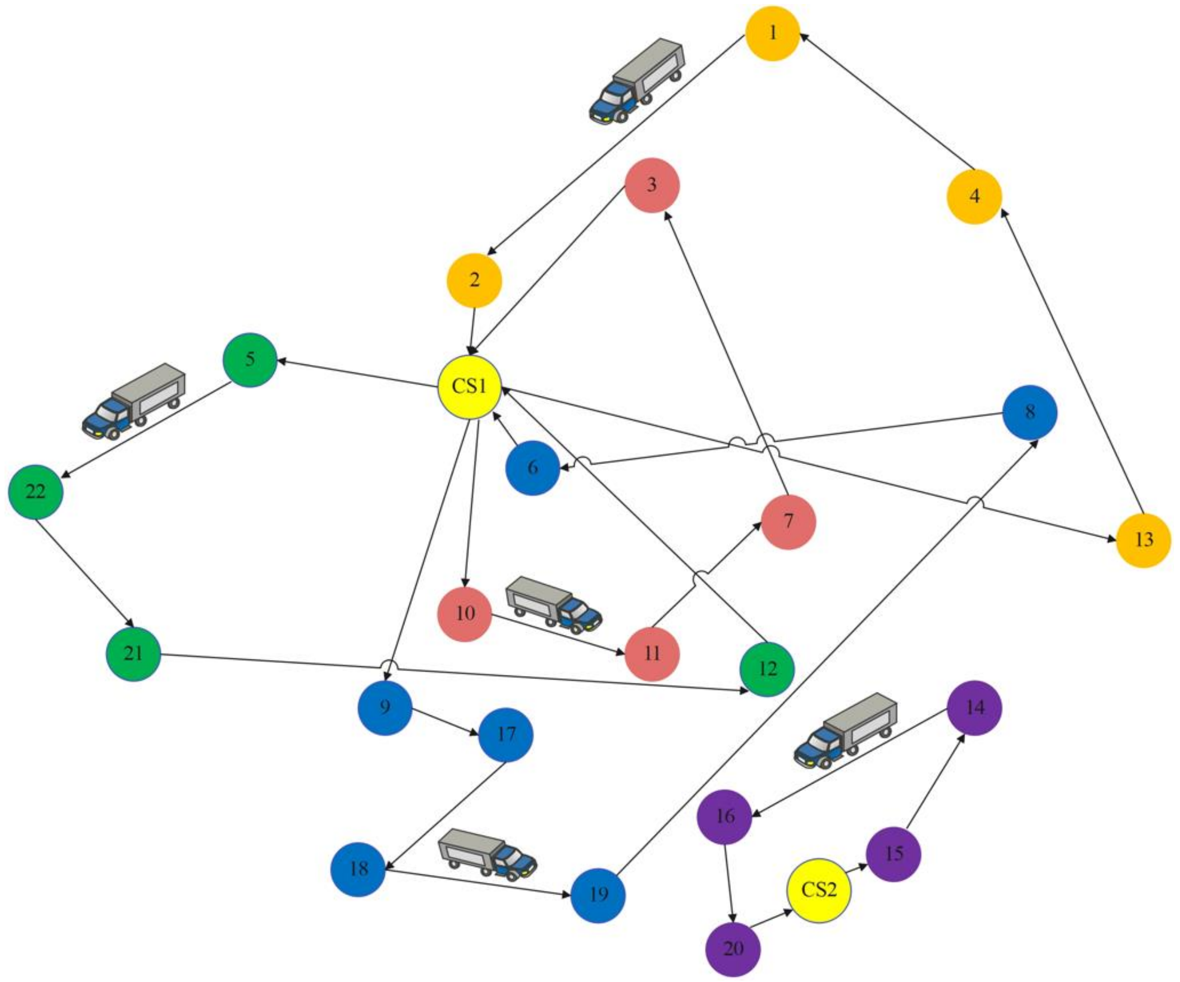

Fig. 6. Optimal vehicle routes for drug distribution among Tehran hospitals based on municipal districts using the modified variant of the proposed method

