Journal of Industrial and Management Optimization

JIMO, 22(12): 6124–6149.
DOI:10.3934/jimo.2026206
Received: 15 April 2026
Revised: 30 September 2026
Accepted: 22 October 2026
Published: 03 November 2026

https://www.aimspress.com/journal/jimo

Research article

Integrated inventory routing and surgical delay for blood products: A bi-objective trade-off approach

Liu Zhao[1,2,3], Nengmin Wang[1,2,3,*], Yinfeng Xu[1] and Meng Zhang[4]

[1] School of Management, Xi'an Jiaotong University, Xi'an Shaanxi 710049, China
[2] ERC for Process Mining of Manufacturing Services in Shaanxi Province, Xi'an Shaanxi 710049, China
[3] The Key Lab of the Ministry of Education for Process Management & Efficiency Engineering, Xi'an Shaanxi 710049, China
[4] School of Economics and Management, Xi'an Technological University, Xi'an Shaanxi 710021, China

* **Correspondence:** Email: wangnm@mail.xjtu.edu.cn.

Abstract: Despite the large and rapidly growing global blood transfusion diagnostics market, it still cannot meet the blood demands worldwide. Medical institutions often resort to postponing or canceling elective surgeries to cope with blood shortages, which not only affects patient treatment but also increases the risk of preventable deaths. In this study, we examined a bi-objective inventory routing problem (IRP) for blood products, taking surgical delays into account. Correspondingly, a bi-objective mixed-integer programming model was developed with two objectives of minimizing the total costs and the total weighted surgical delay time simultaneously. To solve the model, a hybrid algorithm combining the ε-constraint method, the decomposition method, and the variable neighborhood search (VNS) approaches was designed. Specifically, the bi-objective IRP was first transformed into a series of single-objective IRPs based on the framework of the ε-constraint method, and then each single-objective IRP was decomposed into two subproblems of blood product distribution and vehicle routing optimization, which could be solved by the CPLEX solver and the VNS separately. Comparisons with the non-dominated sorting genetic algorithm (NSGA-II) and the augmented ε-constraint method (AUGMECON2) on small- and large-scale instances demonstrated the efficiency and effectiveness of the proposed model and algorithm. Finally, sensitivity analyses revealed that the proposed optimization model provides guidance on setting appropriate hospital inventory capacities under different scenarios, and shows that reducing transportation cost is an effective way to improve the efficiency of the blood supply chain (BSC).

Keywords: bi-objective optimization; IRP; ε-constraint method; decomposition method; VNS
Mathematics Subject Classification: 90B06, 90C59

1. Introduction

Although nearly 120 million units of blood are donated each year, it is still insufficient to meet global blood demands [1]. Many patients cannot obtain safe blood in time, leading to many avoidable deaths. According to a report from the UK, the number of blood transfusion-related deaths in 2024 increased by approximately 55.3% compared to 2023, accounting for about 30.5% of total deaths, with transfusion delay still being the second most common cause. In addition, transfusion delays also caused 6.3% of severe reaction cases [2].

In recent years, due to frequently occurred blood shortages, postponing elective surgeries has become a common practice in medical institutions. In June 2025, the UK's National Health Service issued a warning about a blood shortage to avoid triggering a 'red alert', which means that all non-emergency surgeries and treatments will be forced to be postponed [3]. In July 2025, a blood shortage forced Portugal's largest public hospital to postpone a wave of planned surgeries. Management maintained the operation of trauma and oncology wards by putting all non-life-saving surgeries, from knee replacements to benign tumor removals, on hold [4]. Although the postponement of surgeries due to blood shortages is a common practice in reality, few scholars have taken it into account in issues related to the BSC management.

Most research on IRP for blood products mainly focuses on economic goals, and often quantified blood shortages as a cost. However, the consequences of blood shortages, such as delayed surgeries, untreated patients, and fatalities, are often immeasurable, which differentiate blood product management from that of general perishable goods. Moreover, the limited shelf life of blood products, rendering expired units unusable, further complicates blood product scheduling. On one hand, low blood inventory levels serve to prevent waste caused by expired blood, but they leave healthcare facilities vulnerable to blood shortage risks. On the other hand, hospital blood banks often maintain high inventory levels to ensure timely fulfillment of blood demand, which not only drives up inventory costs but also increases blood product wastage. Therefore, achieving a trade-off between cost-efficiency and the prompt response to patient requirements constitutes a key challenge in the management of BSC.

In this study, we investigate a bi-objective IRP for blood products, considering the postponement of surgeries due to blood shortages. Then, a bi-objective mixed-integer linear programming model is formulated, pursuing the simultaneous minimization of total costs and weighted surgical delay time. To solve the model, a hybrid algorithm based on the ε-constraint and decomposition methods is designed. By converting the bi-objective problem into a set of single-objective problems and subsequently decomposing each of the latter into two subproblems (i.e., blood product distribution and vehicle routing optimization) to be solved individually, this algorithm reduces the complexity of problem-solving while preserving the quality of the obtained solutions. Numerical experiments conducted on small- and large-scale instances confirm that the proposed algorithm exhibits better performance than AUGMECON2 and NSGA-II in terms of solution diversity and convergence. Finally, we conduct two sensitivity analyses to examine the impact of hospital inventory capacity and transportation cost on the two objectives, and several corresponding managerial insights are derived.

The structure of this paper is outlined as follows: In Section 2, we present a concise review of the literature in the field. In Section 3, we elaborate on the description and mathematical formulation of the bi-objective IRP. In Section 4, we provide a detailed illustration of the algorithm proposed in this study, whereas in Section 5, we report the results derived from numerical experiments and sensitivity analyses. Finally, in Section 6, we draw the conclusions and limitations of this research and identify potential directions for future research.

2. Literature review

Given the limited literature on the IRP for blood products, in this section, we mainly review literature in three aspects: (1) IRP-related research for blood products; (2) IRP for perishable products; and (3) solution methods for IRP. Finally, the research gaps are summarized, and the research problem of this paper is presented.

2.1. IRP-related research for blood products

BSC refers to a complete process of collecting blood from donors, then processing, testing, storing, and transporting it to patients for transfusion, including several consecutive stages: collection, production, inventory, and delivery [5]. The IRP for blood products involves two key stages of inventory management and routing planning of the BSC, which are directly related to surgical delays.

Hemmelmayr et al. [6] first proposed a blood delivery scheme for the Austrian Red Cross to reduce travel cost, not allowing blood shortage. Afterward, they considered stochastic blood product usage and introduced four recourse policies to deal with blood shortages and designed a VNS method [7]. Kazemi et al. [8] studied a blood IRP under uncertainty to minimize routing, inventory, and wastage costs and proposed an iterative branch-and-cut algorithm, in which stock-out was not permitted. Jafarkhan and Yaghoubi [9] introduced transshipment and substitution mechanisms into the IRP for red blood cells to decrease the total cost and proposed an adaptive large neighborhood search (ALNS) method. Liu et al. [10] constructed a vendor-managed IRP for blood products to minimize operational costs and developed a decomposition-based algorithm. Vincent et al. [11] investigated a two-stage stochastic vendor-managed IRP for the blood distribution system to minimize transportation and operational costs separately, and proposed a three-stage heuristic algorithm that integrates perturbation heuristics, ALNS, and exact methods. Qing et al. [12] studied a multi-period blood IRP under uncertainties in demand, supply, and travel times, and formulated a robust optimization model to minimize the total cost. They proposed an integrated algorithm based on the *L*-shaped framework combining benders decomposition with a branch-and-price-and-cut to obtain an exact solution.

In recent years, aside from inventory control and route optimization decisions, many scholars have considered more stages of the BSC in their research, such as blood collection, facilities location, and production. In contrast to the IRP for blood products, the BSC literature tends to achieve multiple goals, such as minimizing time-related goals [13–17], reducing the remaining life of blood [18], improving service level [19], increasing network's resilience [15], and reducing the number of workers [15].

2.2. IRP for perishable products

Blood products are a type of perishable product. The perishable IRP refers to the optimization problem where a single supplier is responsible for delivering perishable products to a set of customers within a limited planning period. Shaabani [20] reviewed the perishable IRP across five dimensions: the number of products, the type of product, the type of demand, the number of objective functions, and solution approaches. They focused on the importance of IRP that integrates inventory management with vehicle routing decisions.

Unlike traditional IRPs, the perishable IRP requires special consideration of the product's spoilage characteristics. There are mainly two types of spoilage of perishable products in literature research: One type of product has a fixed lifetime and will become unusable after expiration, and the other type gradually deteriorates over time at a certain rate. Most researchers focus on the first type of perishable product, while for the second type, Azadeh et al. [21] presented an IRP model with transshipment for

a single perishable product, and assumed that products deteriorate at an exponential rate. Timajchi et al. [22] proposed a bi-objective IRP for hazardous and deteriorating pharmaceutical items to minimize the costs and accident loss simultaneously, in which the medicines deteriorate under a constant rate during storage. Dai et al. [23] studied a cyclic IRP for perishable products with stock and price dependent demand, and calculated the loss cost based on the deterioration rate. The sensitivity analysis indicated that an increase in deterioration rate led to higher total costs and required more frequent replenishment to maintain the same inventory level. Cuellar-Usaquén et al. [24] studied an IRP for agri-food with fixed percentage perishability and analyzed the impact of the perishability percentage on the platform's revenue.

Regarding shortages, Diabat et al. [25], Shaabani and Kamalabadi [26], and Azadeh et al. [21] assumed that no stock-out occur in their studies. Hu et al. [27], Alvarez et al. [28], and Zhang et al. [29] indicated that the demands must be fully met through model constraints. Some other scholars use backorder or lost sale to describe stockouts in their research. When stock is insufficient, a backorder refers to unmet demand deferred for later delivery, whereas a lost sale refers to such demand permanently lost. Mirzaei and Seifi [30] presented a perishable IRP model and formulated the cost of lost sale as a linear or an exponential function of the inventory age. Timajchi et al. [22] introduced a parameter to represent the proportion of patients on backorder and explored the impact of this parameter on transshipment strategies in sensitivity analysis, thereby optimizing route selection to reduce accident risk.

