



A multiobjective model for the green capacitated location-routing problem considering drivers' satisfaction and time window with uncertain demand

Kayhan Alamatsaz^{1,2} · Abbas Ahmadi¹ · Seyed Mohammad Javad Mirzapour Al-e-hashem¹

Received: 11 March 2021 / Accepted: 6 August 2021 / Published online: 20 August 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Location-routing problem is a combination of facility location problem and vehicle routing problem. Numerous logistics problems have been extended to investigate greenhouse issues and costs related to the environmental impact of transportation activities. The green capacitated locating-routing problem (LRP) seeks to find the best places to establish facilities and simultaneously design routes to satisfy customers' stochastic demand with minimum total operating costs and total emitted carbon dioxide. In this paper, features that make the problem more practical are: considering time windows for customers and drivers, assuming city traffic congestion to calculate travel time along the edges, and dealing with capacitated warehouses and vehicles. The main novelty of this study is to combine the mentioned features and consider the problem closer to the real-world case uses. A mixed-integer programming model has been developed and scenario production method is used to solve this stochastic model. Since the problem belongs to the class of NP-hard problems, a combination of the progressive hedging algorithm (PHA) and genetic algorithm (GA) is considered to solve large-scale problems. It is the first time, as per our knowledge, that this combination is implemented on a green capacitated location routing problem (G-CLPR) and resulted in satisfactory solutions. Nondominating sorting genetic algorithm II (NSGA-II) and epsilon constraints methods are used to face with the bi-objective problem. Finally, sensitivity analysis is performed on the problem's input parameters and the efficiency of the proposed method is measured. Comparing the results of the proposed solution approach with those of the exact method indicates that the solution approach is computationally efficient in finding promising solutions.

Keywords Green location routing problem · Time window · Bi-objective optimization · Genetic algorithm · Progressive hedging algorithm · NSGA-II

Introduction

While logistics costs are one of the major costs of distribution companies, managing a distribution network has always been one of the most important issues in the supply chain management. Supply chain activities include around 10% of the gross

domestic product (Simchi-Levi et al. 2008). In today's competitive marketing environment, industry efficiency is the key to success. As a result, businesses must improve their efficiency in all areas, particularly in their logistical operations (Nekooghadirli et al. 2014).

Moreover, businesses and governments understand the necessity to assess impacts of their logistics activities on the environment (Nujoom et al. 2018). The importance of the green location-routing problems is reflected in reducing fuel consumption, time loss, traffic volume, and air pollution. Nowadays, most corporations recognize and accept the need to evaluate and decrease the environmental impact of their products' manufacturing and handling. Thus, this study's main motivation is to decrease such logistics costs and total CO₂ emission by considering both location and routing problems in an integrated multiobjective model to come up with better results.

Responsible Editor: Philippe Garrigues

✉ Abbas Ahmadi
abbas.ahmadi@aut.ac.ir

¹ Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, 424 Hafez Ave, Tehran, Iran

² Department of Building, Civil and Environmental Engineering, Concordia University, Montréal, Canada

On the other hand, one of the main challenges of location routing problems (LRP) is to strike a balance between strategic and operational goals. Also, since the LRP should simultaneously deal with location and routing problems simultaneously, many different aspects, features, and parameters must be considered in this study. As a result, we would consider time windows for customers and drivers, assume city traffic jams to calculate travel time along the edges, and deal with the capacitated warehouses and vehicles. The key contributions of this work are given below that set it apart from other studies in the related literature:

- Studying a bi-objective integrated G-CLRP
- Investigating the problem in multiproduct mode
- Considering city traffic depending on vehicles' departure time
- Considering the social factor by focusing on managing the working time of drivers
- Eliminating the assumption that the slope of roads is constant at each edge

In general, the problem discussed in this paper seeks to satisfy the stochastic demand of customers by finding the best places for establishing the warehouses and routing vehicles with the lowest cost as well as by emitting the lowest carbon dioxide. Urban traffic, which has a notable impact on the travel times of product shipments and CO₂ emission is also considered in the problem. Moreover, shortage is allowed and it is assumed to be lost sales. To increase drivers' satisfaction and reduce the risk and destructive effects of overworking, their working hours during the day will be managed. Meanwhile, soft time windows (TW) are considered for each customer. Fig. 1 demonstrates an example solution of the current problem schematically.

The rest of the paper is organized as follows. In the "Literature review" section, we give a summary of discussions about the previous related studies in LRP. The problem will be stated and mathematical formulations will be given in the "Problem description and mathematical formulation" section. The "Solution approach" section provides an overview of the solution approach for the underlying problem. In the "Numerical results" section, numerical and computational results are reported. The proposed framework is implemented on a real case study in the "Isfahan case study" section. The "Sensitivity analysis" section presents the sensitivity analysis and managerial insights and finally, "Conclusion and future works" section provides concluding remarks and enlightens the future research directions.

Literature review

The first location-routing models were studied in the late 1970s and early 1980s; however, Weber (1929) first introduced the concept of location-routing in the 1960s. Over the past few years,

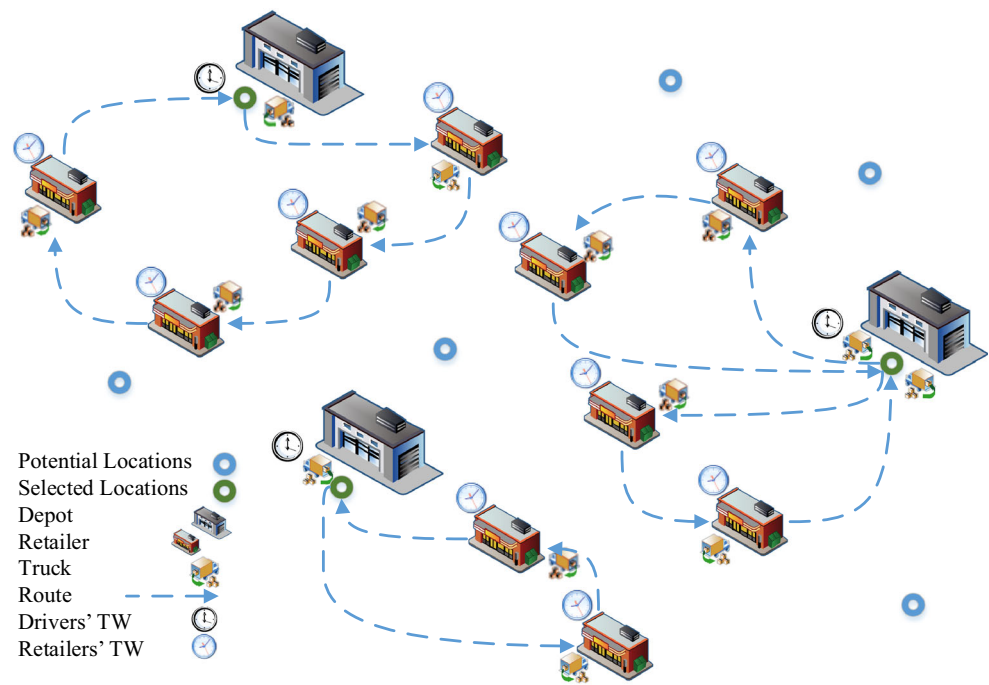
researchers have added other features to the location-routing problem. The emergence of location-routing problem with more accurate assumptions is related to the late 1990s (Min et al. 1998). In LRP, the capacity limit is usually related to the facility or the vehicle, and several researchers have examined this problem with the capacity limit, both for facilities and vehicles. This is known as capacitated location routing problem (CLRP). It means that the amount of cargo that each truck travels in its tour should not exceed the maximum fixed volume of that truck (Prins et al. 2006).

Many studies have examined the environmental aspects of product delivery. Govindan et al. (2014) considered a multiobjective location-routing problem applied to perishable food delivery. The two objective functions were to minimize the total costs and to reduce the total carbon dioxide emissions. Toro et al. (2017) raised a two-objective problem, G-CLRP, intending to minimize operating costs as well as total greenhouse gas emissions. They proposed a new approach for calculating the vehicles' total fuel consumption and subsequently the total greenhouse gas emissions. Finally, they employed the epsilon constraint method to solve the model. Tang et al. (2016) tried to minimize the total costs and the amount of carbon emissions for an inventory-location-routing problem. One of the latest studies on the green location-routing problem is the combination of fuzzy decision-making trial and evaluation (FDEMATEL) and fuzzy analysis network process (FANP). The associated multiobjective mixed-integer programming model aims to minimize total costs while reducing the shortage of uncertain fuzzy demands (Govindan et al. 2020).

The time window constraint is less focused on capacitated location-routing problems, despite its several real-world applications. Jabal-Ameli et al. (2011) presented a mathematical model for the CLRP with soft and hard time windows constraint. The objective function of the problem was to minimize the total costs. Their solution approach was based on a variable neighborhood descent and they validated their method by solving small examples. Nikbakhsh and ZEGORDI (2010) presented a two-level location-routing problem considering soft time windows for customers. They expanded a mathematical modeling and used a heuristic method to tackle the problem. Hu et al. (2019) addressed a multiobjective model for the LRP with city traffic congestion constraints. It aims to minimize total costs and total risks and to maximize customer satisfaction simultaneously. They applied genetic algorithm to tackle the problem and considered a case study to demonstrate the proposed solution approach. Zhong et al. (2020) addressed a multiobjective mixed-integer nonlinear programming model for LRP with stochastic demand to manage and control operational and strategic issues in disaster recovery. The solution approach to solve the problem was based on a genetic algorithm. To encounter the bi-objective model, NSGA-II method has been adopted to produce the Pareto front.

From the social issues point of view, Zhalechian et al. (2016) extended a sustainable mixed-integer model for the inventory-

Fig. 1 Example of a solution to the problem



location-routing problem, which involved a variety of uncertainties in demand. The feature of their proposed model is the study of potential new job opportunities. Minimizing costs and maximizing the job opportunities were the two objectives of this model, which were solved by a metaheuristic hybrid algorithm. Reducing costs and increasing public satisfaction were the two objective functions that Caballero et al. (2005) considered in their multiobjective model. They used metaheuristic methods to solve the problem. Ghaderi and Burdett (2019) studied a two-stage stochastic programming model for a bi-modal transportation network. It simultaneously minimizes the total costs and risks. Abdi et al. (2021) studied a supply chain network problem under uncertainty and suggested a new two-stage stochastic optimization approach. Also, GA was used as one of the solution methods to solve the problem. In another study, Fathollahi-Fard et al. (2020a) investigated a two-stage stochastic programming formulated on a scenario-based approach for a multiobjective problem for a closed-loop supply chain network. Biuki et al. (2020) discussed sustainability in their location-routing-inventory problem of supply chain for perishable products in a four-echelon network. The proposed multiproduct problem assumes that the type of demand is fuzzy. In addition to the particle swarm optimization method, GA has been used as a solution approach to solve this problem. Shen et al. (2019) studied emergency logistics conditions using a two-phase solution algorithm with fuzzy demand. The proposed mixed-integer programming model aims to reduce the total costs and delivery time while minimizing the total amount of CO₂ emissions. Thus, the model considers both economic and social factors.