Table 4. The results of the modified variant of the proposed method

Cluster	Total delivery distance (km)	Route	Total demand
1	41.28	[CS_1, 5, 22, 21, 12, CS_1]	2195
1	25.51	[CS_1, 9, 17, 18, 19, 8, 6, CS_1]	2911
1	19.38	[CS_1, 10, 11, 7, 3, CS_1]	2913
1	34.99	[CS_1, 13, 4, 1, 2, CS_1]	2928
2	18.13	[CS_2, 15, 14, 16, 20, CS_2]	1540

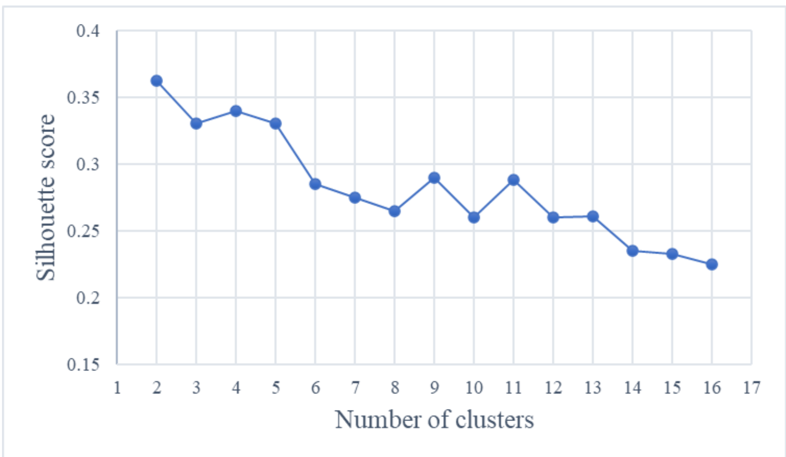

Fig. 7. Silhouette analysis of K-means clustering for optimal cluster number in the method with one step clustering + VRP