2.3. Solution methods for IRP

IRP, as an integrated optimization framework for inventory management, vehicle routing, and delivery scheduling, was first introduced by Bell et al. [31]. Moreover, Coelho et al. [32] provided a comprehensive review of IRPs, summarizing solution methods from two perspectives: exact algorithms and heuristic algorithms.

Most studies on IRP for blood products or perishable items have established single-objective mathematical models. A few scholars adopt branch-and-cut or benders decomposition methods to obtain exact solutions for smaller-scale problems. However, due to the NP-hard nature of the IRP, heuristic algorithms have become the mainstream for solving large-scale practical problems. VNS, ALNS, local search, tabu search, simulated annealing, genetic algorithm, or hybrid metaheuristic algorithms based on them are widely used. In addition, some hybrid heuristic algorithms adopt the idea of a decomposition method to break down complex problems into computationally tractable subproblems to balance solving efficiency and quality [10].

Research on IRP-related BSC management often needs to balance multiple conflicting optimization objectives. General methods for multi-objective optimization problems mainly include: the weighted sum method, the ε-constraint method, and multi-objective evolutionary algorithms [33]. Among them, the ε-constraint method is considered one of the most effective approaches for solving bi-objective optimization problems. This method does not require the unification of the dimensions of objective functions, nor does it rely on the decision maker's prior knowledge, as the weighting method does; moreover, it improves the drawback of the multi-objective evolutionary algorithm, where the quality of the approximate Pareto solutions provided may differ from the optimal solutions to some extent.

2.4. Gap analysis

The literature covered in this paper is summarized in Table 1, which comprises the IRP for blood products and perishable products as well as the BSC involving inventory and routing decisions.

Through a comparative analysis of the reviewed studies in terms of objective functions, treatments of shortages, and methodologies, the research gaps are summarized as follows.

(1) Most BSC studies, including IRP for blood products, primarily pursue cost-oriented objectives for healthcare providers, whereas patient-centered objectives remain largely overlooked. Although the broader BSC literature has embraced multiple objectives, no attempt has been made to jointly consider surgical scheduling alongside blood product delivery. Particularly in the context of blood shortages, deferring elective surgeries is a common practice in clinical settings, and its successful execution is critically dependent on effective BSC management.

(2) Blood shortages pose a persistent challenge in BSC operations. Transshipment and substitution are common strategies in BSC studies to cope with shortages, whereas deferring elective surgeries represents a complementary and underexplored perspective. In perishable IRP, stockouts are typically modeled as backorders or lost sales. Blood products, however, differ from ordinary perishables in that their unmet demand cannot arbitrarily be treated as lost sales; furthermore, the deferral of blood demand is particularly time-sensitive, as it may lead to deterioration of patient conditions and necessitate the reconfiguration of medical resources.

(3) With respect to solution techniques, most studies on IRP for blood or perishable products are designed for single-objective optimization problems and have developed numerous algorithmic approaches accordingly. In contrast, the solution methods for multi-objective IRPs are relatively scarce, rendering this a valuable challenge that deserves dedicated research efforts.

Table 1. Literature review.

Problem	Author	Objective					Treatment of shortages	Method-ology
		Number	Cost	Shortage	Wastage	Others		
IRP for blood products	Hemmelmayr et al. [6]	1	✓				Not permitted	VNS
	Hemmelmayr et al. [7]	1	✓				Recourse policies	VNS
	Jafarkhan and Yaghoubi [9]	1	✓				Transshipment and substitution	ALNS
	Kazemi et al. [8]	1	✓		Cost		Not permitted	BAC
	Liu et al. [10]	1	✓	Cost	Cost			DBA
	Vincent et al. [11]	1	✓	Cost	Cost			HA
	Qing et al. [12]	1	✓	Penalty cost	Cost			HA
	Cheng et al. [34]	2	✓	Penalty cost		Satisfaction	Transshipment and substitution	PDA
BSC involving inventory and routing decisions	Agac et al. [35]	1	✓					GA
	Darestani and Mousazadeh [36]	1	✓				Inventory sharing and substitution	LS and ALNS
	Samani et al. [37]	1	✓	Cost	Cost			FP
	Xu and Szmerekovsky [38]	1	✓	Cost	Cost		Transshipment	SP
	Khalilpourazari and Hashemi Doulabi [13]	2	✓			Transportation time		RO

Continued on next page

Problem	Author	Objective					Treatment of shortages	Method-ology
		Number	Cost	Shortage	Wastage	Others		
BSC involving inventory and routing decisions	Shirazi et al. [16]	2	✓			Blood flow time	Substitution	Multiple heuristics
	Momenitabar et al. [19]	2	✓		Cost	Service level	Transshipment and substitution	FA
	Zahiri et al. [18]	2	✓	Cost		Remaining life of blood		DEA
	Yang et al. [17]	2	✓	Cost		Delivery time	Vertical and horizontal integration	HA
	Farrokhizadeh et al. [39]	2	✓	Unsatisfied demand				ACM and LRA
	Khalilpourazari et al. [14]	3	✓	Unsatisfied demand		Transportation time		NP
	Rekabi et al. [15]	4	✓		Cost	Times, network's resilience and number of workers		ACM
IRP for perishable products	Diabat et al. [25]	1	✓				No	TS
	Shaabani and Kamalabadi [26]	1	✓				No	SAA
	Hu et al. [27]	1	✓				No	HA
	Rohmer et al. [40]	1	✓					ALNS
	Alkaabneh et al. [41]	1				Profit		BD
	Alvarez et al. [28]	1				Profit	No	BAC and HA
	Zhang et al. [29]	1				Profit	No	KS
	Usaquén et al. [24]	1				Reward		ASLP
	Kumar et al. [42]	1	✓		Cost			DEA
	Soysal et al. [43]	1	✓		Cost		Backorder	DA
	Azadeh et al. [21]	1	✓		Cost			GA
	Mirzaei and Seifi [30]	1	✓	Cost			Lost sale	HA
	He et al. [44]	1	✓	Cost			Lost sale	RCR and BD
	Dai et al. [23]	1	✓	Cost			Lost sale	HA
	Nikzad et al. [45]	1	✓	Cost				MP
	Timajchi et al. [22]	2	✓	Cost		Accident loss	Partial backorder	GA
	Niakan and Rahimi [46]	3	✓	Cost	Quantity	Satisfaction and greenhouse gas emissions	Backorder	FA

Notes: BAC: Branch-and-cut; DBA: Decomposition-based algorithm; HA: Hybrid algorithm; PDA: Primal-dual algorithm; GA: Genetic algorithm; LS: Local search; FP: Fuzzy programming; SP: Stochastic programming; RO: Robust optimization; FA: Fuzzy approach; DEA: Differential evolutionary algorithm; ACM: Augmented ε-constraint method; LRA: Lagrangian relaxation approach; NP: Neural-learning process; TS: Tabu search; SAA: Simulated annealing algorithm; BD: Benders

decomposition; KS: Kernel search; ASLP: Adaptive stochastic lookahead policies; DA: Deterministic approximation; RCR: Robust counterpart reformulation; MP: Mathematical programming.

In this study, we propose a bi-objective IRP for blood products, which addresses blood shortages by pursuing not only cost minimization but also the reduction of weighted surgical delay time. To solve the proposed bi-objective mathematical programming model, the ε-constraint method is used to convert the bi-objective IRP into a series of single-objective IRPs. A decomposition-based method is then applied, breaking the single-objective IRP into a blood product distribution subproblem and a set of independent vehicle routing subproblems, which are solved using the CPLEX solver and VNS, respectively. Experimental results demonstrate that the proposed algorithm achieves good diversity and convergence of solutions on small- and large-scale instances. In addition, sensitivity analyses are conducted to investigate the effects of hospital inventory capacity and transportation cost on the total cost and weighted surgical delay time, and corresponding managerial insights are provided.

3. Problem description and formulation

The IRP for blood products can be described by a complete digraph $J=\{V', A\}$, in which $V'=\{0,1,\ldots,n\}$ denotes a set of vertices and $A=\{(i,j): i,j\in V', i\neq j\}$ denotes a set of arcs. The set of vertices V' consists of a blood center numbered 0 and a set of hospitals $V=\{1,2,\ldots,n\}$. Within a time horizon $T=\{1,2,\ldots,|T|\}$, a fleet of homogeneous vehicles $K=\{1,2,\ldots,|K|\}$ with capacity is responsible for transporting blood products from the blood center to hospitals to satisfy the surgical demands. The blood demand for elective surgeries is deterministic and time-varying, which depends on the surgical schedule and the Maximum Surgical Blood Order Schedule (MSBOS) that deterministically quantifies blood units per procedure [47]. The types of surgical procedures vary across hospitals and are denoted as $W=\{1,2,\ldots,|W|\}$, each with distinct requirements for blood products and different priorities for delay. Each hospital has a maximum inventory capacity limit of blood bank. When the blood supply is unable to satisfy the blood demands in the current period, the selective surgeries in hospitals are often postponed and satisfied until the amount of blood products is abundant at later periods. Additionally, because the blood product has a lifespan, expired blood products cannot be used to meet surgical demands and will incur costs.

Based on the above description, a bi-objective IRP for blood products is proposed in this study to optimize inventory control, blood product distribution, and vehicle routing planning simultaneously with two objectives: minimizing the total cost and minimizing the total weighted surgical delay time. The decisions in each period are as follows: 1) which hospitals need to be visited, 2) how much blood products need to be delivered to the hospitals, and 3) how to plan specific delivery routes for the vehicles. The bi-objective IRP for blood products is shown in Figure 1.