Since location-routing is one of the NP-hard problems, researchers mostly have focused on heuristic and metaheuristic

solutions. Quintero-Araujo et al. (2017) considered the LRP with limited vehicle capacity and proposed a metaheuristic algorithm to solve it. Reducing operating costs and minimizing service time were the two objective functions of Nekooghadrilri et al. (2014). In their problem, travel time and customer demands were uncertain. Stochastic travel time was considered in (Fathollahi-Fard et al. 2020b), too. Four metaheuristic methods had been utilized to solve this problem, including multiobjective parallel simulated annealing (MOPSA), nondominated sorting genetic algorithm (NSGA-II), multiobjective imperialist competitive algorithm (MOICA), and Pareto archived evolution strategy (PAES). Zarandi et al. (2013) considered a fuzzy constraint programming and simulated annealing solution approach to solve the CLRP with time windows. Minimizing the total costs and travel distance were the two objective functions of their presented mathematical model.

Assuming lost-sales for unmet demand during the delivery process for LRP could be seen in (Chan et al. 2001), (Caballero et al. 2005), and Bozorgi-Amiri and Khorsi (2016). Bozorgi-Amiri and Khorsi (2016) investigated a multiobjective location-routing model. The objective functions were minimizing the maximum amount of unsatisfied demand, reducing costs, and decreasing travel times. The proposed model was solved by the epsilon constraint method and implemented on a case study. They assumed shortage allowance in both lost-sale and backorder. Rabbani et al. (2019) studied a stochastic location-routing problem for waste disposals recycling centers of hazardous wastes. Their multiobjective model takes into account the inventory decisions, too. The model aims to minimize the total operating costs, site risks, and transportation risks. They used a scenario-based solution approach to solve their mixed-integer

linear programming model and combined NSGA-II and Monte Carlo simulation to face with the multiobjective problem.

Table 1 presents the comparison between previous studies and the current research. As shown, the current paper is covering all of the mentioned features for the green capacitated location routing problem. This paper’s distinctive feature is to consider the combination of attributes such as capacitated trucks, shortage allowance, closed loops, social concerns, time window, and traffic. The combination of such features, which make the problem closer to real-life situations, is the paper’s main contribution.

Problem description and mathematical formulation

In the proposed model, we have a set of nodes called V , which is the collection of facilities (warehouses) I and customers’ nodes J . The problem aims to connect the customers to the facilities to minimize the total costs and CO₂ emissions. Each warehouse has a specific establishing cost, O_i , and a specific capacity, W_i . Each edge (i, j) has the cost of c_{ij} and distance of d_{ij} . Each truck has a fixed cost of F and a fixed capacity of Q_i . Each customer can be served in a specific time interval $[LB_i, UB_i]$. Also, each customer has a certain soft time window $[e_i, r_i]$. Delivery outside of this range leads to earliness and tardiness penalty costs. Drivers, like customers, have a hard $[LK, UK]$ and a soft time window $[ek_i, rk_i]$. The travel time, which affects the arrival time to customers and drivers’ working hours, depends on the time that trucks leave each node, LT_i .

In soft time window, delivery of products outside the specified time window is also allowed. More precisely, for each customer, in addition to a time window, a time interval is defined. The services performed outside the time window interval are penalized. If the fine is considered infinite, then the soft time window becomes a hard time window. For each driver, a soft time window is considered, and for the amount of working hours beyond the soft time windows, additional fees should be paid to the drivers. Thus, the amount of working hours is controlled by the amount of penalty costs paid. Customers’ demands are uncertain, and every customer requires a set of products; in other words, the problem is considered in a multiproduct mode.

This study seeks to minimize the total operational costs and total emission of carbon dioxide by calculating the amount of energy consumed by vehicles. This paper aims to simultaneously select a set of customers to meet their demands, find a set of locations for establishing the distribution centers and design the delivery routes to the customers. City traffic (traffic congestion), which affects delivery time of products and the emission of polluting gasses, is considered in the problem to approach the real world situations. Moreover, the real travel times are obtained as follows. The minimum travel time between all pairs of nodes when the streets are solitary will be

calculated. Then incremental coefficients were added to estimate the travel time in different periods of the day. Other features of the problem include closed-loop, homogenous fleet and soft time window.

The following list of assumptions are taken from (Toro et al. 2017):

- Each customer is served only once and only by one truck.
- Each customer’s demand can be satisfied only by one warehouse with limited capacity.
- Each tour starts from a warehouse and ends at the same warehouse.
- The cost of establishing warehouses is fixed.
- The cost of using vehicles is fixed.
- The size of the vehicles’ fleet is fixed and homogeneous and has limited capacities.

And the second list of assumptions are contributed to the problem as follows:

- Demand shortage assumed as lost in sales.
- Customers must be visited within a specified time window; otherwise, either earliness or tardiness penalty costs are applied.
- Drivers have a specific time for service i.e. if the driver’s working time is outside a specific range, earliness and tardiness penalty costs will be added to the total costs.
- The cost of establishing warehouses is fixed.
- The cost of using vehicles is fixed.
- The size of the vehicles’ fleet is fixed and homogeneous and has limited capacities.

The following sets, parameters and decision variables are used for the mathematical description of the problem:

Sets		
I	Set of warehouses	$i = \{1, 2, \dots, I\}$
J	Set of costumers	$j = \{1, 2, \dots, J\}$
V	Set of nodes	$v = \{1, 2, \dots, V\}$
L	Set of time interval	$l = \{1, 2, \dots, L\}$
P	Set of products	$p = \{1, 2, \dots, P\}$
Ω	Set of scenarios $\theta \in \Omega$	$\Omega = \{1, 2, \dots, \Omega \}$
Parameters		
β_{ij}	The gradient of the road from node i to j	
v	Fixed vehicles speed	
$[e_i, r_i]$	Soft time window of costumer i	
P_e	Customers earliness penalty cost	
P_l	Customers tardiness penalty cost	
LB_i	Lower bound of hard time window of customer i	
UB_i	Upper bound of hard time window of customer i	
LK	Lower bound of hard time window for each driver	
UK	Upper bound of hard time window for each driver	

Table 1 Comparing previous researches with the current study

	Objective function	Capacitated vehicles	Shortage	Closed loop	Homogenous fleet	Social issues	Time window	Traffic	Uncertainty
Caballero et al. (2005)	Min total costs, Max social satisfaction	✓	-	✓	✓	Social equity, negative implications and social denial by neighbors on the truck routes	-	-	-
Jabal-Ameli et al. (2011)	Min total costs	✓	-	✓	✓	-	✓	-	-
Zarandi et al. (2013)	Min total costs and distance traveled	✓	-	✓	✓	-	-	-	travel time and demand
Nekooghadirli et al. (2014)	Min total costs and service time	-	✓	✓	-	-	-	-	travel time and demand
Bozorgi-Amiri and Khorsi (2016)	Max reduction in unsatisfied demand, Min costs and travel time	✓	✓	-	-	-	-	-	demand
Zhalechian et al. (2016)	Min total costs, Max the positive effects of network design	✓	-	✓	✓	Job opportunities	-	-	demand
Ghaderi and Burdett (2019)	Min total costs and risks	✓	-	-	-	Risk of the transportation of hazmat	-	-	demand
Govindan et al. (2020)	Min total costs and amount of shortage	-	✓	✓	✓	-	-	-	demand
Biuki et al. (2020)	sustainable approach	✓	✓	-	-	Customer satisfaction, employee numbers, human rights, and labor working condition	-	-	demand
Rabbani et al. (2019)	Min total costs, site risk, and transportation risk	✓	✓	✓	-	-	-	-	people at risk and generated waste
Shen et al. (2019)	Min total costs, delivery time and CO ₂ emissions	✓	-	-	-	customer satisfaction	-	-	demand
Hu et al. (2019)	Min total costs and risk, Max customer satisfaction	-	-	✓	✓	customer satisfaction	✓	✓	-
Zhong et al. (2020)	Min the CVaR-R ^a of the overall cost and the waiting time	✓	-	✓	✓	-	✓	-	demand
This study	Min total costs and amount of CO ₂ emissions	✓	✓	✓	✓	Drivers satisfaction	✓	✓	demand

^a Conditional value at risk with regret (CVaR-R)

PK_e Earliness penalty cost per unit of drivers' workload time
 PK_l Tardiness penalty cost per unit of drivers' workload time
 $[ek_i, rk_i]$ Soft time window of driver i
 O_i Cost of activating of warehouse i
 W_{ip} Capacity of warehouse i of product p

M A big number
 F Fixed cost of using trucks
 Q Capacity of each truck
 $D_{jp\theta}$ Demand of customer j of product p in scenario θ
 sc Penalty cost of each product shortage

c_{ij}	Travel cost between nodes i and j
d_{ij}	Distance between nodes i and j
g	Gravity constant
b	Constant depending on the terrain
m_0	Mass of each truck
E	Total CO ₂ emissions per unit of energy (kg of CO ₂ /J)
σ_l	Coefficients for travel time between two nodes regarding trucks leaving time
$time_{ij}$	Travel time between nodes i and j when the streets are solitary
EPM	Amount of CO ₂ emissions while the vehicle is in place (working but not moving)
Decision variables	
$x_{ij\theta}$	1, if the path between nodes i and $j \in V$ in scenario θ is used; 0 otherwise
y_i	1, if facility $i \in I$ is used; 0 otherwise
$f_{ij\theta}$	1, if the customer at node $j \in J$ is served by a route that starts at the warehouse $i \in I$ in scenario θ ; 0 otherwise
$a_{ij\theta}$	1, if a vehicle uses node j to return from the end of its route (at node j) to a warehouse (at node i) in scenario θ ; 0 otherwise
$t_{ij\theta}$	The volume of load transported between nodes i and j in scenario θ
$z_{j\theta}$	1, if the customer j is the last customer served in a route in scenario θ
$sh_{ip\theta}$	The amount of shortage of product p of customer i in scenario θ
$y_{b_{il\theta}}$	Binary auxiliary variable indicating if the leaving time of customer i is in the time interval l in scenario θ
$LTL_{il\theta}$	Continuous auxiliary variable for linearization of the constraint regarding city traffic
$LT_{i\theta}$	Departure time of truck from node i in scenario θ
$g_{ij\theta}$	Binary auxiliary variable to linearize the second objective function in scenario θ
$T_{i\theta}$	Arriving time of truck in node i in scenario θ
$ye_{i\theta}$	Amount of earliness of delivering for customer i in scenario θ
$yl_{i\theta}$	Amount of tardiness of delivering for customer i in scenario θ
$yke_{i\theta}$	Amount of earliness of driver i service time in scenario θ
$ykl_{i\theta}$	Amount of tardiness of driver i service time in scenario θ

In this section, a scenario-based modeling method is presented. Our model’s goals are to minimize the total costs, (fixed costs, expected travel costs, and expected penalty costs, in all scenarios) and to minimize the total amount of CO₂ emissions. We consider a two-stage stochastic programming approach to deal with the stochastic demands. Location decisions are made in the first stage, and decisions regarding routing and scheduling customers’ visits are made in the second stage.