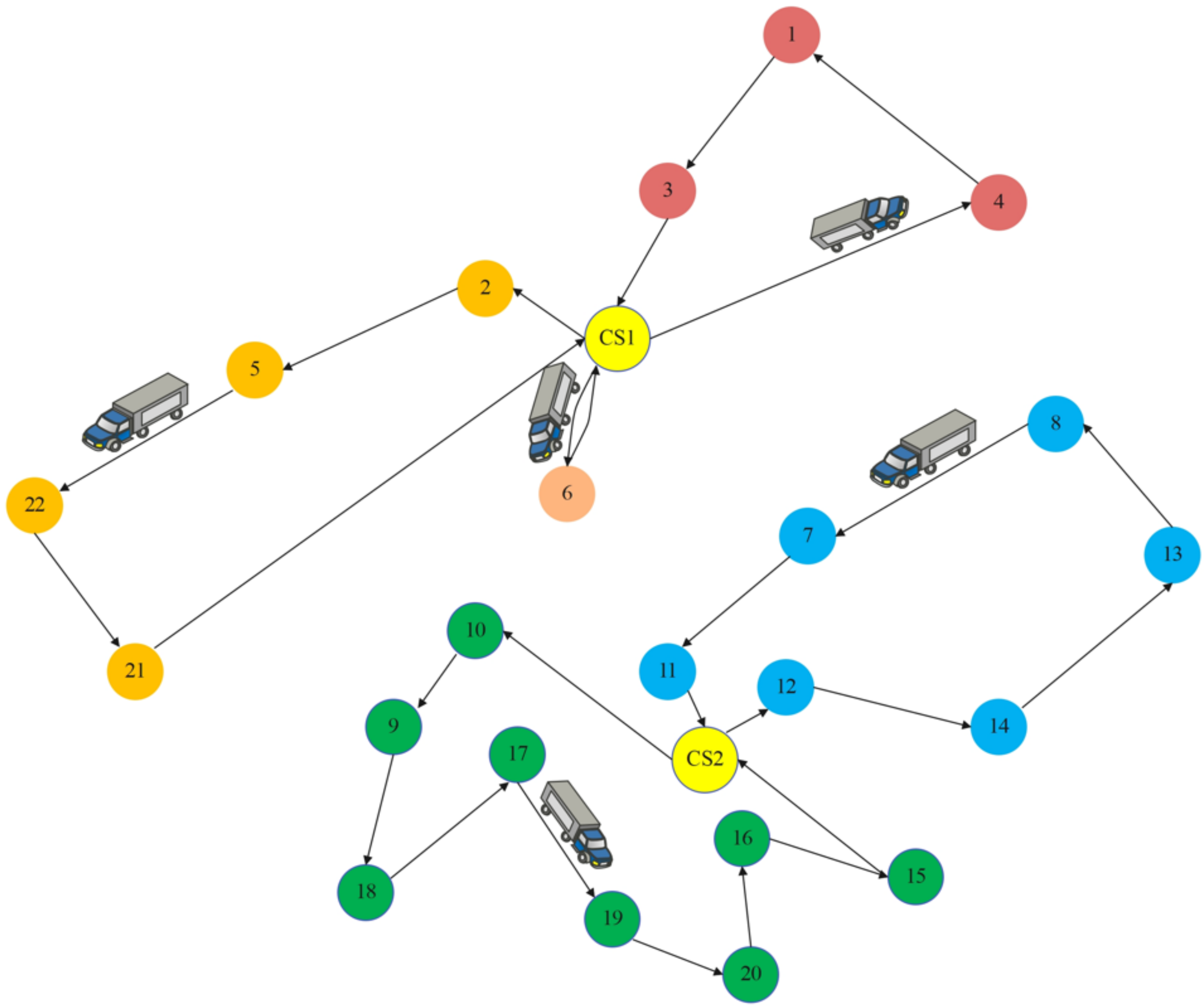

Fig. 8. Optimal vehicle routes for drug distribution among Tehran hospitals based on municipal districts using the method with one step clustering + VRP

Table 5. The results of the method with one step clustering + VRP

Cluster	Total delivery distance (km)	Route	Total demand
1	39.14	[CS_1, 2, 5, 22, 21, CS_1]	2652
1	9.6	[CS_1, 6, CS_1]	2582
1	19.56	[CS_1, 4, 1, 3, CS_1]	2780
2	27.77	[CS_2, 12, 14, 13, 8, 7, 11, CS_2]	2616
2	43.98	[CS_2, 10, 9, 18, 17, 19, 20, 16, 15, CS_2]	1857

The comparison of the proposed method with the two approaches introduced in this section is presented in Table 6. The results reveal that the proposed method effectively meets demands within a 119.68 km radius, showcasing superior performance compared to alternative methods. Notably, the proposed method employs six vehicles, slightly more than the five vehicles utilized by the other two methods, resulting in a 13.88% reduction in car capacity utilization. It's worth highlighting that the proposed method utilizes four distribution centers, whereas the other methods suffice with just two centers. Given the paramount importance of timely medication delivery in terms of both patient well-being and financial costs, the selection criteria for the proposed method prioritizes minimizing distance traveled as the primary objective function. Future studies may delve into a comprehensive analysis of the problem, exploring aspects such as the associated costs of establishing distribution centers and deploying additional vehicles within the objective function. This multi-objective approach can provide additional insights for optimizing the distribution process.