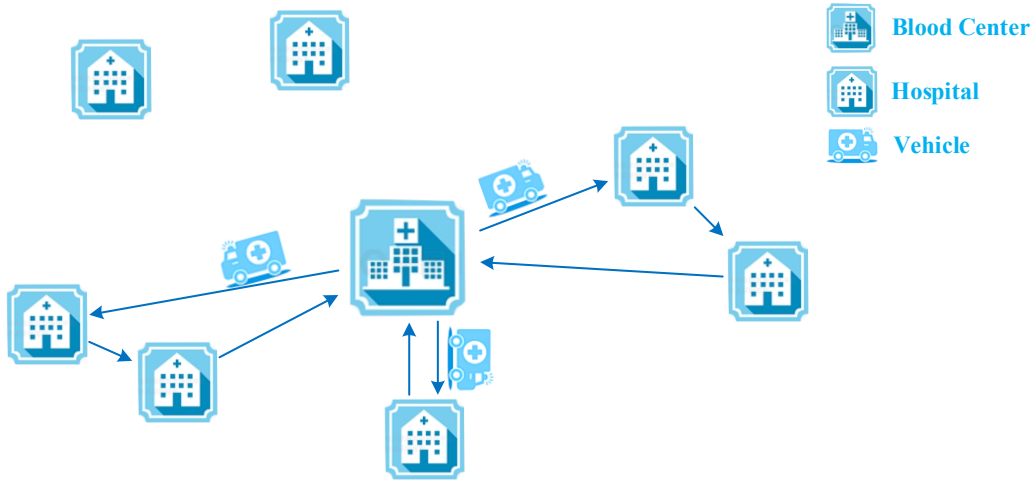


Figure 1. Bi-objective IRP for blood products.

To simplify the problem and focus the optimization model on exploring the trade-off between surgical delays and costs, the following assumptions are made:

• The initial inventory level of each hospital blood bank is 0, which characterizes the extreme conditions arising from blood shortages.

• Blood product transported from the blood center to hospitals is fresh but deteriorates during storage and cannot be used once it exceeds its shelf life.

• Each vehicle starts its journey from the blood center and returns there upon completing its blood transport mission.

• Each hospital can be visited by a vehicle at most once per period, and splitting delivery volumes is not allowed.

To facilitate the formulation of this problem, relevant parameters and variables are defined as follows.

Parameters:

I_i^0 : Initial inventory level of hospital i

S^t : Blood supply to the blood center in period t

U_i : Maximum inventory level of hospital i

D_{iw}^t : The number of type-w surgeries scheduled at hospital i in period t

α_w : Average amount of blood product requirement of type-w surgery

β_w : Priority of type-w surgery

c_{ij} : Distance between vertices i and j

u : Unit transportation cost

h_i : Per-inventory holding cost of hospital i

e : Unit cost of expired blood products

B : Vehicle capacity

L : Shelf life of blood products

M : An arbitrarily large number

θ : An arbitrarily small number

Decision variables:

x_{ij}^{kt} : The binary variable equals one if vertex *i* is visited directly after vertex *j* by vehicle *k* in period *t*; otherwise, it is equal to zero.

y_i^{kt} : The binary variable equals one if vertex *i* is visited by vehicle *k* in period *t*; otherwise, it is equal to zero.

q_i^{kt} : Quantity of blood products delivered to hospital *i* by vehicle *k* during period *t.*

$d_{iw}^{tt'}$: The number of type-*w* surgeries that scheduled at hospital *i* in period *t* and satisfied by blood products delivered in period t' (if $t > t'$, the surgeries scheduled in period *t* are satisfied by the remaining blood inventory at the beginning of period *t*; if $t < t'$, they are satisfied with delay. Notably, $t' = |T|+1$ indicates that the surgeries are canceled).

I_i^t : Inventory level of hospital *i* at the end of period *t.*

E_i^t : Quantity of expired blood products of hospital *i* in period *t.*

Based on the above description and notation, the bi-objective inventory routing problem (BOIRP) for blood products is formulated as follows.

BOIRP

$$\text{Min} \quad f_1 = \sum_{t\in T}\sum_{k\in K}\sum_{i,j\in V'} uc_{ij}x_{ij}^{kt} + \sum_{t\in T}\sum_{i\in V} h_i \max\left(0, I_i^t\right) + eE_i^t \tag{1}$$

$$\text{Min} \quad f_2 = \sum_{i\in V}\sum_{w\in W}\sum_{t,t'\in T} \beta_w d_{iw}^{tt'} \max\left(0, t'-t\right) + M\sum_{i\in V}\sum_{w\in W}\sum_{t\in T}\beta_w d_{iw}^{t,|T|+1} \tag{2}$$

$$\text{s.t.} \quad I_i^t = I_i^{t-1} + \sum_{k\in K} q_i^{kt} - E_i^t - \sum_{w\in W}\alpha_w D_{iw}^t \qquad \forall i\in V, t\in T \tag{3}$$

$$\sum_{i\in V}\sum_{k\in K} q_i^{kt} \le S^t \qquad \forall t\in T \tag{4}$$

$$I_i^t \le U_i \qquad \forall i\in V, t\in T \tag{5}$$

$$d_{iw}^{t't} = 0 \qquad \forall i\in V, 1\le t < |T|-L, t+L < t' \le |T|, w\in W \tag{6}$$

$$\sum_{t'\in T\cup\{|T|+1\}} d_{iw}^{tt'} = D_{iw}^t \qquad \forall i\in V, t\in T, w\in W \tag{7}$$

$$\sum_{k\in K} q_i^{kt} \ge \sum_{w\in W}\sum_{t'\in T}\alpha_w d_{iw}^{t't} \qquad \forall i\in V, t\in T \tag{8}$$

$$E_i^t = \sum_{k\in K} q_i^{kt} - \sum_{w\in W}\sum_{t'\in T}\alpha_w d_{iw}^{t't} \qquad \forall i\in V, t\in T \tag{9}$$

$$\sum_{j\in V'} x_{ij}^{kt} = \sum_{j\in V'} x_{ji}^{kt} = y_i^{kt} \qquad \forall i\in V, t\in T, k\in K \tag{10}$$

$$\sum_{k\in K} y_i^{kt} \le 1 \qquad \forall i\in V, t\in T \tag{11}$$

$$q_i^{kt} \le U_i y_i^{kt} \quad \forall i \in V, t \in T, k \in K \tag{12}$$

$$q_i^{kt} \ge \theta y_i^{kt} \quad \forall i \in V, t \in T, k \in K \tag{13}$$

$$\sum_{i \in V} q_i^{kt} \le B y_0^{kt} \quad \forall t \in T, k \in K \tag{14}$$

$$\sum_{i \in C} \sum_{j \in C} x_{ij}^{kt} \le |C| - 1 \quad \forall C \subseteq V, |C| \ge 2, t \in T, k \in K \tag{15}$$

$$\begin{gathered} x_{ij}^{kt}, y_i^{kt} \in \{0,1\}, \ E_i^t, q_i^{kt} \ge 0, \ d_{iw}^{tt'} \in \mathbb{Z}_{\ge 0} \\ \forall i, j \in V, k \in K, t \in T, t' \in T \cup \{|T|+1\}, w \in W \end{gathered} \tag{16}$$

Objective function (1) minimizes the total costs of transportation, inventory holding, and expired blood products. Objective function (2) minimizes the total weighted surgical postponement time, where the postponement time of canceled surgeries is treated as a very large number. Constraint (3) represents the inventory balance at the hospitals. Constraint (4) ensures that the total quantity of blood products delivered to hospitals in each period does not exceed the available blood supply. Constraint (5) limits the inventory levels to the maximum capacity of the hospitals. Constraint (6) prohibits the use of expired blood products to meet surgical demands, while Constraint (7) implies that surgical demands in hospitals can be delayed or canceled. Constraint (8) specifies that blood products delivered to hospitals are used to meet surgical demands occurring in different periods. Constraint (9) calculates the quantity of expired blood products. Flow conservation is enforced through Constraint (10). Constraint (11) ensures that each hospital is served by a vehicle at most once per period. Constraint (12) specifies that each vehicle delivers blood products only to the hospitals it visits, with the delivery quantity not exceeding the hospital's inventory capacity. Constraint (13) prevents the vehicle from visiting nodes with zero delivery quantity. The vehicle capacity constraint is imposed by Constraint (14), and the sub-tour elimination is enforced by Constraint (15). Finally, the ranges and domains of the decision variables are defined in Constraint (16).

4. Solution method

Due to the complexity of the proposed bi-objective IRP model, it is time-consuming to obtain exact Pareto fronts. Therefore, we design an ε-constraint and decomposition-based algorithm to obtain approximate Pareto fronts. First, the algorithmic framework of the ε-constraint method is presented, which converts the bi-objective IRP into a series of single-objective IRPs. Then, the decomposition of a single-objective IRP is elaborated, and then the algorithm is summarized.

4.1. ε-constraint method

The ε-constraint method was first proposed by Haimes et al. [48], which is widely applied to solve bi-objective optimization problems. This method facilitates the acquisition of Pareto fronts by transforming a bi-objective problem into a sequence of single-objective problems and iteratively solving them to obtain a set of Pareto solutions.

For the BOIRP in this study, objective function (2) is transformed into a constraint subject to ε limitations expressed as constraint (17), yielding a single-objective inventory routing problem

(SOIRP).

SOIRP

$$\text{Min} \quad f_1 = \sum_{t\in T}\sum_{k\in K}\sum_{i,j\in V'} uc_{ij}x_{ij}^{kt} + \sum_{t\in T}\sum_{i\in V} h_i \max\left(0, I_i^t\right) + eE_i^t$$

$$\text{s.t.} \sum_{i\in V}\sum_{w\in W}\sum_{t,t'\in T} \beta_w d_{iw}^{tt'} \max\left(0, t'-t\right) + M\sum_{i\in V}\sum_{w\in W}\sum_{t\in T} \beta_w d_{iw}^{t,|T|+1} \le \varepsilon \tag{17}$$

Constraints (3)–(16)

For the original BOIRP, the ideal point $\left(f_1^{\min}, f_2^{\min}\right)$ and nadir point $\left(f_1^{\max}, f_2^{\max}\right)$ can be obtained by solving the following models, where X denotes the set of feasible solutions.

$$f_1^{\min} = \min f_1(x) \quad \text{s.t.} \quad x \in X \tag{18}$$

$$f_2^{\min} = \min f_2(x) \quad \text{s.t.} \quad x \in X \tag{19}$$

$$f_1^{\max} = \min f_1(x) \quad \text{s.t.} \quad f_2 = f_2^{\min}, \quad x \in X \tag{20}$$

$$f_2^{\max} = \min f_2(x) \quad \text{s.t.} \quad f_1 = f_1^{\min}, \quad x \in X \tag{21}$$

Accordingly, the ε-constraint method can be presented in Algorithm 1.