$$FS = \sum_{i \in I} O_i y_i \tag{1}$$

FS indicates decision variables and their values in the first stage. The formula selected for the second stage of the i_{th} objective function, SS_i , are as follows: (For more details regarding the calculation of total CO₂ emission, see (Toro et al. 2017))

$$SS_1 = \sum_{i \in I, j \in J} F a_{ij\theta} + \sum_{i, j \in V} c_{ij} x_{ij\theta} + \sum_{i \in I, j \in J} c_{ij} a_{ij\theta} + \sum_{i \in I} (P_e \times ye_{i\theta} + P_l \times yl_{i\theta}) + \sum_{i \in I} (PK_e \times yke_{i\theta} + PK_l \times ykl_{i\theta}) + \sum_{j \in J, p \in P} sc \times sh_{jp\theta} \tag{2}$$

$$SS_2 = E \left[\sum_{i, j \in V} m_0 g \left(b \cos \beta_{ij} + \sin \beta_{ij} \frac{v^2}{2gd_{ij}} \right) d_{ij} x_{ij} + \sum_{i \in I, j \in J} d_{ij} a_{ij} \right] + E \left[\sum_{i, j \in V} g \left(b \cos \beta_{ij} + \sin \beta_{ij} \frac{v^2}{2gd_{ij}} \right) t_{ij} d_{ij} \right] + EPM \sum_{i, j \in V, l \in L} \sigma_l (LT_i) * time_{ij} x_{ij} \tag{3}$$

where Ψ_1 is the sum of first stage value and the expected value of second stage variables of the first stage variables and Ψ_2 denotes the expected value of second stage variables of the second objective function. More details have been provided in the appendix (section (a)) regarding the two objective functions. The complete model is presented below:

$$\Psi_1 = FS + E(SS_{\Psi_1}) \tag{4}$$

$$\Psi_2 = E(SS_{\Psi_2}) \tag{5}$$

$$Min \Psi_1, \Psi_2 \tag{6}$$

$$\sum_{i \in V} x_{ij\theta} \leq 1 \quad \forall j \in J, \forall \theta \in \Omega \tag{7}$$

$$\sum_{k \in J} x_{jk\theta} + \sum_{i \in I} a_{ij\theta} = \sum_{i \in V} x_{ij\theta} \quad \forall j \in J, \forall \theta \in \Omega \tag{8}$$

$$\sum_{j \in J} x_{ij\theta} = \sum_{j \in J} a_{ij\theta} \quad \forall i \in I, \forall \theta \in \Omega \tag{9}$$

$$x_{ij\theta} + x_{ji\theta} \leq 1 \quad \forall I, j \in V, \forall \theta \in \Omega \tag{10}$$

$$\sum_{i \in V, i \neq j} t_{ij\theta} = \sum_{k \in V, k \neq j} t_{jk\theta} + \sum_{p \in P} D_{jp\theta} - \sum_{p \in P} sh_{ip\theta} \quad \forall j \in J, \forall \theta \in \Omega \tag{11}$$

$$\sum_{i, j \in V} x_{ij\theta} \leq |J| \quad \forall \theta \in \Omega \tag{12}$$

$$\sum_{i \in I} f_{ij\theta} \leq 1 \quad \forall j \in J, \forall \theta \in \Omega \tag{13}$$

$$t_{ij\theta} \leq Q x_{ij\theta} \quad \forall I, j \in V, \forall \theta \in \Omega \tag{14}$$

$$\sum_{j \in J} t_{ij\theta} \leq \sum_{i \in I} W_{ip} y_i \quad \forall i \in I, \forall \theta \in \Omega, p \in P \tag{15}$$

$$\sum_{k \in I} x_{jk\theta} = 1 - z_{j\theta} \quad \forall j \in J, \forall \theta \in \Omega \tag{16}$$

$$1 + a_{ij\theta} \geq f_{ij\theta} + z_{j\theta} \quad \forall i \in I, \forall j \in J, \forall \theta \in \Omega \tag{17}$$

$$-(1 - x_{ju\theta} - x_{uj\theta}) \leq f_{ij\theta} - f_{iu\theta} \quad \forall i \in I, \forall j, u \in V, \forall \theta \in \Omega \tag{18}$$

$$f_{ij\theta} - f_{iu\theta} \leq (1 - x_{ju\theta} - x_{uj\theta}) \quad \forall i \in I, \forall j, u \in V, \forall \theta \in \Omega \tag{19}$$

$$f_{ij\theta} \geq x_{ij\theta} \quad \forall i \in I, j \in J, \forall \theta \in \Omega \quad (20)$$

$$\sum_i y_i \geq \frac{\sum_j \sum_p (D_{jp\theta} - sh_{ip\theta})}{\sum_i \sum_p W_{ip}} \quad \forall i \in I, \forall \theta \in \Omega \quad (21)$$

$$\sum_{j \in J} x_{ij\theta} \leq \frac{\sum_p W_{ip}}{Q} \quad \forall i \in V, \forall \theta \in \Omega \quad (22)$$

$$\sum_{i \in I, j \in J} x_{ij\theta} \geq \sum_{j \in J} \frac{\sum_p (D_{jp\theta} - sh_{ip\theta})}{Q} \quad \forall \theta \in \Omega \quad (23)$$

$$T_{i\theta} + \sum_{l \in L} g_{ijl\theta} \times time_{ij} \times (1 + \sigma_l) \leq T_{j\theta} + M(1 - x_{ij\theta}) \quad \forall i, j \in V, \forall \theta \in \Omega \quad (24)$$

$$T_{i\theta} + \sum_{l \in L} g_{ijl\theta} \times time_{ij} \times (1 + \sigma_l) \leq T_{j\theta} + M(1 - a_{ij\theta}) \quad \forall i, j \in V, \forall \theta \in \Omega \quad (25)$$

$$T_{i\theta} \geq LB_i \quad \forall i \in V, \forall \theta \in \Omega \quad (26)$$

$$T_{i\theta} \leq UB_i \quad \forall i \in V, \forall \theta \in \Omega \quad (27)$$

$$ye_{i\theta} \geq e_i - T_{i\theta} \quad \forall i \in V, \forall \theta \in \Omega \quad (28)$$

$$yl_{i\theta} \geq T_{i\theta} - r_i \quad \forall i \in V, \forall \theta \in \Omega \quad (29)$$

$$T_{i\theta} \geq LK \quad \forall i \in V, \forall \theta \in \Omega \quad (30)$$

$$T_{i\theta} \leq UK \quad \forall i \in V, \forall \theta \in \Omega \quad (31)$$

$$yke_{i\theta} \geq ek_i - T_{i\theta} \quad \forall i \in V, \forall \theta \in \Omega \quad (32)$$

$$ykl_{i\theta} \geq T_{i\theta} - rk_i \quad \forall i \in V, \forall \theta \in \Omega \quad (33)$$

$$\sigma_l(LT_{i\theta}) = \sum_{l \in L \setminus L_{end}} \sigma_l \times yb_{il\theta} \quad \forall i \in V, \forall \theta \in \Omega \quad (34)$$

$$L_l \leq LT_i \leq L_{l+1} \quad \forall i \in V, \forall l \in L \setminus L_{end} \quad (35)$$

$$T_{0\theta} \geq 6 \quad \forall \theta \in \Omega \quad (36)$$

$$T_{i\theta} \leq 24 \quad \forall \theta \in \Omega \quad (37)$$

$$x_{ij\theta}, y_i, f_{ij\theta}, z_{i\theta}, a_{ij\theta}, yb_{ij\theta}, g_{ijl\theta} \in \{0, 1\} \quad \forall i, j \in V, \forall \theta \in \Omega, \forall \theta \in \Omega \quad (38)$$

$$t_{ij\theta}, LTL_{i\theta} \in \mathbb{R} \quad \forall i, j \in V, \forall \theta \in \Omega \quad (39)$$

$$sh_{ip\theta}, LT_{i\theta}, T_{i\theta}, ye_{i\theta}, yl_{i\theta}, yke_i, ykl_i \in \mathbb{R}^+ \quad \forall i \in I, \forall p \in P, \forall \theta \in \Omega \quad (40)$$

Constraint (7) ensures that a maximum of one edge is connected to each customer, meaning that any customer is visited by only one truck. Constraint (8) states that the input and output edges of a node must be balanced, i.e., the number of input paths to the node must be equal to the node's output paths. Constraint (9) balances the number of output paths from each facility with the number of edges entering the facility. Constraint (10) is applied to prevent duplication of the arcs at the same time. Constraint (11) sets the balance of the truck's load and demand of each customer and the amount of its shortage. The minimum number of edges needed to connect customers is expressed in constraint (12). This

eliminates subtours in the problem. Constraint (13) guarantees that customers' demand on each route must be connected to a warehouse. Constraint (14) prevents cargo overload from truck capacity. Constraint (15) limits the amount of cargo associated with each warehouse according to its capacity. Constraint (16) ensures the last visited customer on each route to return to the warehouse. Constraint (17) ensures that an edge exists which connects the last customer to a warehouse. Constraints (18) and (19) ensure that all the active edges are connected to warehouses. If an edge connects warehouse i to customer j , it is guaranteed by constraint (20) that customer j is connected to warehouse i . Constraint (21) indicates a lower bound for the number of warehouses used based on total demand and warehouse capacity. By constraint (22), the number of routes started from a warehouse is limited. Constraint (23) ensures the number of routes is sufficient to meet all customers' demands. Constraints (24) and (25) guarantee that if a truck travels from node i to node j , the time to reach node j is greater than the time spent to reach node i , according to the time spent along the edge ij . Constraints (26) and (27) state that trucks must deliver products to each customer within their hard time windows. Constraints (28) and (29) calculate the earliness and tardiness of product deliveries to each customer regarding each customer's soft time window. Constraints (30) and (31) ensure that drivers' working time should be in the required range, from 6 a.m. to 12 p.m. Constraints (32) and (33) calculate the earliness and tardiness of each driver's working hours. The arrival time of the customer j is considered as the arrival time of the previous customer plus the travel time between customers i and j , which is a function of traffic congestion at the time of leaving the customer i , $\sigma_l(LT_{i\theta})$. This is satisfied by constraint (34). In this constraint, L_{end} is referred to as the last range of time interval. Constraints (35) guarantee that the departure time of each truck be in range of time intervals. Constraints (36) and (37) ensure the daily planning interval. Constraints (38) to (40) determine the variables of the problem.