Table 6. Comparison of the proposed method with other two approaches

Method	Number of depots	Number of vehicles	Filled vehicle capacity (pack)	Distance (km)
The method with one step clustering + VRP	2	5	2497.4 (83.25%)	140.05
The modified variant of the proposed method	2	5	2497.4 (83.25%)	139.29
The proposed method	4	6	2081.16 (69.37%)	119.68

Managerial discussion

The findings of this study significantly contribute to informed decision-making within the realm of drug supply chain management. The primary objective is identifying the most efficient routes for drug distribution to hospital pharmacies. The benefits of finding these routes include reducing drug delivery time, cost savings, and efficient resource allocation. Reducing the time of drug delivery is imperative because any delay can directly impact people’s health and lives. Spoilage or shortages of essential drugs can have severe consequences [7]. Inefficient routes lead to increased fuel consumption and vehicle wear and tear, incurring unnecessary costs. The study's results highlight substantial cost reductions and savings achievable through route optimization. Moreover, by incorporating factors such as pharmacy demand distribution, vehicle capacity, and operational accessibility in the clustering process, the research effectively determined the optimal number and locations of distribution centers. This approach results in well-allocated paths and optimal resource utilization, enhancing overall operational efficiency. Additionally, by Leveraging regression methods, pharmacy demands are estimated accurately. These robust estimates enhance decision-makers confidence, enabling them to allocate resources effectively.

Conclusion

Healthcare facilities and hospitals experience delays in drug delivery and distribution because of inadequate infrastructure and the absence of decentralized central pharmacies. In this study, an innovative two-step clustering method was employed to determine optimal locations for establishing new distribution centers and to develop routing plans for drug distribution at the municipal district level in Tehran. The numerical results indicate that the demand of hospitals in Tehran can be met using five vehicles from four distribution centers, with a total travel distance of 119.68 km. The optimal number of clusters (distribution centers) was determined using the silhouette score, which calculated as 0.3567 for four clusters. The performance of the proposed method was compared with two other methods, and the results showed its superiority in minimizing the objective function, with an impressive improvement of over 14% in reducing the total travel distance to meet the demands. The proposed approach has successfully designed an effective drug distribution system, which not only reduces costs associated with storage and supplies but also enhances service quality and patient satisfaction. But it is important to note the key limitations as well as potential opportunities for our study.

In the current study, the focus is on reducing the distance traveled to deliver the drug to the demand points. However, as observed from the results, the number of CS and vehicles required for the proposed method exceeds that of alternative methods. Future studies could explore multi-objective MDVRP to further investigate the performance of the proposed method. Additionally, to enhance realism, incorporating a variety of drugs with different transportation considerations and utilizing diverse types of vehicles could be considered in addressing the problem.

References

[1] J.K. Jain, S. Banerjee, J. Jain, G.S. Dangayach, G. Agarwal, Supply Chain Management: Literature Review and Some Issues, 2010. https://www.researchgate.net/publication/268361779.

[2] M. EL Midaoui, M. Qbadou, K. Mansouri, A novel approach of smart logistics for the health-care sector using genetic algorithm, Advances in Science, Technology and Engineering Systems 5 (2020) 1143–1152. https://doi.org/10.25046/aj0506138.

[3] A. Ghaffar, A. Rashidian, W. Khan, M. Tariq, Verbalising importance of supply chain management in access to health services, J Pharm Policy Pract 14 (2021). https://doi.org/10.1186/s40545-021-00352-5.

[4] M. Arora, Y. Gigras, Importance of Supply Chain Management in Healthcare of Third World Countries, International Journal of Supply and Operations Management 5 (2018) 101–106. www.ijsom.com.

[5] Sathyanarayanan S, K.S. Joseph, S.K.V. Jayakumar, A hybrid population seeding technique based Genetic Algorithm for stochastic Multiple Depot Vehicle Routing Problem, in: 2015 International Conference on Computing and Communications Technologies (ICCCT), IEEE, 2015: pp. 119–127. https://doi.org/10.1109/ICCCT2.2015.7292731.

[6] S. Zhang, W. Zhang, Y. Gajpal, S.S. Appadoo, Ant Colony Algorithm for Routing Alternate Fuel Vehicles in Multi-depot Vehicle Routing Problem, in: 2019: pp. 251–260. https://doi.org/10.1007/978-981-13-0860-4_19.