Algorithm 1 ε-constraint method

1. Determine the range of ε, that is $\left[f_2^{\min}, f_2^{\max}\right]$;
2. Create a set F and add point $\left(f_1^{\min}, f_2^{\max}\right)$ to it;
3. For $\varepsilon = f_2^{\max} - \sigma$ to $f_2^{\min}$ // σ is a small positive number
4. Solve the SOIRP with the current value of ε to generate an optimal solution $\hat{x}$;
5. Add point $\left(f_1(\hat{x}), f_2(\hat{x})\right)$ to set F;
6. Let $\varepsilon = f_2(\hat{x}) - \sigma$;
7. End for
8. Delete non-dominated solutions from set F;
9. Output set F, which constitutes the Pareto front of the BOIRP.

4.2. Decomposition method for SOIRP

When iteratively solving the SOIRP in Algorithm 1, finding solutions within a reasonable time frame is challenging due to the complexity inherent in this NP-hard model. Therefore, it is imperative to decompose the SOIRP into tractable subproblems and solve them individually to enhance the solution efficiency of the problem. In this study, the SOIRP can be divided into a blood product distribution subproblem and a series of independent vehicle routing subproblems, which are elaborated as follows.

4.2.1. Blood product distribution subproblem (BPDP)

The BPDP determines the quantity of blood products delivered to hospitals regardless of detailed routes, which is derived from the SOIRP without route-related variables and constraints. In the BPDP, decision variables y_i^{kt} and q_i^{kt} of the SOIRP are replaced by Y_i^t and Q_i^t with similar meanings, by assuming an aggregated vehicle with an accumulated capacity responsible for the blood product transportation task. Accordingly, the transportation cost can be estimated by assuming direct delivery of blood products from the blood center to the hospitals. Subsequently, the BPDP can be established as follows.

BPDP

$$\text{Min} \quad f_{ec} = \sum_{t\in T}\sum_{i\in V} u2c_{i0}Y_i^t + \sum_{t\in T}\sum_{i\in V} h_i \max\left(0, I_i^t\right) + eE_i^t \tag{22}$$

$$\text{s.t.} \quad I_i^t = I_i^{t-1} + Q_i^t - E_i^t - \sum_{w\in W} \alpha_w D_{iw}^t \qquad \forall i \in V, t \in T \tag{23}$$

$$\sum_{i\in V} Q_i^t \le S^t \qquad \forall t \in T \tag{24}$$

$$Q_i^t \ge \sum_{w\in W}\sum_{t'\in T} \alpha_w d_{iw}^{t't} \qquad \forall i \in V, t \in T \tag{25}$$

$$E_i^t = Q_i^t - \sum_{w\in W}\sum_{t'\in T} \alpha_w d_{iw}^{t't} \qquad \forall i \in V, t \in T \tag{26}$$

$$Q_i^t \le U_i Y_i^t \qquad \forall i \in V, t \in T \tag{27}$$

$$Q_i^t \ge \theta Y_i^t \qquad \forall i \in V, t \in T \tag{28}$$

$$\sum_{i\in V} Q_i^t \le B\left|K\right| \qquad \forall t \in T \tag{29}$$

$$Y_i^t \in \{0,1\}, \quad E_i^t, Q_i^t \ge 0, \quad d_{iw}^{tt'} \in \mathbb{Z}_{\ge 0} \qquad \forall i \in V, t \in T, t' \in T \cup \left\{\left|T\right|+1\right\}, w \in W \tag{30}$$

Constraints (5)–(7) and (17)

Objective function (22) minimizes the approximate total costs, including estimated transportation cost, inventory cost, and expiration cost. Constraints (23)–(29) convey the same meanings as Constraints (3), (4), (8), (9), (12), (13), and (14), respectively. Constraint (30) defines the ranges of the decision variables.

As the BPDP does not involve the subtour elimination constraint presented in the SOIRP, it can be solved more quickly using the CPLEX solver. However, a solution to the BPDP that satisfies Constraint (29) may not satisfy Constraint (14), which implies that no feasible solution to the SOIRP can be found. This is because the presumed vehicle performing the transportation assignment complies with the aggregated capacity limit but may violate the capacity requirement for individual vehicles. To address this issue, we find that verifying the feasibility of the BPDP solution with respect to the individual vehicle capacity constraint can be modeled as a classic bin packing problem, which can be

effectively solved by the First Fit Descending heuristic described in Algorithm 2. The idea of Algorithm 2 is: Iteratively assign the hospital with the largest blood product delivery quantity to the vehicle with the largest remaining capacity until all hospitals are allocated; if this cannot be achieved, it indicates that the BPDP solution fails to satisfy the individual vehicle capacity constraint.

Algorithm 2 Bin-packing based feasibility check for individual vehicle capacity

1. Input the set of visited hospitals $H^t = \left\{ i \in V \middle| Y_i^t = 1 \right\}$ in period t and corresponding delivery quantity Q_i^t obtained by solving the BPDP.
2. Initialize remaining capacity R_k of each vehicle in set K to B, i.e., $R_k = B$.
3. Set $tag^t \leftarrow 0$; //tag^t is a binary variable and indicates whether all visited hospitals can be assigned to vehicles in period t or not.
4. Sort hospitals in set H^t in descending order of Q_i^t to obtain set $\hat{H}^t$.
5. For $i = 1$ to $\left|\hat{H}^t\right|$.
6. Find the vehicle $\tilde{k}$ with the maximum remaining capacity $R_{\tilde{k}}$.
7. If $\left(Q_i^t \le R_{\tilde{k}}\right)$
8. Assign hospital i to vehicle $\tilde{k}$.
9. Update $R_{\tilde{k}} \leftarrow R_{\tilde{k}} - Q_i^t$.
10. Else
11. $tag^t \leftarrow 1$; //Unallocated hospitals exist in period t.
12. End if
13. End for
14. Output tag^t.

The solution obtained from the BPDP is feasible for the SOIRP only if the visited hospitals can be assigned to the set of vehicles across all periods. If not, we can adjust the delivery quantity of blood products by adding Constraint (31) to the BPDP and solve it iteratively until a feasible solution is found.

$$\sum_{i=|K|}^{\left|\hat{H}^t\right|} Q_i^t \le B \quad \forall t \in T \middle| tag^t = 1 \tag{31}$$

Constraint (31) restricts the last $\left|\hat{H}^t\right| - |K| + 1$ hospitals in set $\hat{H}^t$ with smaller delivery quantities to be served by a single vehicle. Thus, the first $|K| - 1$ hospitals with larger delivery quantities can each be served by a separate vehicle, ensuring the solution satisfies the individual vehicle capacity constraint.

4.2.2. Independent vehicle routing subproblem (IVRP)

Based on the set of visited hospitals H^t and corresponding delivery quantity Q_i^t obtained from solving the BPDP, a set of IVRPs can be formulated to determine the specific delivery route in each

period, as shown below. The decision variables X_{ij}^k and Z_i^k are derived from x_{ij}^{kt} and y_i^{kt} in the SOIRP, with similar meanings, respectively.

IVRP

$$\text{Min} \quad f_{rc} = \sum_{k \in K} \sum_{i,j \in H^t \cup \{0\}} uc_{ij} X_{ij}^k \tag{32}$$

$$\text{s.t.} \quad \sum_{j \in H^t \cup \{0\}} X_{ij}^k = \sum_{j \in H^t \cup \{0\}} X_{ji}^k = Z_i^k \quad \forall i \in H^t, k \in K \tag{33}$$

$$\sum_{i \in H^t} Q_i^t Z_i^k \le B \quad \forall k \in K \tag{34}$$

$$\sum_{i \in C} \sum_{j \in C} X_{ij}^k \le |C| - 1 \quad \forall C \subseteq H^t, |C| \ge 2, k \in K \tag{35}$$

$$\sum_{k \in K} Z_i^k = 1 \quad \forall i \in H^t \tag{36}$$

$$X_{ij}^k, Z_i^k \in \{0,1\} \quad \forall i, j \in H^t \cup \{0\}, k \in K \tag{37}$$

Objective function (32) minimizes the exact transportation cost. Constraints (33)–(35) hold the similar meanings to constrains (10), (14), and (15), respectively. Constraint (36) ensures that all hospitals in set H^t must be visited by a vehicle. The binary decision variables are defined in Constraint (37).

As one of the methods extensively used to solve vehicle routing problems, VNS is introduced in this study to solve the IVRP. The VNS enlarges the search scope by changing neighborhood structures from the simplest to more complex ones, which is powerful for exploitation. The four popular neighborhood structures used in this study are listed below.

Insert within a path: A hospital is relocated on the same path.

Swap within a path: Two hospitals on a path exchange their positions.

Insert between two paths: A hospital is moved from one path to another path.

Swap between two paths: Two hospitals on different paths exchange their positions.

The VNS employed to solve the IVRP is described in Algorithm 3, where the initial solution is derived from checking the feasibility of the BPDP solution via Algorithm 2.

Algorithm 3 VNS for vehicle routing problem

1. Input a set of neighborhood structures G_s $\left(s = 1, \ldots, s_{\max}\right)$.
2. Generate an initial solution x via Algorithm 2.
3. Randomly reinsert 50% of the hospitals in the initial solution x to form a new solution x'.
4. For $s = 1$ to $s_{\max}$
5. Generate a local optimum solution x'' from the neighborhood $G_s\left(x'\right)$.
6. If $\left(f_{rc}\left(x''\right) < f_{rc}\left(x'\right)\right)$ then
7. $x' \leftarrow x''$; $s \leftarrow 1$.
8. Else

9. $s \leftarrow s+1$.
10. End if
11. End for
12. Output the best solution x^* and $f_{rc}\left(x^*\right)$.

4.3. Decomposition-based heuristic for SOIRP

Based on the decomposition method detailed in Section 4.2, a heuristic combining the CPLEX solver and VNS is designed to solve the SOIRP. The BPDP is first solved by the CPLEX solver and then checked via Algorithm 2, producing a set of visited hospitals and the corresponding delivery quantities of blood products. Furthermore, the total weighted surgical delay time, inventory cost, and expiration cost can be computed. Then, using the initial route derived from Algorithm 2, VNS is applied to address the IVRPs to determine the delivery route and the exact transportation cost. As a result, the total cost can be calculated.