Model linearization

The second objective function of the proposed model and the constraints of urban traffic are nonlinear. The term $\sum_{i,j \in V, l \in L} \sigma_l(LT_i) \times time_{ij} x_{ij}$ is nonlinear in the second objective function. Consider σ_l as a function of LT_i is a linear expression. By adding an auxiliary binary variable, yb_{il} , assume that $\sigma_l(LT_i) = \sum_l \sigma_l \times yb_{il}$. Since x_{ij} and yb_{il} are two binary variables; thus, according to the technique used in (Liberti 2007),

Table 2 PHA algorithm

1.	$k=0$
2.	for all $\theta \in \Omega, y_{\theta}^k = \operatorname{argmin}_{y, x_{\theta}} (cy + f_{\theta} x_{\theta}) : (y, x_{\theta}) \in \Phi(y, \Omega_{\theta})$
3.	$\bar{y}^k = \sum_{\theta \in \Omega} \operatorname{Pr}_{\theta} y_{\theta}^k$
4.	for all $\theta \in \Omega, w_{\theta}^k = \tau(y_{\theta}^k - \bar{y}^k)$
5.	$k=k+1$
6.	for all $\theta \in \Omega, y_{\theta}^k = \operatorname{argmin}_{y, x_{\theta}} (cy + w_{\theta}^{k-1}y + \frac{\tau}{2} \ y - \bar{y}^{k-1}\ ^2 + f_{\theta} x_{\theta}) : (y, x_{\theta}) \in \Phi(y, \Omega_{\theta})$
7.	$\bar{y}^k = \sum_{\theta \in \Omega} \operatorname{Pr}_{\theta} y_{\theta}^k$
8.	for all $\theta \in \Omega, w_{\theta}^k = w_{\theta}^{k-1} + \tau(y_{\theta}^k - \bar{y}^k)$
9.	$\pi^k = \sum_{\theta \in \Omega} \operatorname{Pr}_{\theta} \ y_{\theta}^k - \bar{y}^k\ $
10.	if $\pi^k < \epsilon$ then go to step 5. Otherwise, terminate.

g_{ijl} , which is equal to $x_{ij} \times y_{bil}$ would a binary variable. As a result, Ψ_2 would be linearized as follows:

$$\Psi_2 = E \left[\sum_{i,j \in V} m_0 g \left(b \cos \beta_{ij} d_{ij} x_{ij} + \sin \beta_{ij} \frac{v}{2g} \right) + \sum_{i \in I, j \in J} d_{ij} a_{ij} \right] + E \left[\sum_{i,j \in V} g \left(b t_{ij} d_{ij} \cos \beta_{ij} + \sin \beta_{ij} \frac{v}{2g} \right) \right] + EPM \sum_{i,j \in V, l \in L} g_{ijl} \sigma_l \times \text{time}_{ij} \tag{41}$$

Assume we have 9 time intervals from 06:00 AM to 12:00 AM. Then, the nonlinear format of their constraints would be as follow:

$$6 \leq LT_i \leq 8 \rightarrow \sigma_1 = 1.24 \quad \forall i \in V \tag{42}$$

$$8 \leq LT_i \leq 10 \rightarrow \sigma_2 = 0.52 \quad \forall i \in V \tag{43}$$

$$10 \leq LT_i \leq 12 \rightarrow \sigma_3 = 0.64 \quad \forall i \in V \tag{44}$$

$$12 \leq LT_i \leq 14 \rightarrow \sigma_4 = 0.60 \quad \forall i \in V \tag{45}$$

$$14 \leq LT_i \leq 16 \rightarrow \sigma_5 = 0.71 \quad \forall i \in V \tag{46}$$

$$16 \leq LT_i \leq 18 \rightarrow \sigma_6 = 1.25 \quad \forall i \in V \tag{47}$$

$$18 \leq LT_i \leq 20 \rightarrow \sigma_7 = 1.23 \quad \forall i \in V \tag{48}$$

$$20 \leq LT_i \leq 22 \rightarrow \sigma_8 = 0.55 \quad \forall i \in V \tag{49}$$

$$22 \leq LT_i \leq 24 \rightarrow \sigma_9 = 0.17 \quad \forall i \in V \tag{50}$$

As it was mentioned before, σ_l is a function of LT_i and based on (Sherali and Adams 1998) the best way to linearize such constraints with the minimum number of added constraints is to add the auxiliary binary variable y_{bil} and a new binary variable $LTL_{il\theta}$. Moreover, y_{bil} indicates whether the leaving time of truck i is in the range of time interval of l , and $LTL_{il\theta}$ represents the leaving time of trucks with respect to

the time intervals. Thus, to linearize the constraints related to taking urban traffic at different hours of the day, the following constraints will be considered:

$$\sigma_l(LT_i) = \sum_l \sigma_l \times y_{bil} \tag{51}$$

s.t

$$L_l \times y_{bil\theta} \leq LTL_{il\theta} \leq L_{l+1} \times y_{bil\theta} \quad \forall i \in V, \forall \theta \in \Omega, \forall l \in L \setminus L_{end} \tag{52}$$

$$LTL_{i\theta} = \sum_{l=1}^{L \setminus L_{end}} LTL_{il\theta} \quad \forall i \in V, \forall \theta \in \Omega \tag{53}$$

$$\sum_{l=1}^{L \setminus L_{end}} y_{bil\theta} = 1 \quad \forall i \in V, \forall \theta \in \Omega \tag{54}$$

Constraint (51) represents the linear expression of coefficients for travel time between two nodes as a function of trucks leaving time. Constraint (52) ensures that the leaving time of each truck should be in the range of time intervals associated with the auxiliary variables for the linearization and constraint (53) indicates the increase in travel time ratio according to different hours of the day. Constraint (54) ensures that each truck could leave only one node in each time interval.

Solution approach

The method that has been chosen in this paper to encounter the uncertainty of demand is the combination of progressive hedging algorithm (PHA) method and genetic algorithm (GA), called PHA-GA. In the following subsections,

Table 3 Solution set of the represented chromosome

Locations	2	6	9
Customers	{3, 8, 10}	{6, 2}, {1, 4}	{7, 9}, {5}

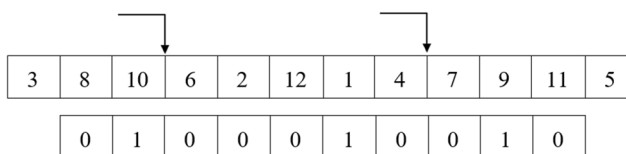


Fig. 2. An example of chromosome representation for decision variables

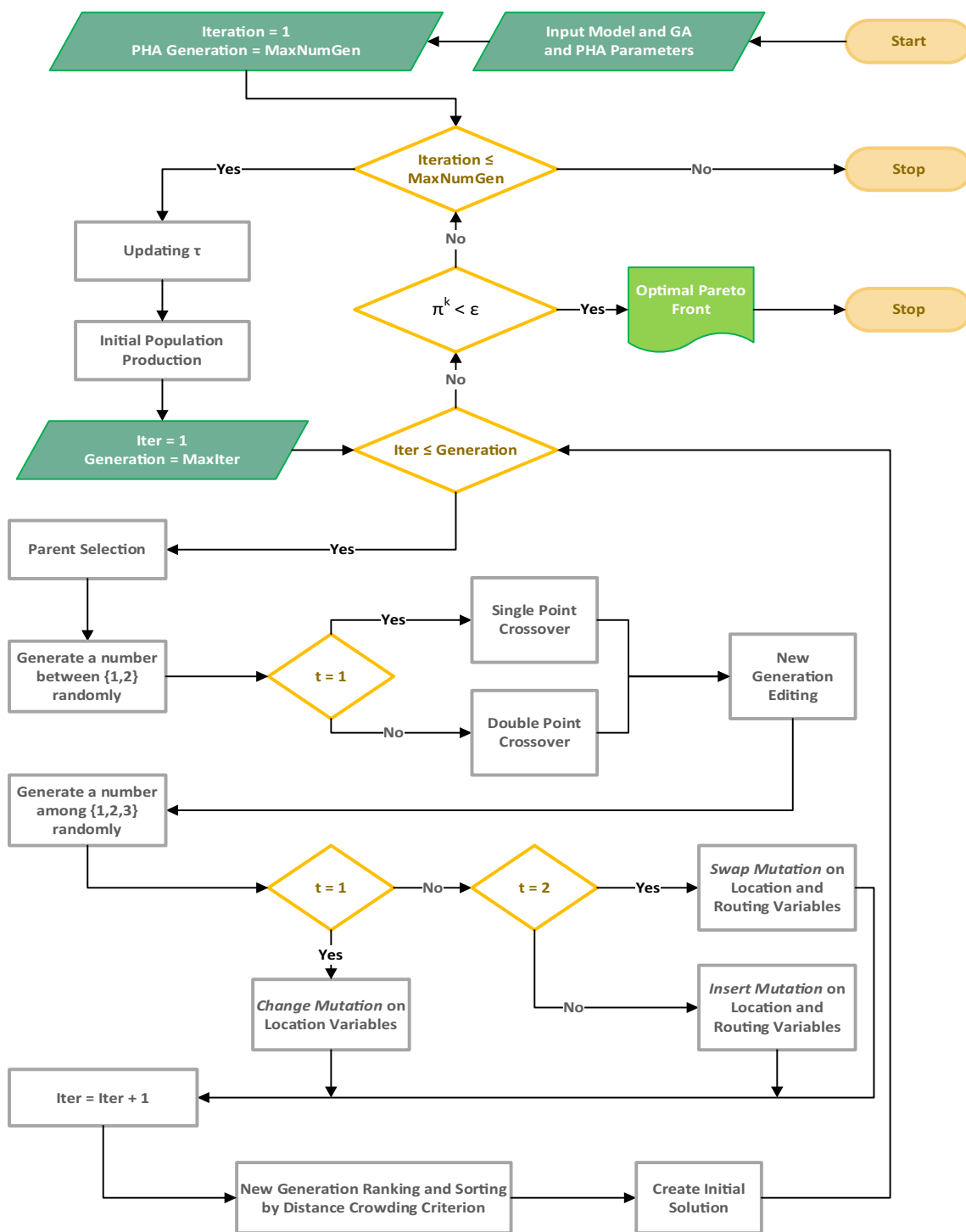


Fig. 3. The flowchart of the PHA-GA method

individually, PHA and GA methods, are briefly described, then combination of them is illustrated.

Progressive hedging algorithm

One way to solve stochastic problems based on problem division is the PHA method, introduced by Rockafellar and Wets

(1991). The problems that are less solvable due to memory limitation and computational time are made smaller and solvable by PHA. This method analyzes uncertainties based on scenarios. The important feature of this method is that when the stochastic problem space is convex, the PHA converges to the optimal solution (Løkketangen and Woodruff 1996), (Fan and Liu 2010), and (Watson and Woodruff 2011).

Table 4 Ranges from which the parameters were randomly generated

Parameter	Corresponding random distribution	Description
Q	$\sim\text{uniform}(0.5\bar{w}, 0.9\bar{w})$	\bar{w} is the average facilitates capacities
W_{ip}	$\sim\text{uniform}(1.5\bar{D}, 2.5\bar{D})$	\bar{D} is the sum of the total customers' demands for the number of potential locations
$[e_i, r_i]$	$\sim\text{uniform}(7, 10)$	In the morning
$[ek_i, rk_i]$	$\sim\text{uniform}(8, 12)$	Before noon
O_i	$\sim\text{uniform}(100, 200)$	
sc	$\sim\text{uniform}(5, 10)$	
$time_{ij}$	$\sim\text{uniform}(0.5, 2)$	

The PHA method repeatedly solves the scenarios separately and tries to gradually approach to a viable solution by adding a penalty cost to the objective function(s). Table 2 provides a general overview of the PHA algorithm. Assume that $\tau > 0$ is a penalty factor and ϵ is a convergence threshold.

The index k indicates the number of iterations, and the Euclidean distance is expressed in the k_{th} iteration as π^k . The vectors y_s^k , \bar{y}^k , and w_s^k , respectively, represent the decision variables for scenario θ in iteration k , the weighted decision variable in iteration k , and penalty cost for scenario θ at iteration k .

Genetic Algorithm (GA)

Genetic algorithm is one of the most popular metaheuristic solution approaches for solving NP-hard problems. In the proposed genetic algorithm, first, the problem parameters are determined, including the number of customers, the number of potential locations for warehouses, the number of the initial population, and customers' demands. Then, we produce the initial population given by pop .

Chromosome representation

We define the individual chromosomes for the location and routing variables by illustrating an example presented in Fig. 2. Let's suppose we have ten customers and ten potential locations for locating warehouses. Numbers greater than the number of customers (10 in this case) determine which vehicle is assigned to which customers. Then for location variables, three locations

from the set of available locations are randomly selected, and the customers of each section of the first array (chromosome) are allocated to the warehouses in the specified locations. Table 3 shows the allocation of customers to warehouses with different vehicles.

Population size and content

We use a random method to produce the initial population for location and routing variables. It is assumed that n is the number of potential locations for establishing warehouses, pop denotes the number of primary population, r is the total number of customers, and veh indicates the number of vehicles and d is the number of selected facilities.