[7] S. Khoukhi, C. Bojji, Y. Bensouda, A review of medical distribution logistics in pharmaceutical supply chain, 2019.

[8] B. Bouziyane, B. Dkhissi, M. Cherkaoui, Multiobjective optimization in delivering pharmaceutical products with disrupted vehicle routing problem, International Journal of Industrial Engineering Computations 11 (2020) 299–316. https://doi.org/10.5267/j.ijiec.2019.7.003.

[9] C. Oliveira, J. Pereira, E. Santos, T.M. Lima, P.D. Gaspar, Optimization of the COVID-19 Vaccine Distribution Route Using the Vehicle Routing Problem with Time Windows Model and Capacity Constraint, Applied System Innovation 6 (2023). https://doi.org/10.3390/asi6010017.

[10] M. de Magalh, J. Pinho de Sousa, Dynamic VRP in pharmaceutical distribution-a case study, 2006.

[11] A. Perwira Redi, F. Rohmatul Maula, F. Kumari, N. Utami Syaveyenda, N. Ruswandi, A. Uswatun Khasanah, A. Chandra Kurniawan, J. Teuku Nyak Arief, K. Kby Lama, K. Jakarta Se latan, Simulated annealing algorithm for solving the capacitated vehicle routing problem: a case study of pharmaceutical distribution, Jurnal Sistem Dan Manajemen Industri 4 (2020) 41–49. https://doi.org/10.30656/jsmi.v4i1.2115.

[12] E. Köse, B. Düzenlİ, S. Çakmak, D. Vural, Medicine distribution problem between pharmacy warehouse and pharmacies, Sādhanā 47 (2022) 171. https://doi.org/10.1007/s12046-022-01938-8.

[13] Aji Thomas, Sushma Duraphe, Arvind Gupta, Harnessing the Power of Metaheuristic Algorithms for Optimal Logistics Management in Epidemic Response, International Journal of Engineering Technology and Management Sciences 7 (2023) 627–633. https://doi.org/10.46647/ijetms.2023.v07i04.087.

[14] K.S. Harms, Single-Depot Vehicle Routing Problem for Home Delivery of Single-Depot Vehicle Routing Problem for Home Delivery of Pharmaceutical Products Pharmaceutical Products, 2013. https://scholarworks.wmich.edu/masters_theses.

[15] C.-C. Chiou, W.-H. Ouyang, T.-Y. Lin, An Empirical Study of Vehicle Routing Problem with Clustering Approach for Medical Consumable Materials Delivery, in: Research and Developments in Engineering Research Vol. 1, B P International (a part of SCIENCEDOMAIN International), 2023: pp. 54–66. https://doi.org/10.9734/bpi/rader/v1/9716F.

[16] G. Pérez Lechuga, Optimal logistics strategy to distribute medicines in clinics and hospitals, J Math Ind 8 (2018). https://doi.org/10.1186/s13362-018-0044-5.

[17] Y. Ji, Optimal scheduling in home health care pharmacy-hospital-patient's vehicle routing problem, in: ACM International Conference Proceeding Series, Association for Computing Machinery, 2019. https://doi.org/10.1145/3331453.3361310.

[18] X. Shao, S. Lu, Optimizing Dual Distribution Scheme in Pharmaceutical Cold Chain for Cost and Carbon Emissions Reduction, Applied Sciences 13 (2023) 5524.

