Article:
Wu, W, Ma, J, Liu, R orcid.org/0000-0003-0627-3184 et al. (1 more author) (2022) Multi-class hazmat distribution network design with inventory and superimposed risks. Transportation Research Part E: Logistics and Transportation Review, 161. 102693. ISSN 1366-5545

https://doi.org/10.1016/j.tre.2022.102693

Manuscript accepted for publication in Transportation Research Part E, March 2022

Multi-class hazmat distribution network design with inventory and superimposed risks

Weitiao Wu[a1], Jian Ma[a], Ronghui Liu[b], Wenzhou Jin[a]

a. School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China

b. Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, United Kingdom

Abstract: Transportation and inventory are essential to hazardous materials logistics, while different classes of hazardous materials are often transported over a network simultaneously. Despite their in-transit and storage incompatibility, the superimposed risks among different materials, which results from possible chemical reaction once accidents (e.g., leakage, explosion) happen, further complicate the comprehensive plans. In this study, we introduce a new multi-class hazmat distribution network design problem with inventory and superimposed risks (MHND) in a multi-echelon supply chain, where the planning of locations, inventory, and routes are made together. The long-term detour cost/risk and the cyclic time windows penalty costs under the time-dependent (periodic) road closure policy are explicitly formulated. We further propose a new population-based risk definition that evaluates the risk for the population at any location and any time with respect to its multi-class hazmat logistics system. In particular, to capture the interactions between different types of materials, we introduce risk superposition coefficients to capture possible superimposed risks among different hazmat that accommodate a general system with more than two hazmat types. We develop a knowledge-based NSGA-II algorithm with cyclic dissimilarity-based elitist selection (NSGA-II-CD) to solve the problem. The devised cyclic dissimilarity-based elitist selection (CD) operator can tackle the issue of speeding proliferation, which greatly improves the solution quality. Our model is applied to a metropolitan-wide real-world case study in Guangzhou, China. The results suggest that, from the perspective of the traffic management sector, the periodic road closures policy in Guangzhou could be possibly upgraded to a full-time prohibition. Moreover, the results provide the following insights to authorities (1) there is a positive convex relation between risk minimization and risk equilibration. The authorities should not try to find a perfect distribution of risk, and they should make a trade-off between the risk equity and total exposed risk; (2) there is a positive correlation between the level of service and total risk. Thus, in practice, the agencies should make a trade-off between economic viability of the system, exposed risk, and maintaining good service for customers; and (3) the interactions between different types of hazmat considerably affect the distribution network design; specifically, the route overlapping ratio for different types of hazmat decreases when their interactions intensify.

[1] Corresponding author, E-mail: ctwtwu@scut.edu.cn

Keywords: Distribution network design; Multi-class hazmat; Risk assessment; Multi-objective optimization; cyclic dissimilarity-based elitist selection

1. Introduction

Hazardous materials (hazmat) form an important integral part of industrial raw material and economic development in many industrialized countries. In China for instance the annual tonnage of ethylene shipped has reached 18.218 million tons, while Guangzhou ranks the first in production with 13.6%. Daya Bay in Huizhou city has become one of the seven largest petrochemical production bases in China. Large amounts of hazmat are frequently transported routinely. However, the tremendous use of hazmat also brings about growing risks that may seriously imperil humans, property, and the environment. As such, the management of hazmat logistics is essential to sustain our industrial lifestyle in that the transportation cost and safety should be considered concurrently.