To generate the initial population for location and routing variables, we first generate a matrix by dimension $pop \times (r + veh)$ as a random displacement of customers. Moreover, veh is randomly selected in the range of $[0, r/2]$ and is numbered in greater order than the customer index in the routing variables. This random setup determines the sequence of customers' visits and trucks' tour. In the next step, we randomly divide each chromosome by a random number in the range of $[1, n - 1]$. Then, according to the number of sections separated from each chromosome, the selected locations will be generated randomly for construction. Binary arrays will be created for each chromosome; 1 for the location where a warehouse is located and 0 for locations without located warehouse. (To find details about parent selection procedure, crossover, and mutation, see Appendix section (b))

Fitness function

Two fitness functions are examined to compare the populations produced in different generations. Total operating costs and total CO₂ emissions are the two objective functions that should be considered as fitness functions for the genetic algorithm.

$$\Psi_1 = FS + E(SS_{\Psi_1}) \tag{55}$$

$$\Psi_2 = E(SS_{\Psi_2}) \tag{56}$$

Table 5 Candidate amount of parameters for the proposed GA

Level	MaxIteration	PopSize	PCrossover, PMutation
1	100	20	0.6 , 0.4
2	200	50	0.7 , 0.3
3	500	75	0.8 , 0.2
4	1000	100	0.9 , 0.1

Fig. 4. Main effects plot of factors for GA



PHA-GA approach

It is the first time, in our best knowledge, that combination of progressive hedging algorithm and genetic algorithm is used and implemented for a green location routing problem. The “Experimental results” section presents the promising solutions of the proposed approach which demonstrates that PHA-GA approach is computationally efficient and produces satisfactory results. The flowchart of the proposed solution approach is demonstrated in Fig. 3.

The problem consists of $|\Omega|$ scenarios that their probabilities are given by pr_{θ} . We start solving the problem in order from the first scenario. We store the optimal solution for the location variables and obtain their mean based on each scenario’s probability. According to Table 2, a penalty function is multiplied by the difference in the solution of the location variables from the obtained average and is added to the cost function. This algorithm iterates until the convergence of the solutions.

In the PHA-GA method, if a solution is found in which the decision variables are equal after several iterations, the method has converged. After reaching the maximum number of generations, $MaxNumGen$, the algorithm terminates and the results are obtained.

NSGA-II

Nondominated sorting genetic algorithm II (NSGA-II) is among the most popular evolutionary algorithms for dealing with multiobjective problems. As the name implies, the genetic algorithm has an important role in this algorithm. The difference between this algorithm and single-objective genetic algorithms is in the process of sorting solutions.

The nondominating sorting algorithm quickly offers a wide range of solutions and its good performance is evident in

many two-objective problems (Rabbani et al. 2017), (Rabbani et al. 2019), (Karampour et al. 2020). The NSGA-II uses a Pareto front range and adopts the nondominating sorting mechanism to maintain the best solutions. To get acquainted with this algorithm, the concept of *rank* and *crowding distance* should be defined. See Appendix section (c) to find more information regarding NSGA-II and the concepts of rank and crowding distance, see Appendix.

Numerical results

To analyze the performance of the proposed solution method, several numerical examples in different scales are examined. These examples are solved once with CPLEX software to provide a basis for comparing the proposed algorithm solutions with the exact solutions. All instances run on the Intel Core™ i7-8750 CPU 2.20 GHz processor with 16GB of RAM.

Examples are shown in the format of R–F, where R returns the number of customers and F returns the number of candidate spots for placing warehouses. For example, format 4–3 indicates that the number of customers and potential warehouse locations are 4 and 3, respectively. Customers and facility coordinates are also randomly assigned in the range of [1,100]. The distance between nodes in the problem network is calculated based on the Euclidean distance. Other parameters are illustrated in Table 4.

Parameter tuning

The optimization design method was first used by Ronald Fisher (Yang and Tang 1998) in 1920 and Taguchi presented the *Taguchi* method as an effective and systematic approach

Table 6 Computational results

Problem size: small ; R-F ≈4–3				Problem size: small ; R-F ≈5–3			
NSGA-II (Run time=1305s)		ε-constraint (Run time=264s)		NSGA-II (Run time=1307s)		ε-constraint (Run time=264s)	
Ψ_1	Ψ_2	Ψ_1	Ψ_2	Ψ_1	Ψ_2	Ψ_1	Ψ_2
11966.49	0.00	11966.49	0.00	23036.28	0.00	23036.28	0.00
11766.49	540.87	11766.49	410.81	22918.27	369.49	22849.97	347.86
10237.31	597.69	9627.37	821.62	22476.17	668.43	19347.16	695.72
9773.49	934.42	8830.37	1232.43	21147.94	768.52	19347.16	1043.58
8731.48	1372.52	5929.49	1643.24	19348.33	963.76	18155.93	1391.44
8238.18	1627.23	5927.68	2054.05	18064.24	1391.55	16378.04	1739.30
6037.37	2083.39	5816.85	2464.86	17774.24	1868.71	16378.04	2087.16
4138.29	2281.70	2647.16	2875.67	15079.84	2133.77	13249.57	2435.02
3777.54	2951.49	2647.16	3286.47	13274.76	2679.63	12146.33	2782.88
2947.18	3286.48	2126.74	3697.28	12063.49	3132.74	11566.18	3130.74
Problem size: medium ; R-F ≈10–4				Problem size: medium ; R-F ≈15–4			
NSGA-II (Run time=1396s)		ε-constraint (Run time=10800s)		NSGA-II (Run time=1393s)		ε-constraint (Run time=10800s)	
Ψ_1	Ψ_2	Ψ_1	Ψ_2	Ψ_1	Ψ_2	Ψ_1	Ψ_2
12006.92	213.83	11993.64	0.00	24140.62	395.16	24140.62	0.00
11815.92	540.87	11793.64	418.70	23167.44	501.34	20646.81	446.33
10286.74	597.69	9654.52	837.41	22375.32	946.61	20646.81	892.66
9815.92	934.42	8857.52	1256.11	21497.51	1472.38	19452.58	1338.98
8773.91	1372.52	5956.64	1674.81	19672.94	1934.86	17670.69	1785.31
8283.61	1627.23	5954.83	2093.52	19210.38	2268.19	17668.69	2231.64
6083.80	2083.39	2674.31	2512.22	17923.07	2549.28	14545.22	2677.97
4183.72	2281.70	2674.31	2930.92	15625.81	2716.69	13441.98	3124.29
3829.97	2951.49	2674.31	3349.63	14377.29	3248.49	12857.83	3570.62
2992.61	3286.48	2153.89	3768.33	13087.24	3793.51	12644.61	4016.95
Problem size: large ; R-F ≈50–5				Problem size: large ; R-F ≈75–5			
NSGA-II (Run time=7058s)		ε-constraint (Run time≥10800s)		NSGA-II (Run time=7058s)		ε-constraint (Run time≥10800s)	
Ψ_1	Ψ_2	Ψ_1	Ψ_2	Ψ_1	Ψ_2	Ψ_1	Ψ_2
80475.29	7679.38	92384.73	0.00	81775.94	7694.88	93675.38	0.00
80226.83	7934.82	92384.73	7816.02	81527.48	7960.32	93683.38	7982.65
80216.82	13751.20	91388.47	15632.03	81516.47	13769.70	93683.38	15965.30
79971.65	17431.53	89341.26	23448.05	81269.30	17455.03	93683.38	23947.95
79216.08	24167.38	89341.26	31264.06	80514.73	24191.88	90639.91	31930.60
78467.24	36168.72	89341.26	39080.08	79766.89	36191.22	90641.91	39913.26
76926.19	45379.61	86493.10	46896.09	78226.84	45404.11	90641.91	47895.91
75733.81	53034.29	84379.68	54712.11	77034.46	53064.79	85674.33	55878.56
74268.67	64187.24	84379.68	62528.12	75560.32	64204.74	84272.16	63861.21
72631.94	68154.31	81375.46	70344.14	73924.59	68175.81	82670.11	71843.86

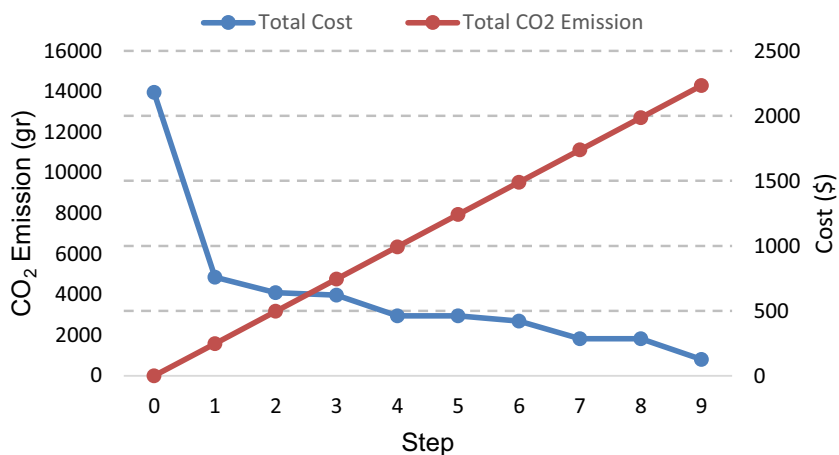
Table 7 Comparing the performance of NSGA-II and ε-constraint method

Instance	1	2	3	4	5	6
#R – F	4 – 3	5 – 3	10 – 4	15 – 4	50 – 5	75 – 5
HR	1.02	1.04	1.06	1.07	0.86	0.83

to obtain optimal parameters (Davidson et al. 2008). In this paper, for parameter tuning of the proposed genetic algorithm, the Taguchi method was implemented in Minitab 19.2. This method dramatically reduces the number of necessary tests using orthogonal arrays. To use the proposed GA, we set 3 factors as the input parameters of this algorithm. Candidate values for the parameters are determined and are given in Table 5:

Four comparison criteria are considered to validate the proposed GA: Quality metric (QM), Mean ideal distance (MID),

Fig. 5. Inverse relationship between two objective functions



Diversification metric (DM), and Spacing metric (SM) (Nekooghadirli et al. 2014). QM and MID are two quantitative criteria and DM and SM are two qualitative criteria. The weights of qualitative and quantitative criteria are 1 and 2, respectively (Asefi et al. 2014). The utility function will be as follows.

$$UF = \sqrt{(QM)^2 + (MID)^2 + (DM)^1 + (SM)^1} \tag{57}$$

The results of the implementation of the GA to investigate the direct effect of factors using the *Taguchi* method can be seen in Fig. 4:

Based on Fig. 4, it is recommended that *MaxIteration*, *PopSize*, *PMutation*, and *PCrossover* be rated at their third level so that the genetics algorithm can provide the best possible result.

Scenario reduction

In this paper, we use a scenario-based approach to deal with the stochastic problem. The number of scenarios is determined

regarding the best solution’s accuracy and can be calculated by the value of the expected objective function, given by,

$$S(sn) = \sqrt{\frac{\sum_{s=1}^{sn} (E[Cost] - Cost_s)^2}{sc-1}} \tag{58}$$

where *sn* shows the number of scenarios, *Cost_s* is the total value of the objective function for the scenario *s* and *E[Cost]* indicates the value of the expected objective function.