https://doi.org/10.3390/app13095524.
[19] M. Cieśla, B. Mrówczyńska, Problem of medicines distribution on the example of pharmaceutical wholesale, in: Mechanisms and Machine Science, Kluwer Academic Publishers, 2017: pp. 51–65. https://doi.org/10.1007/978-3-319-39020-8_4.
[20] N. Mostafa, A. Eltawil, Solving the Heterogeneous Capacitated Vehicle Routing Problem using K-Means Clustering and Valid Inequalities, in: Rabat, Morocco, 2017.
[21] A.N. Alfiyatin, W.F. Mahmudy, Y.P. Anggodo, K-means clustering and genetic algorithm to solve vehicle routing problem with time windows problem, Indonesian Journal of Electrical Engineering and Computer Science 11 (2018) 462–468. https://doi.org/10.11591/ijeecs.v11.i2.pp462-468.
[22] H. Moussa, Using Recursive KMeans and Dijkstra Algorithm to Solve CVRP, ArXiv (2021).
[23] R. He, W. Xu, J. Sun, B. Zu, Balanced K-means algorithm for partitioning areas in large-scale vehicle routing problem, in: 3rd International Symposium on Intelligent Information Technology Application, IITA 2009, 2009: pp. 87–90. https://doi.org/10.1109/IITA.2009.307.
[24] J.R. Montoya-Torres, J. López Franco, S. Nieto Isaza, H. Felizzola Jiménez, N. Herazo-Padilla, A literature review on the vehicle routing problem with multiple depots, Comput Ind Eng 79 (2015) 115–129. https://doi.org/10.1016/j.cie.2014.10.029.
[25] S.K. Routroy, S. Yadav, Vehicle routing problem: recent literature review of its variants, 2018.
[26] M. El Midaoui, M. Qbadou, K. Mansouri, Logistics chain optimization and scheduling of hospital pharmacy drugs using genetic algorithms: Morocco case, International Journal of Web-Based Learning and Teaching Technologies 16 (2021) 54–64. https://doi.org/10.4018/IJWLTT.2021030104.
[27] A. Burmakova, D. Kalibatienė, Applying Fuzzy Inference and Machine Learning Methods for Prediction with a Small Dataset: A Case Study for Predicting the Consequences of Oil Spills on a Ground Environment, Applied Sciences (Switzerland) 12 (2022). https://doi.org/10.3390/app12168252.
[28] E.W. Steyerberg, M.J.C. Eijkemans, F.E. Harrell, J. Dik, F. Habbema, Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets, 2000.
[29] J.M. Lichtenberg, Simple Regression Models, 2016.
[30] S. Rekabi, Z. Sazvar, F. Goodarzian, A machine learning model with linear and quadratic regression for designing pharmaceutical supply chains with soft time windows and perishable products, Decision Analytics Journal 9 (2023). https://doi.org/10.1016/j.dajour.2023.100325.

Appendices

Table A.1. Separation of Tehran Hospitals Based on Municipal Areas

Districts	Hospital Names	Number of beds	Demands (pack)
1	Noor Afshar Hospital, Shohadaye Tajrish, Nikan, Artesh 505, Ramtin, Farmaniyeh, Mahak Charity and Hospital, Chamran, Taleqani, Rezai Psychiatric Hospital, Akhtar, Rofaydeh Rehabilitation Center and Masih Daneshvari	2537	1024
2	Laleh, Parsian, Behgar Tavan Surgery Center, Rasoul Akram, Modarres, Milad, Atieh, Erfan Niayesh, Sina, Niayesh Psychiatry and Bahman	3198	1291
3	Ali Asghar Children's, Dey General, Negah Eye Hospital, HMC, Rajaie, Cardiovascular Medical and Research Center, Mofid Children's Hospital, Hedayat Hospital, Moheb Mehr Hospital, Hashemi Nezhad Hospital, Javaheri Hospital, Iranmehr Hospital, Tandis Hospital, Golestan Hospital, Kian Hospital, Mottahari Burns Hospital, Baqiyatallah Hospital, Artesh 504 and Khatam-al-Anbya Hospital	3102	1252
4	Fakhrizadeh dental specialty, Labbafinezhad Hospital, Tehranpars nuclear medicine, Arash Women’s Hospital and Alghadir Hospital	1247	504
5	Sarem Hospital, West Nikan Hospital, Ebnesina Hospital ‹Payambaran Hospital, Omid Hospital and Farhikhtegan Hospital	1057	427
6	Pars Hospital, Mehdi Clinic Hospital, Fatima Plastic and Reconstructive Surgery Hospital, Hajar Hospital, Mehr Hospital, Massoud Clinic, Mom	6399	2582