Based on the above, the decomposition-based heuristic algorithm for solving the SOIRP is described in Algorithm 4.

Algorithm 4 Decomposition-based heuristic for SOIRP

1. Solve the BPDP with the current value of ε using the CPLEX solver to generate a set of visited hospitals and corresponding delivery quantities.
2. Check the BPDP solution by Algorithm 2 until it is feasible for individual vehicle capacity;
3. Calculate f_2, IC, and EC based on the checked BPDP solution; // IC denotes the inventory cost and EC denotes the expiration cost.
4. Solve the IVRP by Algorithm 3 to determine the delivery route and f_{rc}.
5. Calculate $f_1 \leftarrow f_{rc}+IC+EC$.
6. Output f_1 and f_2.

4.4. Complete algorithm for BOIRP

According to the framework of the ε-constraint method and decomposition mechanism elaborated in previous sections, the complete heuristic algorithm for solving the BOIRP is developed in this section. Like the ε-constraint method, the first step of the complete algorithm is to acquire the range of ε before the iteration starts. In Algorithm 1, the range of ε presented as $\left[f_2^{\min}, f_2^{\max}\right]$ is determined by solving three mixed-integer programming models exactly. Specifically, $f_2^{\min}$ is derived from Formula (19), while $f_2^{\max}$ is obtained by solving Formula (21) before $f_1^{\min}$ is determined from Formula (18).

Since the objective function of Formula (19) is not related to delivery routes, we approximate the lower bound of ε by assuming an aggregated vehicle to deliver blood products, as in the BPDP model constructed in Section 4.2.1. The approximate lower bound of ε is denoted by $\hat{f}_2^{\min}$ and is obtained by solving the following Model 1. This approximation reduces computational time by removing vehicle routing decisions.

Model 1

$$\text{Min} \quad f_2 = \sum_{i\in V}\sum_{w\in W}\sum_{t,t'\in T}\beta_w d_{iw}^{tt'}\max\left(0,t'-t\right)+M\sum_{i\in V}\sum_{w\in W}\sum_{t\in T}\beta_w d_{iw}^{t,|T|+1}$$

$$\text{s.t.} \quad Q_i^t \le U_i \qquad \forall i\in V, t\in T \tag{38}$$

$$E_i^t, Q_i^t \ge 0, \quad d_{iw}^{tt'} \in \mathbb{Z}_{\ge 0} \qquad \forall i\in V, t\in T, t'\in T\cup\left\{|T|+1\right\}, w\in W \tag{39}$$

Constraints (5)–(7), (23)–(26) and (29)

Constraint (38) serves the same purpose as Constraint (27), and Constraint (39) specifies the ranges of the decision variables.

To obtain the upper bound of ε, we consider the worst case: no blood products are delivered to hospitals by vehicles in any period, and all surgeries are cancelled. In this case, the total cost, including transportation, inventory, and blood expiration costs, is zero, i.e., $\hat{f}_1^{\min}=0$. The total weighted surgical delay time can be calculated by $\hat{f}_2^{\max}=M\sum_{i\in V}\sum_{w\in W}\sum_{t\in T}\beta_w D_{iw}^t$.

Thus, the approximate range of ε in the complete heuristic, i.e., $\left[\hat{f}_2^{\min},\hat{f}_2^{\max}\right]$, is determined. Then, the SOIRP is decomposed and solved by Algorithm 4 to acquire a set of approximate Pareto solutions. The complete algorithm for solving the BOIRP based on the ε-constraint and decomposition methods is illustrated as Algorithm 5, with its flowchart shown in Figure 2.

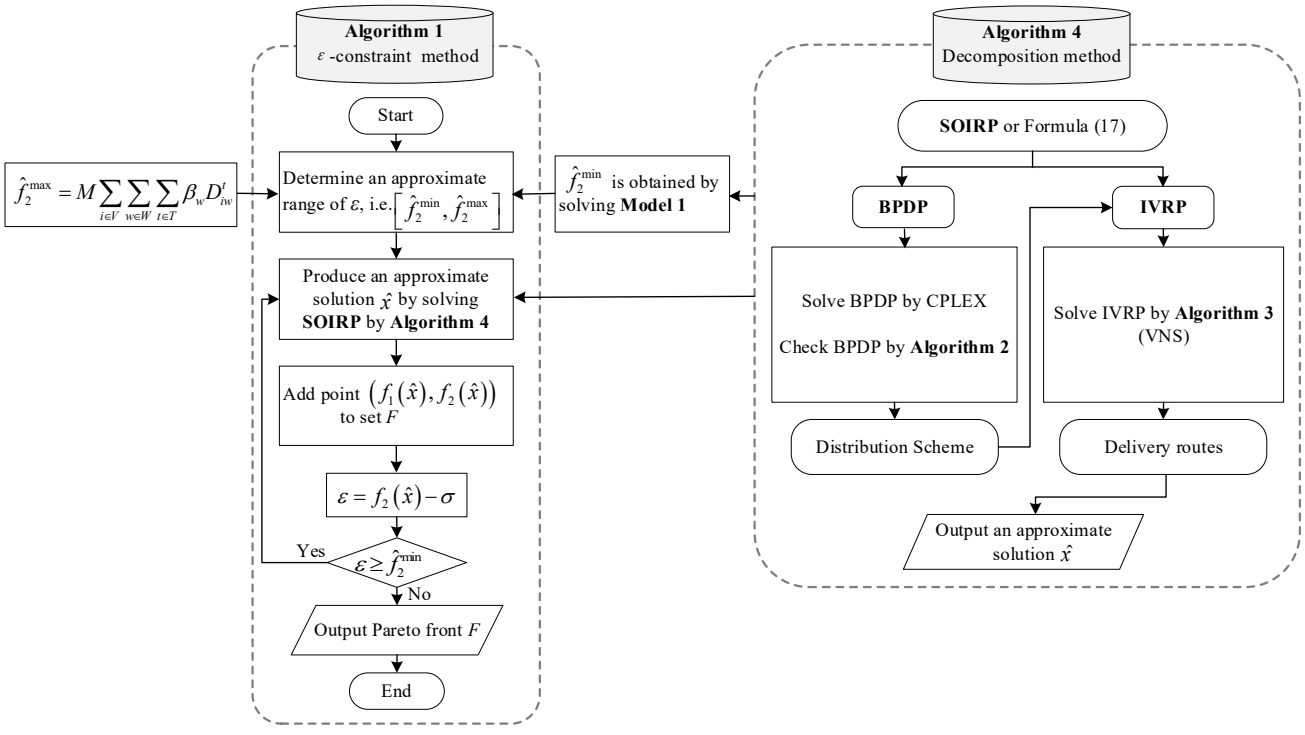


Figure 2. Flowchart of the complete algorithm.

5. Computational experiments

To showcase the contributions of this study, the following experiments are conducted. First, the acquisition of the test data is introduced in Section 5.1. Then, in Sections 5.2 and 5.3, we present

comparative experiments on small- and large-scale instances, respectively, where the proposed algorithm (Algorithm 5) is compared with AUGMECON2 and NSGA-II. Similarly, decomposition methods are also employed in these two algorithms. Finally, two sensitivity analyses are performed in Section 5.4, and managerial insights are derived. All algorithms are coded in Java with CPLEX 12.6.3, and the numerical experiments are run on a personal computer equipped with an Intel i5-13500H CPU (2.60 GHz) and 16.0 GB of RAM.

5.1. Test instances

Since it is difficult to obtain real-world data or existing benchmarks suitable for our research, the experimental data in this paper are drawn from the literature or generated randomly. Vincent et al. [11] investigated an IRP for platelet based on historical data in Surabaya, Indonesia, which has one blood center and 30 hospitals. We derive the coordinates of the blood center and hospitals from the map and calculate the distances between them. The blood supply and demand are randomly generated based on the information provided in this literature. Here, we set three surgery types ($W = \{1,2,3\}$), requiring 1, 2, and 3 units of platelet and with priorities of 1, 2, and 3, respectively. Other parameter values required by the mathematical model are listed in Table 2.

Table 2. Parameters of the optimization model.

Parameter	Value	Resource	Parameter	Value	Resource
S^t	[100, 300]	Generate randomly	U_i	10	Literature [11]
D_{iw}^t	[0, 5]	Generate randomly	u	645	Literature [11]
c_{ij}		Calculate based on Literature [11]	h_i	680	Literature [11]
α_w	{1, 2, 3}		e	1400	Literature [11]
β_w	{1, 2, 3}		B	50	Literature [11]
I_i^0	0		L	5	Literature [11]
θ	0.001		M	100	

Since the values of Objective function (2) are integers, it suffices to obtain the Pareto fronts of the ε-constraint method and Algorithm 5 by setting σ to 0.5. The parameters of AUGMECON2 and NSGA_II shown in Table 3 are tuned, and the processes are provided in Appendix A.

Table 3. Parameters of AUGMECON2 and NSGA_II.

AUGMECON2		NSGA_II	
Parameter	Value	Parameter	Value
Number of grid points	3000	Population size	100
		Crossover probability	0.9
		Mutation probability	0.5
		Number of generations	30

5.2. Results for small-scale instances

In this section, we compare the approximate Pareto fronts yielded by the proposed algorithm (Algorithm 5), AUGMECON2, and NSGA-II with the exact Pareto fronts obtained by the ε-constraint

method (Algorithm 1) using CPLEX. Specifically, the comparison considers the number of solutions, the number of optimal solutions, the distance to the exact Pareto front, and the computational time. Before computing the distance between Pareto fronts, the objective values of all solutions are normalized to [0,1] using the ideal and nadir points of the exact Pareto front.

Fifteen instances across 3 scales are tested, with each scale characterized by 3 periods, 3 vehicles, and a different number of hospitals (3, 4, or 5). Each scale has 5 instances, and each instance is run 3 times to obtain the average value. Since the planning horizon in these instances is shorter than the lifetime of platelets, expiration is not considered. The ε-constraint method with CPLEX fails to obtain the complete exact Pareto front within 4 hours for larger-scale instances with more periods or hospitals.