The $1 - \alpha$ confidence interval is given below,

$$\left[E[Cost] - \frac{z_{\alpha/2} S(sn)}{\sqrt{sn}}, E[Cost] + \frac{z_{\alpha/2} S(sn)}{\sqrt{sn}} \right] \tag{59}$$

where $z_{\alpha/2}$ is the standard deviation so that $1 - \alpha/2$ is the standard normal distribution in $z \sim N(0, 1)$, $Pr(z \leq z_{\alpha/2}) = 1 - \alpha/2$ covers. If the sample estimator *S(sn)* and the optimal confidence interval *H* are given, we can obtain the minimum number of required scenarios by,

$$N = \left[\frac{z_{\alpha/2} S(sn)}{H} \right]^2 \tag{60}$$

Fig. 6. Pareto front of a sample instance

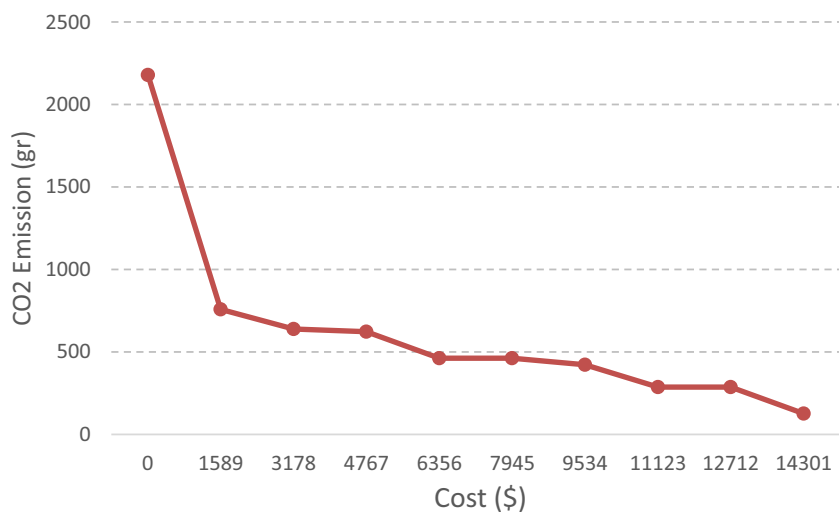
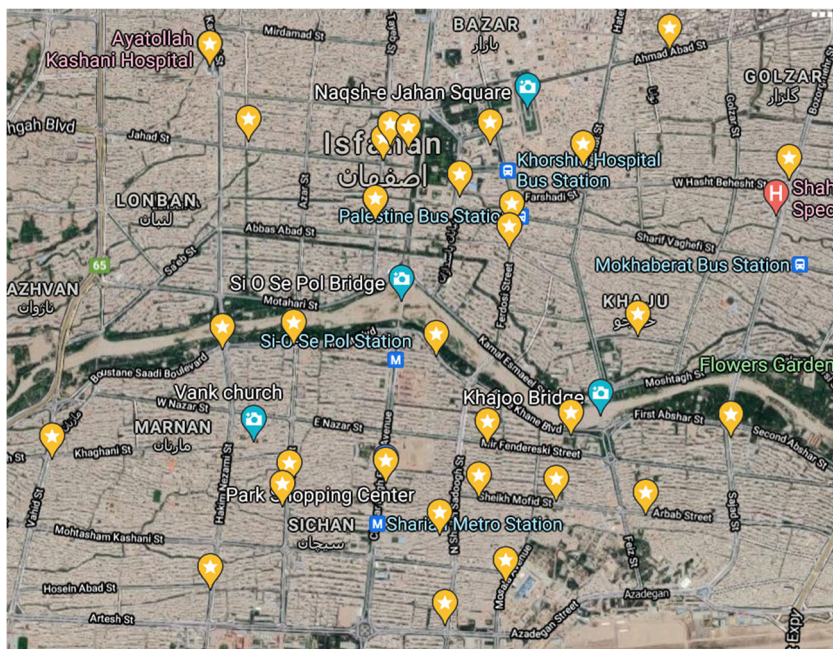


Fig. 7. Location of pharmacies and health centers in Isfahan



Thus, to specify the number of scenarios N , we will consider a small stochastic model with sn number of scenarios. We may obtain the number of scenarios regarding the desired confidence interval in order to determine the number of samples $S(sn)$.

To estimate the required number of scenarios, we first consider 30 scenarios in which the demand from the first to the last scenarios are 5, 10, 15, ..., 150, respectively. Then, we obtain the required number of scenarios, which is 8. We divide the customer demand values in the range of 5 to 150 for each product in eight intervals to get the demand for eight scenarios.

Experimental results

This section examines the results of the proposed NSGA-II method and epsilon-constraint method on the problem. The proposed PHA-GA and the exact method are coded, respectively, in MATLAB and CPLEX. The calculation time of obtaining the Pareto front of ten Pareto optimal solutions is given by *Run time*. According to Table 6, for medium and large problems, the *run time* of the NSGA-II is significantly shorter than that of the epsilon-constraint method. The maximum execution time for each step in the epsilon-constraint method is set to 1080 seconds. Hence, the obtained Pareto

Fig. 8. Pareto front of the case study

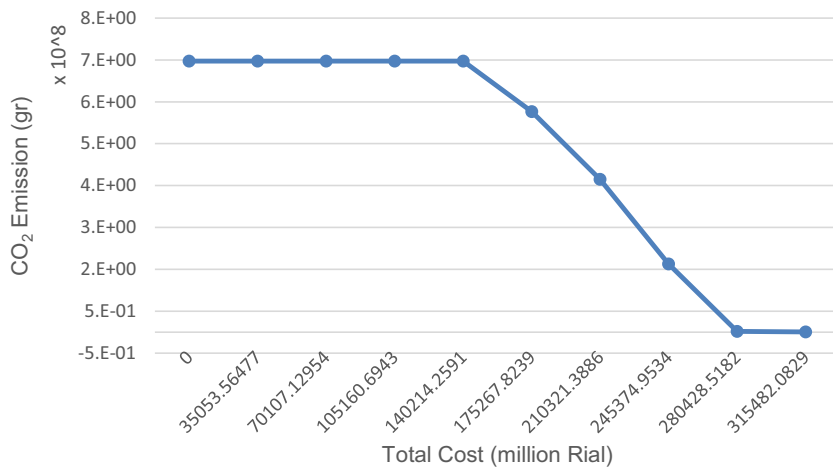


Table 8 Computational results of trucks' fixed cost changes

	1	2	3	4	5	6	7	8	9	10
Vehicle fixed cost	25	30	35	45	65	100	200	500	1000	5000
Ψ_1	926	941	956	986	1046	1151	1451	2351	3851	15851
Ψ_2	14775	14775	14775	14775	14775	14775	16703	16147	16135	16197

front solutions of medium and large problems are not optimum. As it was mentioned before, instances are presented in the format of R–F, where R returns the number of customers and F returns the number of candidate locations for placing warehouses. According to Table 6, the execution time for instances more than R–F $\approx 15 - 4$ would drastically increase due to the large scale complexity of the problem.

The *hyperarea* is a criterion used to compare the proposed solution approach's results with the exact solution's results regarding each Pareto front polygons area. The value of this criterion for each algorithm is calculated as follows (Van Veldhuizen 1999),

$$HA^A = (obj_1^1 \times obj_1^2) + \sum_{num=2} ((obj_{num}^2 - obj_{num-1}^2) \times obj_{num}^1) \quad (61)$$

where obj_{num}^1 and obj_{num}^2 are the values corresponding to the first and second objective functions in the num^{th} nondominated solution. The *hyperarea ratio* is shown by:

$$HR = \frac{HA^{NSGA-II}}{HA^{\epsilon\text{-constraint}}} \quad (62)$$

The corresponding value of this metric for instances in Table 6 are shown in Table 7. If *HR* be greater than 1, it shows that the performance of the proposed solution method is better than the epsilon-constraint method and vice versa. According to Table 7, the epsilon-constraint method performs better for small-scale problems. However, the proposed solution

method generates superior Pareto fronts for medium and large problems in a shorter time.

As depicted in Fig. 5, the total operating costs and total CO₂ emissions are in conflict. By reducing the total costs, the amount of emitted carbon dioxide increases. Such Pareto front helps managers make the right decision regarding corporations' current priorities. Subsequently, the optimal Pareto front is presented in Fig. 6.

Isfahan case study

The case study is to locate the drug depots of a small local drug distribution company (in Isfahan) and to schedule its trucks to meet the demand of a number of pharmacies and health centers, in a way that operating costs and emitted carbon dioxide will be minimized. In this case study, the drug distribution company has provided a dataset related to the demand of 30 pharmacies and 6 health centers in Isfahan. Also, the number of potential places to establish warehouses is 8. The number of health centers and pharmacies that the drug distribution company should serve is 36, drawn in Fig. 7.

The distance between customers and the distance between customers and depots are given. The slope between nodes in the city network is obtained by an approximation of Google Earth and ArcGIS software, and the travel times between nodes in the best and most secluded states are also provided. Google

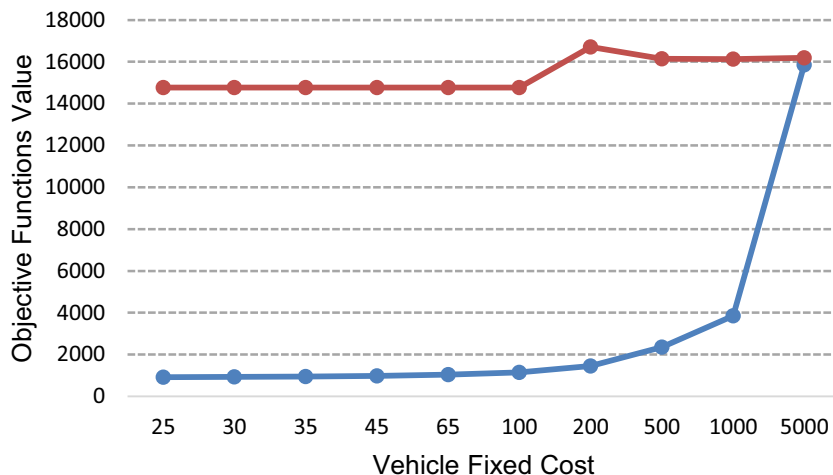
Fig. 9. Objective Functions trends vs. fixed truck costs

Table 9 Computational results of changing O_i

	Fixed facility establishment cost	Ψ_1	Ψ_2	sh	y
1	{100–150–200}	890	36194	400	{1,2,3}
2	{200–300–400}	1340	36192	400	{1,2,3}
3	{500–750–1000}	2690	36033	400	{1,2,3}
4	{1000–1500–2000}	4593	36001	3640	{1,2}
5	{1500–2250–3000}	5183	35960	6880	{1}

Map software was used to estimate the travel time between nodes.

Figure 8 shows the Pareto front of the case study. This chart helps managers make decisions and enables them to select a point in the chart according to the company’s needs and priorities.

Sensitivity analysis

We evaluate the performance of the model and its accuracy by applying changes to the problem parameters in this section. We have produced several small-scale problems in 4–3 format and in each of them, we change a specific parameter and examine its effect on the solution set and the objective functions.

We consider the effect of fixed cost changes in this problem. First, we consider the effect of fluctuations on the fixed costs of using each truck. Then, we examine the effects of fixed cost changes on establishing warehouses on the objective functions. We change the fixed cost of using each truck and keep the other parameters of the problem constant. Table 8 presents the comparative results of the objective function’s values.