Districts	Hospital Names	Number of beds	Demands (pack)
	Fertility and Infertility Center, Kasra Hospital, Army 600 Dental Hospital, Imam Khomeini Hospital, Shariati Hospital, Sasan Hospital, Tehran Clinic Hospital, Imam Reza Hospital(Army 501), Waliasr Naja super specialty, Asia Hospital, Yas Sepid Hospital, Pasteurno Hospital, Firoozgar Hospital, Children's Medical Center, Gandhi Hotel Hospital, Panzdah-e Khordad Hospital, Sadr Hospital, Tehran Clinic Hospital, Toos General Hospital, Tehran Heart Center, Valiasr Hospital, Mehregan Hospital, Jam Hospital, Apadana Hospital, Aban Hospital, Shahram Hospital, Marvasti Hospital, Tehran National Bank Hospital, Tehran Children Hospital, Moheb Kosar Hospital, Fatima Plastic and Reconstructive Surgery Hospital and Tehran Children Hospital Taleghani		
7	Moheb Yas General Hospital, Noor Ophthalmology, Arad Hospital, Army 502 Hospital, Mehrad Hospital, Iranshahr Hospital, Pasargad Maternity Hospital, Imam Hossein Hospital, Army Family Hospital, Bahrami Children's Hospital, Qamar Bani Hashem Hospital, Resalat Hospital, Madaran Hospital and Imam Sajjad Hospital	1783	720
8	Bina Ophthalmology	195	79
9	Shariat Razavi Hospital	172	70
10	Shahryar Hospital, Lulager Hospital, Baavar Clinic, Babak Hospital, Azadi Hospital, Iqbal Hospital in Tehran and Maymant Psychiatry	733	296
11	Shahid Shourideh Hospital, Najmiyah Hospital, Shahid Fahmideh Hospital, Rozbeh Hospital, Justice Hospital, Loqman Hakim Hospital, Baharlo Hospital, Madaen Hospital, Parsa Hospital, Alborz Hospital in Tehran, Sorena Hospital, Chashm Farabi Eye Specialist and Baher Hospital	1597	645
12	Naft Company Hospital, Shahid Andrazgo Bazarganan Hospital, Amir Alam Hospital, Seyed Al Shahada Hospital, Tarafa Hospital, Third Shaaban Hospital, Akbarabadi Educational and Treatment Center, Dr. Sapir Hospital, Shafa Yahyayan Hospital, Ayatollah Ashrafi Isfahani Hospital and Shahid Moairi Hospital	2066	834
13	Bo Ali Hospital, Martyr Dr. Mohammad Baqer Lavasani Hospital and Fajr Hospital	270	109
14	Ba'ath Nahaja Hospital, People's Hospital and East Special Medical Center	567	229
15	Mahdia Hospital and Shohada Gomnam Specialty and Subspecialty Hospital	339	137
16	Mafarah Hospital, Ayatollah Kashani Hospital, Amir Al Mominin Javadieh Hospital, Delaram Sina Psychiatry and Salim Overnight Clinic	476	193
17	Ziaian Hospital, Barzoye Surgery Center and District 17 Municipal Clinic	142	58
18	Ghayathi Charity Hospital and Yaftabad Martyrs Hospital	223	91
19	Razi Skin Hospital	75	31
20	Firouzabadi Hospital, Razi Psychiatry Center, Valiasr Charity Hospital, Hazrat Siddiqa Zahra Hospital, Tehran Children's Medical Center and Shahada Haftam Tir Hospital	2429	981
21	Khaira Sahib Kausar Hospital, Shahid Sanikhani Hospital, Fayaz Bakhsh Hospital, Iran Psychiatry and Refrans Karim Health Care Center	983	397
22	Eram Surgery Center Hospital, Dr. Sami's Neurology and Neurology, Trita Hospital and Iranian Psychiatry	1330	537