The experimental results based on small-scale instances are shown in Table 4. The instance name 'xTyN_z' contains three numbers, x, y, and z, denoting the number of periods, hospitals, and instance index, respectively.

Table 4. Experimental results for small-scale instances.

Instance	Number of solutions				Number of optimal solutions			Distance to the exact Pareto front			Computation time (s)			
	A1	A5	AA	AN	A5	AA	AN	A5	AA	AN	A1	A5	AA	AN
3T3N_1	16	19	16	40	3	1	0	0.05	0.04	0.14	10.66	6.5	5.11	28.75
3T3N_2	10	15	14	38	3	2	0	0.05	0.06	0.2	9.18	2.97	3.03	63.75
3T3N_3	15	14	14	35	5	1	0	0.05	0.05	0.2	10.38	3.02	3.43	42.56
3T3N_4	18	20	16	51	2	1	0	0.05	0.05	0.15	28.21	4.56	3.65	79.43
3T3N_5	53	23	15	43	3	1	0	0.05	0.05	0.14	20.12	4.32	4.21	46.43
3T4N_1	16	25	17	43	3	1	0	0.1	0.08	0.29	25.22	5.12	5.11	107.17
3T4N_2	16	22	18	48	3	2	0	0.06	0.08	0.26	79.97	4.56	6.87	156.76
3T4N_3	32	27	19	40	3	2	0	0.1	0.11	0.27	28.45	4.72	4.27	163.14
3T4N_4	18	29	17	47	2	1	0	0.08	0.09	0.27	39.65	8.54	6.98	102.34
3T4N_5	21	20	18	47	2	2	0	0.06	0.07	0.25	62.21	7.61	6.99	95.01
3T5N_1	41	28	24	34	2	1	0	0.08	0.09	0.26	84.56	14.24	9.14	91.33
3T5N_2	28	26	23	38	0	1	0	0.2	0.1	0.31	56.35	12.45	7.7	57.92
3T5N_3	28	23	31	35	0	1	0	0.09	0.08	0.29	42.38	13.43	6.87	140.54
3T5N_4	32	36	21	34	2	1	0	0.08	0.07	0.26	123.13	14.65	6.97	59.40
3T5N_5	21	27	24	45	2	1	0	0.09	0.08	0.32	97.32	10.30	8.97	147.78

Notes: A1: Algorithm 1; A5: Algorithm 5; AA: AUGMECON2; AN: NSGA_II.

From Table 4, it can be observed that Algorithm 5 and AUGMECON2 yield a number of Pareto solutions closer to that of the exact Pareto front. NSGA-II obtains far more Pareto solutions than the other algorithms, and this number is almost unaffected by the instance scale. In terms of the number of optimal solutions contained, Algorithm 5 outperforms AUGMECON2 and NSGA-II, while NSGA-II finds almost no optimal solutions. Regarding the distance to the exact Pareto front, Algorithm 5 and AUGMECON2 are evidently closer to the exact Pareto front, yet no clear winner can be distinguished between these two algorithms. In terms of computational time, AUGMECON2 is the fastest, Algorithm 5 is also faster than Algorithm 1, while NSGA-II performs the worst.

In summary, for small-scale instances, Algorithm 5 can obtain Pareto solutions closer to the exact Pareto front than AUGMECON2 and NSGA-II. The distance between the Pareto front obtained by

AUGMECON2 and the exact Pareto front is comparable to that of Algorithm 5, and AUGMECON2 has an advantage in computational time. NSGA-II, however, performs the worst in terms of solution accuracy and computational time, although it is able to obtain the largest number of Pareto solutions.

5.3. Results for large-scale instances

For large-scale instances, we test 15 instances across 5 problem scales with 6, 7, 8, or 9 hospitals, 3 or 6 periods, and 3 vehicles, to compare the performance of the proposed algorithm (Algorithm 5) with that of AUGMECON2 and NSGA-II. Larger-scale instances cannot yield the complete Pareto front within 4 hours by Algorithm 5. Each problem scale contains 3 instances, each of which is solved 3 times to obtain the average value. If the planning horizon is shorter than the platelet lifetime, expiration is not taken into account.

The performance of all algorithms is evaluated using the following four metrics. Since the ranges of objective values are too large, all Pareto fronts need to be normalized to [0,1] before calculating metric values, using the ideal and nadir points of the combined approximate Pareto fronts of the three algorithms.

(1) Number of Pareto solutions: The higher the number of solutions, the better the algorithm's performance. This metric measures how well the algorithm covers the Pareto front.

(2) Diversification matrix [49]: This measure serves as a diversity indicator, evaluating the spread of solutions within the Pareto front. Higher diversification matrix values suggest better diversification capability. It is calculated as follows:

$$\sqrt{\left(\hat{f}_1^{\max}-\hat{f}_1^{\min}\right)^2+\left(\hat{f}_2^{\max}-\hat{f}_2^{\min}\right)^2} \tag{40}$$

(3) Mean ideal distance [49]: This metric represents the convergence rate of the Pareto front, specifically measuring the proximity of solutions to the ideal point. Smaller values of this metric indicate a superior performance. In this study, the ideal point is defined as (0, 0), which can be calculated as follows:

$$\frac{\sum_{m\in F}\sqrt{\left(f_1^m\right)^2+\left(f_2^m\right)^2}}{|F|} \tag{41}$$

where f_1^m and f_2^m denote the values of the two objective functions of the mth solution on the Pareto front F.

(4) Computational time.

Table 5 compares the five metric values among the three algorithms. The instance names follow the same pattern as in Table 4.

Table 5. Experimental results for large-scale instances.

Instance	Number of Pareto solutions			Diversification matrix			Mean ideal distance			Computation time (s)		
	A5	AA	AN	A5	AA	AN	A5	AA	AN	A5	AA	AN
3T6N_1	39	26	49	1.23	1.25	0.76	0.64	0.67	0.74	22.34	11.87	157.33
3T6N_2	35	27	39	1.26	1.21	0.83	0.62	0.65	0.77	14.32	10.96	154.73
3T6N_3	35	28	52	1.29	1.31	0.67	0.62	0.64	0.76	15.86	8.33	148.89
3T7N_1	48	27	47	1.29	1.27	0.77	0.64	0.68	0.76	53.97	14.75	162.25
3T7N_2	52	29	44	1.27	1.25	0.66	0.63	0.66	0.79	113.25	19.17	136.47
3T7N_3	38	28	37	1.27	1.24	0.8	0.66	0.68	0.77	34.63	13.72	134.26
3T8N_1	42	31	38	1.29	1.22	0.62	0.66	0.67	0.8	77.48	15.62	167.38
3T8N_2	42	28	42	1.24	1.27	0.64	0.63	0.66	0.82	59.02	42.98	106.99
3T8N_3	47	32	35	1.31	1.19	0.62	0.63	0.65	0.78	515.89	17.18	168.89
3T9N_1	57	29	33	1.28	1.26	0.59	0.65	0.69	0.81	71.79	13.42	92.95
3T9N_2	57	34	40	1.26	1.25	0.7	0.72	0.74	0.79	77.01	13.43	86.83
3T9N_3	47	30	44	1.26	1.29	0.72	0.69	0.72	0.78	320.01	15.58	144.54
6T3N_1	44	24	45	1.31	1.27	0.93	0.64	0.67	0.76	57.06	21.34	72.69
6T3N_2	34	27	45	1.31	1.29	0.77	0.65	0.65	0.78	49.95	19.83	149.98
6T3N_3	41	24	43	1.31	1.3	0.91	0.65	0.65	0.77	109.37	28.57	120.53

Notes: A5: Algorithm 5; AA: AUGMECON2; AN: NSGA_II.

Table 5 indicates that Algorithm 5 and NSGA-II produce comparable numbers of Pareto solutions, both markedly exceeding that of AUGMECON2. The diversification matrix values of Algorithm 5 and AUGMECON2 are close to each other, yet both are considerably larger than that of NSGA-II. Regarding the mean ideal distance, NSGA-II performs slightly worse than the other two algorithms, while Algorithm 5 is marginally outperforming AUGMECON2 across all instances. In terms of computational time, AUGMECON2 is the fastest, whereas NSGA-II is the slowest.

Given the above, for large-scale problems, Algorithm 5 and AUGMECON2 outperform NSGA-II in terms of solution diversity, convergence, and computational time, even though NSGA-II yields a larger number of Pareto solutions. Algorithm 5 and AUGMECON2 are comparable in diversity, whereas Algorithm 5 possibly exhibits slightly better convergence, as reflected in its marginally higher mean ideal distance values. Moreover, AUGMECON2 requires the shortest computational time, and this advantage becomes more pronounced as the instance scale increases, but it obtains the fewest Pareto solutions. Although Algorithm 5 takes longer than AUGMECON2, it is able to generate more Pareto solutions.

5.4. Sensitivity analyses

In this section, we discuss the sensitivity analysis results with respect to hospital inventory capacity and transportation cost, based on instance '3T6N_2'. Other parameters such as vehicle capacity, hospital inventory cost, and expiration cost are also tested, but the results are insensitive to these parameters. All experiments are carried out using Algorithm 5.

5.4.1. Sensitivity to hospital inventory capacity

To analyze the sensitivity to hospital inventory capacity, the parameter is set to 10 (the base instance), 20, 30, 40, and 50. The experimental results are illustrated in Figure 3.