According to Fig. 9, operating costs increase as fixed vehicle cost increases, but CO₂ emissions do not have a significant relationship with the fixed costs. The noteworthy point is the sudden increase in the fixed cost of vehicles. Instead of optimizing the cost of shortages by increasing the customers’ demand, the

model seeks to reduce the use of high-cost fixed vehicles. In other words, the amount of added costs caused by the demand shortage is less than the use of vehicles with such fixed cost.

To study the effects of the fixed cost for establishing facilities, several problems are generated with different fixed costs, which are illustrated in Table 9:

As seen in Fig. 10, when fixed warehouse costs increase, the total operating costs increase. The important point is an increase in the shortage due to the fixed cost of establishing warehouses and satisfying customer demand. In other words, in the trade-off between the demand shortage and establishing the warehouses, the first one is chosen.

Managerial insights

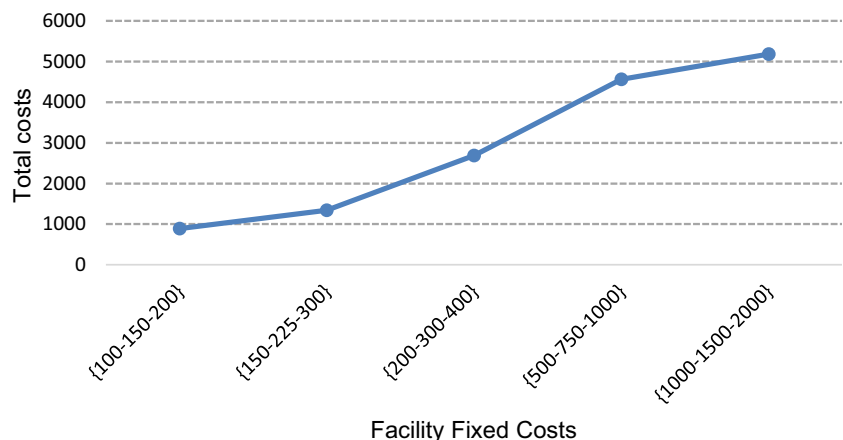
The points of view and approaches that help the distribution company managers regarding the issue studied in this paper are as follows:

- Negotiate with customers (pharmacies and health centers) to increase the time window of product delivery to reduce early and late fines for exceeding the time window.
- Negotiate with customers to share sales information to reduce costs associated with shortages and fluctuations in demand for various products.
- Operate the delivery processes on weekends due to light traffic congestion.

Conclusion and future works

Since the LRP consists of a combination of location and routing problems, and each of them is NP-hard, we encountered the problem with extraordinarily high solution time to solve, especially for large-scale problems. That is why we used genetic algorithm to solve the problem. To address the uncertainty in the problem, we also considered using the

Fig. 10 Trend of total costs towards the changes in O_i



progressive hedging algorithm. We investigated the problem by separating the main variables of the problem: location and routing decisions. As a result, we considered the problem as a two-stage problem. By using the progressive hedging algorithm, we identified the variables of the first step, which were the location variables, and then solved the routing problem. The implementation of the PHA method was performed using the genetic algorithm.

In general, the proposed problem is related to the integrated decision-making of location and routing in a wide range of areas. Researchers have done a lot of research and development over the years, but this topic is still attractive for future research because of the applicability of this problem. The following remarks are a number of aspects for developing the model:

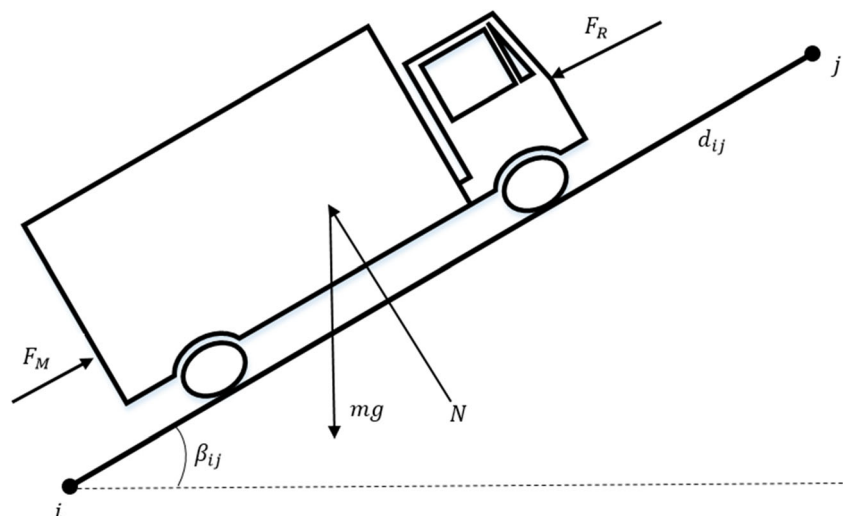
- Considering service time for each customer according to the customer's demands
- Considering the probability of disruption occurrences such as vehicle crashes and breakdowns
- Development of problem based on heterogeneous transport fleet
- Allowance of backorders for demand shortage.

Appendix

Section a) CO₂ emission formula

In Toro et al. (2017), all the forces toward the truck are discussed. Below the schematic of such forces on the truck is depicted. Using mathematical relations and static mechanics, the total amount of energy consumed, and the total amount of fuel consumed, followed by the total amount of carbon dioxide emissions, are calculated (Fig. 11).

Fig. 11. Forces act on the truck



\vec{F}_R represents the forces that are opposed to the movement of the vehicle, \vec{F}_M represents the forces created by the engine and transmitted to the tire of the vehicle, mg is the weight of the vehicle, and \vec{N} is the surface force on the vehicle.

$$\sum F_x = ma_x \quad a_x = 0 \quad F_M - F_R - mg \sin \beta_{ij} = 0 \quad (63)$$

$$\sum F_y = ma_y \quad a_y = 0 \quad N - mg \cos \beta_{ij} = 0 \quad (64)$$

$\vec{F}_{R,tires}$ represents the force created between the wheels without traction and the ground, which is opposite to the movement of the vehicle. $\vec{F}_{R,wind}$ is the force exerted by the wind against the motion of the vehicle. $\vec{F}_{R,internal}$ represents the equivalent force of internal forces opposing the motion of the vehicle, and $\frac{m v^2}{2d_{ij}}$ is the force required by the vehicle to achieve steady state dynamic energy. The mass of the loaded vehicle is the sum of the unloaded vehicle's mass (the mass of the vehicle itself) m_0 and the load carried between nodes i and j .

$$\vec{F}_R = \vec{F}_{R,tires} + \vec{F}_{R,wind} + \vec{F}_{R,internal} + \frac{m v^2}{2d_{ij}} \quad (65)$$

$$m = m_0 + t_{ij} \quad (66)$$

$$F_{R,tires} = Nb \quad (67)$$

$$\vec{F}_M = (mg \cos \beta_{ij})b + \vec{F}_{R,wind} + \vec{F}_{R,internal} + \frac{m v^2}{2d_{ij}} + mg \sin \beta_{ij} \quad (68)$$

In this step, we calculate the work function (U_{ij}) done by the truck, which is the same force in the displacement amount. Next, the total amount of gas emissions will be obtained.

$$U_{ij} = \left[(m_0 + t_{ij})gbcos\beta_{ij} + \overrightarrow{F_{R,wind}} + \overrightarrow{F_{R,internal}} + \frac{(m_0 + t_{ij})v^2}{2gd_{ij}} + (m_0 + t_{ij})g \sin \beta_{ij} \right] d_{ij} \tag{69}$$

$$U_{ij} = \left[m_0g \left(bcos\beta_{ij} + sin\beta_{ij} \frac{v_{ij}^2}{2gd_{ij}} \right) + F_{R,wind} + F_{R,internal} \right] d_{ij} \tag{70}$$

$$+ \left[g \left(bcos\beta_{ij} + sin\beta_{ij} \frac{v_{ij}^2}{2gd_{ij}} \right) \right] t_{ij}d_{ij}$$

The amount of fuel (diesel) required to do the whole work function, i.e., $\sum_{i,j \in V} U_{ij}$, is obtained by the conversion factor E_1 (gallon diesel/joule). Another conversion factor obtains each fuel unit’s emission rate, E_2 (grams of CO₂/gallon diesel). Therefore, the amount of carbon dioxide emissions is calculated as follows:

$$E_1 \times E_2 \times \sum_{i,j \in V} U_{ij} = E \times \sum_{i,j \in V} U_{ij} \tag{71}$$

Another part of the second objective function that needs to be addressed is the amount of carbon dioxide emitted in terms of the amount of time the vehicle is in traffic congestion. We can calculate this by the following equation of carbon dioxide emissions.

$$E_2 \times E_3 \times \sum_{i,j \in V, l \in L} \sigma_l(LT_i) * time_{ij}x_{ij}$$

$$= EPM \sum_{i,j \in V, l \in L} \sigma_l(LT_i) * time_{ij}x_{ij} \tag{72}$$

In this regard, the *EPM*, which is the conversion factor of the idle hours of truck operation to the amount of carbon dioxide produced, is obtained using two conversion factors. The first conversion factor, E_3 , is for gallons of diesel and the second conversion factor is equal to the same factor E_2 .

Section b) GA parent selection and operators

1. Parent selection

In the proposed GA, two types of parent selection procedures, including random selection and roulette wheel selection, are involved in the model. However, based on the results of the genetic algorithm’s implementation in several identical examples, the roulette wheel selection method has performed better than another method (Yu et al. 2019). According to (Amal and Chabchoub 2018), half of the chromosomes with the highest fitness function will be selected for the next generation.

2. Crossover

Two different types of crossover are used randomly with the same probability to produce offspring from two selected parents: single-point crossover and two-point crossover (See

(Asefi et al. 2014) for more details). The two types of crossover are depicted in Figs. 12 and 13.

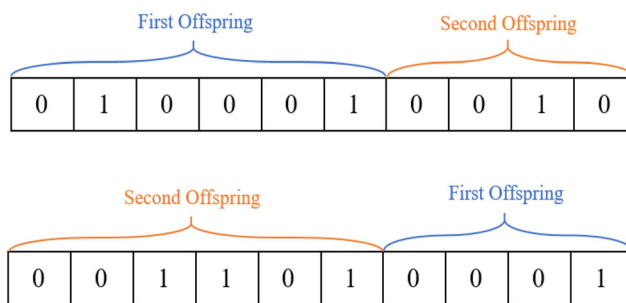


Fig. 12 An example of single-point crossover on location variable

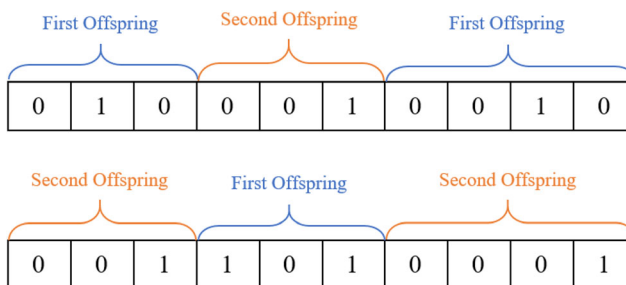


Fig. 13 An example of two-point crossover on location variable

3. Mutation

Three mutation operators will be used for location variables, and two mutation operators will be used for routing variables in the proposed GA. Mutation operators for location variables are *change*, *swap*, and *insertion*, and routing variables are *swap* and *insertion*.