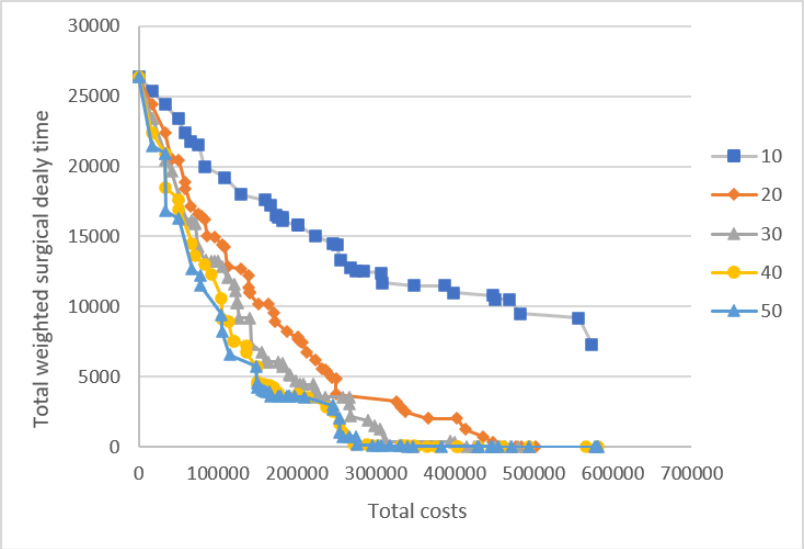


Figure 3. Sensitivity to hospital inventory capacity.

Figure 3 illustrates the trade-off between total cost and total weighted surgical delay time under different hospital inventory capacities. Overall, the larger the inventory capacity, the more the curve shifts toward the lower-left region. Specifically, the curve for capacity 10 lies in the uppermost-right position and performs the worst, whereas the curves for capacities 20–50 shift markedly toward the lower-left, indicating significant improvement, although the marginal benefit gradually diminishes. In addition, the curves for capacities 20–50 can drive the surgical delay time close to zero at progressively lower cost levels, whereas capacity 10 results in a total weighted surgical delay time of approximately 9,000 even at the highest cost level (about 600,000). This indicates that an excessively small inventory capacity (the base instance) makes it difficult to eliminate surgical delays even with substantial investment, and that enlarging inventory capacity can effectively reduce cost and surgical delays, while the marginal benefit becomes very limited once the inventory capacity exceeds 40.

Therefore, if the budget is limited, capacities 20–30 can be selected to achieve satisfactory surgical delay control at moderate cost; if low delay is pursued, capacity 40 can reduce the delay to nearly zero at a cost of about 300,000; capacity 50 is not recommended, since it is almost indistinguishable from capacity 40 and thus offers poor cost-effectiveness. Managers should determine the investment scale near the inflection point of the curve (around a cost of 300,000), where cost-effectiveness is maximized.

5.4.2. Sensitivity to transportation cost

To analyze the sensitivity with respect to transportation cost, this parameter is set to −50%, −25%, 0%, 25%, and 50% of its base value. The corresponding results are displayed in Figure 4.

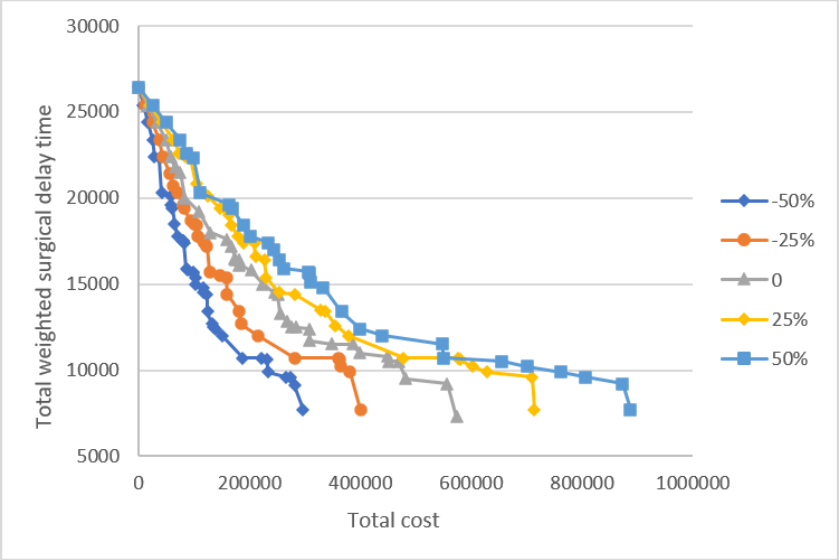


Figure 4. Sensitivity to transportation cost.

Figure 4 illustrates the trade-off between total cost and total weighted surgical delay time under different transportation cost levels (−50%, −25%, 0%, 25%, and 50%). The figure exhibits a weakly monotonic relationship: As the transportation cost decreases, the curve shifts toward the lower-left and achieves better performance. Notably, the final minimum delay levels of all curves are similar, converging to approximately 7,000–8,000; however, the lower the transportation cost, the less total cost is required to reach this level. This indicates that the main benefit of reducing transportation cost lies in achieving the same delay level with less total investment, rather than in pushing the delay below its lower bound.

Therefore, reducing transportation cost is an effective way to improve the efficiency of the BSC, but its primary value lies in saving total cost rather than significantly reducing the surgical delay time. Managers should prioritize optimizing the transportation process and determine the investment scale near the knee of the curve, where the marginal benefit is the highest.

6. Conclusion

In this study, we present a bi-objective IRP for blood products, addressing total cost and weighted surgical delay time in the context of blood shortages. The proposed model integrates inventory control and route optimization while considering the real-world challenge of elective surgery postponements due to insufficient blood supply. The solution method combines the ε-constraint method with a decomposition strategy, dividing the IRP into a blood product distribution subproblem and independent vehicle routing subproblems. By conducting numerical experiments on small-scale and large-scale instances, the proposed algorithm achieves the best balance between solution diversity and convergence, AUGMECON2 is the most computationally efficient but yields the fewest Pareto solutions, and NSGA-II performs the worst despite producing the largest number of Pareto solutions. The sensitivity analyses concerning hospital inventory capacity and transportation cost yield several management sights. Managerially, when the budget is limited, a hospital inventory capacity of 20–30 is cost-effective for satisfactory surgical delay control, whereas a capacity of 40 is preferable for pursuing near-zero delay at a cost of about 300,000, and more capacity offers little additional benefit. Reducing transportation cost mainly saves total cost rather than lowering the surgical delay, and the investment scale should be set near the knee of the curve.

This research advances the field by proposing a novel bi-objective IRP approach for blood product scheduling, where cost and surgical delay time are considered as dual objectives. However, this study has several limitations, including deterministic demand and supply, zero initial inventory, and no differentiation of surgeries by urgency level. Researchers could therefore extend the model to stochastic problems, incorporate more refined emergency prioritization, and adapt the framework to other healthcare contexts, such as multi-product blood systems and regional coordination across hospitals or regions. Developing more efficient algorithms to enhance robustness and scalability for larger and more complex instances would further strengthen its practical relevance.

Author contributions

Liu Zhao: Conceptualization, Data curation, Investigation, Methodology, Software, Validation, Visualization, Writing–original draft, and Writing–review and editing. Nengmin Wang: Funding acquisition, Project administration, Supervision, and Writing–review and editing. Yinfeng Xu: Supervision. Meng Zhang: Writing–review and editing.

Use of Generative-AI tools declaration

During the preparation of this manuscript, the authors used DeepSeek for language polishing and expression improvement. All AI-assisted content was critically reviewed and revised by the authors, who take full responsibility for the final published version.

Acknowledgments

This work was supported in part by the Major Project of National Natural Science Foundation of China under Grant 72192830 and Grant 72192834.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. World Health Organization, Blood transfusion safety, 2025. Available from: https://www.who.int/health-topics/blood-transfusion-safety#tab=tab_1.
2. S. Narayan, D. Poles, Headline data: Deaths, major morbidity and ABO-incompatible transfusions, *SHOT UK Annual Report 2024,* 2025. Available from: https://www.shotuk.org/wp-content/uploads/2025/07/3.-Headline-Data-Deaths-Major-Morbidity-and-ABO-Incompatible-Transfusions-2024.pdf.
3. K. Ouyang, The UK blood supply is in crisis and needs 200,000 new donors to join, China News Service, 2025. Available from: https://www.chinanews.com.cn/gj/2025/06-10/10429424.shtml.
4. The Portugal Post, Summer blood drought forces Portuguese hospitals to delay operations, 2025, Available from: https://theportugalpost.com/posts/summer-blood-drought-forces-portuguese-hospitals-to-delay-operations.
5. A. F. Osorio, S. C. Brailsford, H. K. Smith, A structured review of quantitative models in the blood supply chain: A taxonomic framework for decision-making, *Int. J. Prod. Res.*, **53** (2015), 7191–7212. https://doi.org/10.1080/00207543.2015.1005766

6. V. Hemmelmayr, K. F. Doerner, R. F. Hartl, M. W. Savelsbergh, Delivery strategies for blood products supplies, *OR Spectrum*, **31** (2009), 707–725. https://doi.org/10.1007/s00291-008-0134-7
7. V. Hemmelmayr, K. F. Doerner, R. F. Hartl, M. W. Savelsbergh, Vendor managed inventory for environments with stochastic product usage, *Eur. J. Oper. Res.*, **202** (2010), 686–695. https://doi.org/10.1016/j.ejor.2009.06.003
8. S. M. Kazemi, M. Rabbani, R. Tavakkoli-Moghaddam, F. A. Shahreza, Blood inventory-routing problem under uncertainty, *J. Intell. Fuzzy Syst.*, **32** (2017), 467–481. https://doi.org/10.3233/JIFS-152175
9. F. Jafarkhan, S. Yaghoubi, An efficient solution method for the flexible and robust inventory-routing of red blood cells, *Comput. Ind. Eng.*, **117** (2018), 191–206. https://doi.org/10.1016/j.cie.2018.01.029
10. W. Liu, G. Y. Ke, J. Chen, L. Zhang, Scheduling the distribution of blood products: A vendor-managed inventory routing approach, *Transp. Res. Pt. e-Logist. Transp. Rev.*, **140** (2020), 101964. https://doi.org/10.1016/j.tre.2020.101964
11. F. Y. Vincent, N. Y. Salsabila, A. Gunawan, N. Siswanto, A three-stage matheuristic for the blood stochastic inventory routing problem, *Transp. Res. Pt. e-Logist. Transp. Rev.*, **200** (2025), 104143. https://doi.org/10.1016/j.tre.2025.104143
12. L. Qing, Y. Yin, J. Ignatius, D. Wang, Robust multi-period blood inventory routing under multiple uncertainties, *Eur. J. Oper. Res.*, **328** (2026), 137–161. https://doi.org/10.1016/j.ejor.2025.05.036
13. S. Khalilpourazari, H. Hashemi Doulabi, A flexible robust model for blood supply chain network design problem, *Ann. Oper. Res.*, **328** (2023), 701–726. https://doi.org/10.1007/s10479-022-04673-9
14. S. Khalilpourazari, S. Soltanzadeh, G. W. Weber, S. K. Roy, Designing an efficient blood supply chain network in crisis: neural learning, optimization and case study, *Ann. Oper. Res.*, **289** (2020), 123–152. https://doi.org/10.1007/s10479-019-03437-2
15. S. Rekabi, Z. Sazvar, R. Tavakkoli-Moghaddam, A data-driven approach to optimize a blood supply chain within the Industry 5.0 framework: A stochastic optimization model, *Expert Syst. Appl.*, **289** (2025), 127960. https://doi.org/10.1016/j.eswa.2025.127960
16. H. Shirazi, R. Kia, P. Ghasemi, A stochastic bi-objective simulation–optimization model for plasma supply chain in case of COVID-19 outbreak, *Appl. Soft Comput.*, **112** (2021), 107725. https://doi.org/10.1016/j.asoc.2021.107725
17. H. Yang, Y. Yin, D. Wang, T. C. E. Cheng, R. Zhang, H. Hu, An integrated blood supply chain network design during a pandemic, *Int. J. Prod. Res.*, **63** (2025), 3384–3399. https://doi.org/10.1080/00207543.2024.2396511
18. B. Zahiri, S. A. Torabi, M. Mohammadi, M. Aghabegloo, A multi-stage stochastic programming approach for blood supply chain planning, *Comput. Ind. Eng.*, **122** (2018), 1–14. https://doi.org/10.1016/j.cie.2018.05.041
19. M. Momenitabar, Z. Dehdari Ebrahimi, M. Arani, J. Mattson, Robust possibilistic programming to design a closed-loop blood supply chain network considering service-level maximization and lateral resupply, *Ann. Oper. Res.*, **328** (2023), 859–901. https://doi.org/10.1007/s10479-022-04930-x
20. H. Shaabani, A literature review of the perishable inventory routing problem, *Asian J. Shipping Logist.*, **38** (2022), 143–161. https://doi.org/10.1016/j.ajsl.2022.05.002
21. A. Azadeh, S. Elahi, M. H. Farahani, B. Nasirian, A genetic algorithm-Taguchi based approach to inventory routing problem of a single perishable product with transshipment, *Comput. Ind. Eng.*, **104** (2017), 124–133. https://doi.org/10.1016/j.cie.2016.12.019

22. A. Timajchi, S. M. J. M. Al-e-Hashem, Y. Rekik, Inventory routing problem for hazardous and deteriorating items in the presence of accident risk with transshipment option, *Int. J. Prod. Econ.*, **209** (2019), 302–315. https://doi.org/10.1016/j.ijpe.2018.01.018
23. Z. Dai, K. Gao, B. C. Giri, A hybrid heuristic algorithm for cyclic inventory-routing problem with perishable products in VMI supply chain, *Expert Syst. Appl.*, **153** (2020), 113322. https://doi.org/10.1016/j.eswa.2020.113322
24. D. Cuellar-Usaquén, M. W. Ulmer, C. Gomez, D. Álvarez-Martínez, Adaptive stochastic lookahead policies for dynamic multi-period purchasing and inventory routing, *Eur. J. Oper. Res.*, **318** (2024), 1028–1041. https://doi.org/10.1016/j.ejor.2024.06.011
25. A. Diabat, T. Abdallah, T. Le, A hybrid tabu search based heuristic for the periodic distribution inventory problem with perishable goods, *Ann. Oper. Res.*, **242** (2016), 373–398. https://doi.org/10.1007/s10479-014-1640-4
26. H. Shaabani, I. N. Kamalabadi, An efficient population-based simulated annealing algorithm for the multi-product multi-retailer perishable inventory routing problem, *Comput. Ind. Eng.*, **99** (2016), 189–201. https://doi.org/10.1016/j.cie.2016.07.022
27. W. Hu, A. Toriello, M. Dessouky, Integrated inventory routing and freight consolidation for perishable goods, *Eur. J. Oper. Res.*, **271** (2018), 548–560. https://doi.org/10.1016/j.ejor.2018.05.034
28. A. Alvarez, J. F. Cordeau, R. Jans, P. Munari, R. Morabito, Formulations, branch-and-cut and a hybrid heuristic algorithm for an inventory routing problem with perishable products, *Eur. J. Oper. Res.*, **283** (2020), 511–529. https://doi.org/10.1016/j.ejor.2019.11.015
29. Y. Zhang, F. Chu, A. Che, Y. Li, Closed-loop inventory routing problem for perishable food with returnable transport items selection, *Int. J. Prod. Res.*, **62** (2024), 501–521. https://doi.org/10.1080/00207543.2023.2275639
30. S. Mirzaei, A. Seifi, Considering lost sale in inventory routing problems for perishable goods, *Comput. Ind. Eng.*, **87** (2015), 213–227. https://doi.org/10.1016/j.cie.2015.05.010
31. W. J. Bell, L. M. Dalberto, M. L. Fisher, A. J. Greenfield, R. Jaikumar, P. Kedia, et al., Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer, *Interfaces*, **13** (1983), 4–23. https://doi.org/10.1287/inte.13.6.4
32. L. C. Coelho, J. F. Cordeau, G. Laporte, Thirty years of inventory routing, *Transp. Sci.*, **48** (2014), 1–19. https://doi.org/10.1287/trsc.2013.0472
33. Y. Li, F. Chu, C. Feng, C. Chu, M. Zhou, Integrated production inventory routing planning for intelligent food logistics systems, *IEEE T. Intell. Transp.*, **20** (2019), 867–878. https://doi.org/10.1109/TITS.2018.2835145
34. J. Cheng, Y. Zhou, K. L. Teo, Bi-level programming for emergency blood transshipment-allocation considering substitution against public health events, *J. Ind. Manag. Optim.*, **20** (2024), 2697-2718. https://doi.org/10.3934/jimo.2024022
35. G. Agac, B. Baki, I. M. Ar, H. T. Kahraman, A supply chain network design for blood and its products using genetic algorithm: A case study of Turkey, *J. Ind. Manag. Optim.*, **19** (2023), 5407-5446. https://doi.org/10.3934/jimo.2022179
36. S. A. Darestani, S. Mousazadeh, Modeling a production-inventory-routing problem of blood products using heuristic solution methods, *J. Intell. Fuzzy Syst.*, **37** (2019), 5589–5609. https://doi.org/10.3233/JIFS-190723
37. M. R. G. Samani, S. M. Hosseini-Motlagh, S. Homaei, A reactive phase against disruptions for designing a proactive platelet supply network, *Transp. Res. Pt. e-Logist. Transp. Rev.*, **140** (2020), 102008. https://doi.org/10.1016/j.tre.2020.102008

38. Y. Xu, J. Szmerekovsky, The impact of transshipment on an integrated platelet supply chain: A multi-stage stochastic programming approach, *Comput. Ind. Eng.*, **176** (2023), 108991. https://doi.org/10.1016/j.cie.2023.108991
39. E. Farrokhizadeh, S. A. Seyfi-Shishavan, S. I. Satoglu, Blood supply planning during natural disasters under uncertainty: a novel bi-objective model and an application for red crescent, *Ann. Oper. Res.*, **319** (2022), 73–113. https://doi.org/10.1007/s10479-021-03978-5
40. S. U. K. Rohmer, G. D. H. Claassen, G. Laporte, A two-echelon inventory routing problem for perishable products, *Comput. Oper. Res.*, **107** (2019), 156–172. https://doi.org/10.1016/j.cor.2019.03.015
41. F. Alkaabneh, A. Diabat, H. O. Gao, Benders decomposition for the inventory vehicle routing problem with perishable products and environmental costs, *Comput. Oper. Res.*, **113** (2020), 104751. https://doi.org/10.1016/j.cor.2019.07.009
42. M. Kumar, D. Kumar, P. Saini, S. Pratap, Inventory routing model for perishable products toward circular economy, *Comput. Ind. Eng.*, **169** (2022), 108220. https://doi.org/10.1016/j.cie.2022.108220
43. M. Soysal, J. M. Bloemhof-Ruwaard, R. Haijema, J. G. A. J. V. D. Vorst, Modeling a green inventory routing problem for perishable products with horizontal collaboration, *Comput. Oper. Res.*, **89** (2018), 168–182. https://doi.org/10.1016/j.cor.2016.02.003
44. Z. He, Y. Liu, K. Liu, Robust optimization approaches for the perishable product inventory routing problem with demand uncertainty, *J. Ind. Manag. Optim.*, **20** (2024), 2740–2769. https://doi.org/10.3934/jimo.2024024
45. E. Nikzad, M. Bashiri, F. Oliveira, Two-stage stochastic programming approach for the medical drug inventory routing problem under uncertainty, *Comput. Ind. Eng.*, **128** (2019), 358–370. https://doi.org/10.1016/j.cie.2018.12.055
46. F. Niakan, M. Rahimi, A multi-objective healthcare inventory routing problem; a fuzzy possibilistic approach, *Transp. Res. Pt. e-Logist. Transp. Rev.*, **80** (2015), 74–94. https://doi.org/10.1016/j.tre.2015.04.010
47. A. Bishara, M. Lin, J. Digitale, J. W. Yim, H. Soleimani, D. Robinowitz, AI-guided surgical blood readiness: overcoming real-world challenges in prospective validation for safer, more efficient blood preparation, *NEJM AI*, **2** (2025), AIoa2401198. https://doi.org/10.1056/aioa2401198
48. Y. Haimes, On a bicriterion formulation of the problems of integrated system identification and system optimization, *IEEE Trans. Syst. Man Cybern.*, **1** (1971), 296–297. https://doi.org/10.1109/TSMC.1971.4308298
49. N. Ouertani, H. Ben-Romdhane, S. Krichen, A decision support system for the dynamic hazardous materials vehicle routing problem, *Oper. Res.*, **22** (2022), 551–576. https://doi.org/10.1007/s12351-020-00562-w

AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)