In the *change* mutation operator, two random points are selected from each chromosome. Two points are chosen randomly along the chromosome. Then, as depicted in Fig. 14, if the chosen point is 0, it will be converted to 1, and vice versa. Other mutation operators are illustrated in (Asefi et al. 2014). Figs. 15 and 16 clearly show how such operators perform.

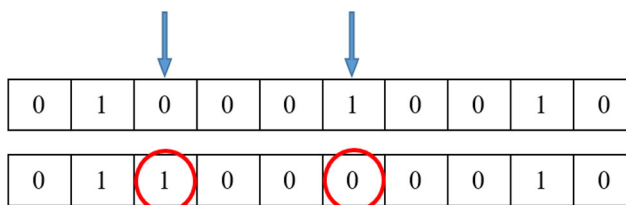


Fig. 14. An example of *change* mutation for location variables

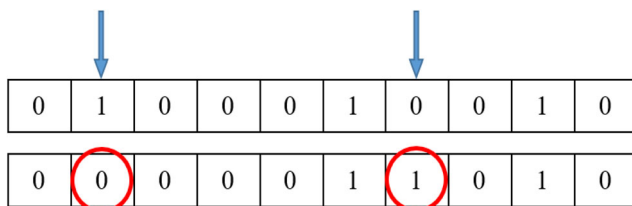


Fig. 15. An example of *swap* mutation for location variable

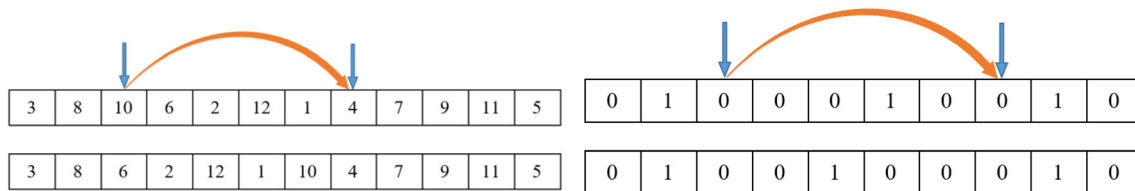


Fig. 16. An example of *insertion* mutation

Section c) NSGA-II

Ranking means that we want to rank the solutions according to the concept of quality. To rank the initial population solutions, we put the solutions that never dominated in the first place. Then, we remove non-dominated solutions from the solution set and again compare the rest of the solutions. Once again, we put non-dominated solutions on the second frontier. In this way, all the solutions will be categorized into different boundaries.

If the solutions are ranked only by the *rank* criterion, there is no need for the *crowding distance* defined by the neighbors of a solution and the first (best) and last (worst) chromosome of the population (Van Veldhuizen and Lamont 1999).

Authors' contributions Kayhan Alamatsaz: Conceptualization, Methodology, Software, Writing—Original draft, Validation, Formal analysis. Abbas Ahmadi: Supervision, Methodology, Writing—review and editing. Seyed Mohammad Javad Mirzapour Al-e-hashem: Conceptualization, Methodology, Supervision, Software, Writing—review and editing.

Data availability All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abdi A, Abdi A, Fathollahi-Fard AM, Hajiaghaei-Keshteli M (2021) A set of calibrated metaheuristics to address a closed-loop supply chain network design problem under uncertainty. *Int J Syst Sci* 8(1):23–40

- Amal L, Chabchoub H (2018) SGA: spatial GIS-based genetic algorithm for route optimization of municipal solid waste collection. *Environ Sci Pollut Res* 25(27):27569–27582
- Asefi H, Jolai F, Rabiee M, Araghi MT (2014) A hybrid NSGA-II and VNS for solving a bi-objective no-wait flexible flowshop scheduling problem. *Int J Adv Manuf Technol* 75(5-8):1017–1033
- Biuki M, Kazemi A, Alinezhad A (2020) An integrated location-routing-inventory model for sustainable design of a perishable products supply chain network. *J Clean Prod* 260:120842
- Bozorgi-Amiri A, Khorsi M (2016) A dynamic multi-objective location-routing model for relief logistic planning under uncertainty on demand, travel time, and cost parameters. *Int J Adv Manuf Technol* 85(5-8):1633–1648
- Caballero R, González M, Guerrero FM, Molina J, Parolera C (2005) A metaheuristic procedure for multiobjective location routing. In: 27th International Conference on Information Technology Interfaces, 2005. IEEE, New York, pp 462–467. <https://doi.org/10.1109/ITI.2005.1491171>
- Chan Y, Carter WB, Burnes MD (2001) A multiple-depot, multiple-vehicle, location-routing problem with stochastically processed demands. *Comput Oper Res* 28(8):803–826
- Davidson MJ, Balasubramanian K, Tagore G (2008) Experimental investigation on flow-forming of AA6061 alloy—a Taguchi approach. *J Mater Process Technol* 200(1-3):283–287
- Fan Y, Liu C (2010) Solving stochastic transportation network protection problems using the progressive hedging-based method. *Netw Spat Econ* 10(2):193–208
- Fathollahi-Fard AM, Ahmadi A, Al-e-Hashem SM (2020a) Sustainable closed-loop supply chain network for an integrated water supply and wastewater collection system under uncertainty. *J Environ Manag* 275:111277
- Fathollahi-Fard AM, Ahmadi A, Karimi B (2020b) A robust optimization for a home healthcare routing and scheduling problem considering greenhouse gas emissions and stochastic travel and service times. In: *Green Transportation and New Advances in Vehicle Routing Problems*. Springer, Berlin, pp 43–73
- Ghaderi A, Burdett RL (2019) An integrated location and routing approach for transporting hazardous materials in a bi-modal transportation network. *Transport Res E-Log* 127:49–65
- Govindan K, Jafarian A, Khodaverdi R, Devika K (2014) Two-echelon multiple-vehicle location-routing problem with time windows for optimization of sustainable supply chain network of perishable food. *Int J Prod Econ* 152:9–28
- Govindan K, Mina H, Esmaeili A, Gholami-Zanjani SM (2020) An integrated hybrid approach for circular supplier selection and closed

- loop supply chain network design under uncertainty. *J Clean Prod* 242:118317
- Hu H, Li X, Zhang Y, Shang C, Zhang S (2019) Multi-objective location-routing model for hazardous material logistics with traffic restriction constraint in inter-city roads. *Comput Ind Eng* 128:861–876
- Jabal-Ameli M, Aryanezhad M, Ghaffari-Nasab N (2011) A variable neighborhood descent based heuristic to solve the capacitated location-routing problem. *Int J Ind Eng Comput* 2(1):141–154
- Karampour MM, Hajiaghahi-Keshmeli M, Fathollahi-Fard AM, Tian G (2020) Metaheuristics for a bi-objective green vendor managed inventory problem in a two-echelon supply chain network. *Scientia Iranica*. <https://doi.org/10.24200/sci.2020.53420.3228>
- Liberti L (2007) Compact linearization for binary quadratic problems. *4OR* 5(3):231–245
- Løkketangen A, Woodruff DL (1996) Progressive hedging and tabu search applied to mixed integer (0, 1) multistage stochastic programming. *J Heuristics* 2(2):111–128
- Min H, Jayaraman V, Srivastava R (1998) Combined location-routing problems: a synthesis and future research directions. *Eur J Oper Res* 108(1):1–15
- Nekooghadirli N, Tavakkoli-Moghaddam R, Ghezavati VR, Javanmard (2014) Solving a new bi-objective location-routing-inventory problem in a distribution network by meta-heuristics. *Comput Ind Eng* 76:204–221
- Nikbaksh E, Zegordi S (2010) A heuristic algorithm and a lower bound for the two-echelon location-routing problem with soft time window constraints. *Scientia Iranica*, 17 (1 (TRANSACTION E: INDUSTRIAL ENGINEERING))
- Nujoom R, Mohammed A, Wang Q (2018) A sustainable manufacturing system design: a fuzzy multi-objective optimization model. *Environ Sci Pollut Res* 25(25):24535–24547
- Prins C, Prodhon C, Calvo RW (2006) Solving the capacitated location-routing problem by a GRASP complemented by a learning process and a path relinking. *4OR* 4(3):221–238
- Quintero-Araujo CL, Caballero-Villalobos JP, Juan AA, Montoya-Torres JR (2017) A biased-randomized metaheuristic for the capacitated location routing problem. *Int Trans Oper Res* 24(5):1079–1098
- Rabbani M, Farrokhi-Asl H, Asgarian B (2017) Solving a bi-objective location routing problem by a NSGA-II combined with clustering approach: application in waste collection problem. *J Ind Eng Int* 13(1):13–27
- Rabbani M, Heidari R, Yazdanparast R (2019) A stochastic multi-period industrial hazardous waste location-routing problem: integrating NSGA-II and Monte Carlo simulation. *Eur J Oper Res* 272(3):945–961
- Rockafellar RT, Wets RJ-B (1991) Scenarios and policy aggregation in optimization under uncertainty. *Math Oper Res* 16(1):119–147
- Shen L, Tao F, Shi Y, Qin R (2019) Optimization of location-routing problem in emergency logistics considering carbon emissions. *Int J Environ Res Public Health* 16(16):2982
- Sherali HD, Adams WP (1998) Reformulation-linearization techniques for discrete optimization problems. In: *Handbook of combinatorial optimization*. Springer, Berlin, pp 479–532
- Simchi-Levi D, Kaminsky P, Simchi-Levi E, Shankar R (2008) *Designing and managing the supply chain: concepts, strategies and case studies*. Tata McGraw-Hill Education, New York
- Tang J, Ji S, Jiang L (2016) The design of a sustainable location-routing-inventory model considering consumer environmental behavior. *Sustainability* 8(3):211
- Toro EM, Franco JF, Echeverri MG, Guimarães FG (2017) A multi-objective model for the green capacitated location-routing problem considering environmental impact. *Comput Ind Eng* 110:114–125
- Van Veldhuizen DA (1999) Multiobjective evolutionary algorithms: classifications, analyses, and new innovations. *Air Force Inst of Tech Wright-Pattersonafb Oh School of Engineering*, Ann Arbor
- Van Veldhuizen DA, Lamont GB (1999) Multiobjective evolutionary algorithm test suites. In: *Proceedings of the 1999 ACM symposium on Applied computing*, San Antonio, Texas, USA, pp 351–357
- Watson J-P, Woodruff DL (2011) Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Comput Manag Sci* 8(4):355–370
- Weber A (1929) *Theory of the Location of Industries*. University of Chicago Press, Chicago
- Yang W p, Tamg Y (1998) Design optimization of cutting parameters for turning operations based on the Taguchi method. *J Mater Process Technol* 84(1-3):122–129
- Yu X, Zhou Y, Liu X-F (2019) A novel hybrid genetic algorithm for the location routing problem with tight capacity constraints. *Appl Soft Comput* 85:105760
- Zarandi MHF, Hemmati A, Davari S, Turksen IB (2013) Capacitated location-routing problem with time windows under uncertainty. *Knowl-Based Syst* 37:480–489
- Zhalechian M, Tavakkoli-Moghaddam R, Zahiri B, Mohammadi M (2016) Sustainable design of a closed-loop location-routing-inventory supply chain network under mixed uncertainty. *Transp Res E-Log* 89:182–214
- Zhong S, Cheng R, Jiang Y, Wang Z, Larsen A, Nielsen OA (2020) Risk-averse optimization of disaster relief facility location and vehicle routing under stochastic demand. *Transp Res E-Log* 141:102015

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.