The overall process of hazmat transportation can be divided into several phases with environmental risk, including loading, in transit, storage, unloading, and waste treatment. Due to the nature of hazardous materials, each production, storage, and transportation activity associated with their use presents an inherent risk for both society and the environment. An analysis for 207 hazmat accidents reports that 38.6% of hazmat accidents can produce the domino effect (Kourniotis, et al., 2000). Once an explosion or leakage takes place, catastrophic interlocking accidents will easily occur. To reduce the potential negative impact of catastrophic interlocking accidents, collaborative management for the overall process of hazmat transportation is imperative. Besides technical improvement and regulation on the hazmat routing, the warehouse location and inventory control are also important aspects of risk reduction. Warehouse locations directly affect the routing options, while any specific routes can, in turn, exert influence on the location decision of warehouses (Wu et al., 2021). The Chinese government has enforced regulations to reduce the risk stemming from the facility location and inventory management after the 2015 Tianjin Port explosion. As such, in the hazmat supply chain management, the risk in the transportation and other relevant stages (e.g., production and inventory) should be considered comprehensively. Unfortunately, to date, most previous works on hazmat logistics only investigated these issues separately and failed to integrate the planning of locations, routes, and inventory decisions. Treating these decisions separately may lead to not only excess costs but also a notable risk in managing the hazmat. Motivated by this fact, this paper investigates the distribution network design for hazardous materials, where the planning locations, inventory, and routes are made together.

As reported in the Globally Harmonized System of Classification and Labeling of Chemicals (GHS), hazardous materials can be classified into several types such as flammable, poisonous, and corrosive substances, according to the dimensions in physical, health, and environmental damage. The logistics costs and risks of various hazardous materials are quite distinct, such that the logistics costs are not necessarily higher for the hazmat with higher risk. Moreover, despite their in-transit and storage incompatibility, the superimposed risks among different materials, which results from possible chemical reaction once accidents (e.g., leakage, explosion) happen, complicate the comprehensive plans. The review of the

literature reveals, however, that different streams of research investigate hazmat transportation only from a single-hazmat-type point of view. Thus, we fill this gap by integrating multiple classes of hazmat, and the interaction among different hazmat types in the hazmat network design.

Another challenge of hazmat network design comes from the time-dependent road closures to hazmat carriers, particularly within an urban environment. In a metropolitan area, heavy traffic congestion can take hours in peak hours. To protect the public and environment, regulatory agencies have adopted a series of countermeasures to restrict hazmat transportation, notably with time-dependent (periodic) road closures policy. However, as noted by Kara and Verter (2004), the government intervention in route choices of carriers can significantly reduce transportation risks at the expense of additional high transportation costs. Since the location-allocation involves great capital investment and is unalterable, the long-term effect of time-dependent road closures and resulting detours should be explicitly incorporated in the design of the hazmat distribution network.

Putting together the above challenges, this study contributes to developing a multi-class hazmat distribution network design problem (MHND) in a multi-echelon supply chain. The superimposed risks among different hazmat, which result from a possible chemical reaction, are captured. The costs and risks induced by plant location, inventory level, and vehicle routes are jointly included in the design of the hazmat logistics system. Our model is applied to a real-world case study in Guangzhou, China. Managerial insights are also provided.

2. Literature review and main contributions

Hazmat logistics management problems have received attention from the operations research and transportation community since the 1980s (Erkut et al., 2007), whereas the problems remain a prominent research topic because of their practical significance and inherent complexity. The literature of hazmat logistics management problems can be generally categorized into the following groups: network design, hazmat routing, and scheduling, and risk assessment. In this section, we present an overview of the relevant representative literature.

2.1 hazmat network design

The purpose of the hazmat network design problem is to optimally determine combined facility locations and routing. A typical application lies in the disposal of hazardous waste, which determines various waste facilities and the corresponding routes of hazardous wastes and waste residues between facilities. Kara and Verter (2004) developed a bi-level programming model of hazmat network design to build the relationship between the regulator and carriers. Zhao et al. (2016) investigated the network design problem for regional hazardous waste management. Samanliglu (2013) developed a multi-objective location-routing model for industrial hazardous waste management. The model decides on the locations of treatment and recycling facilities, as well as the routing of collection vehicles. Fontaine and Minner (2018) proposed a multi-cut Benders decomposition method for the hazmat transportation network design problem. Rabbani et al (2019) proposed a stochastic multi-period industrial hazardous waste location routing

problem. Ghaderi and Burdett (2019) investigated the location and routing problem for hazardous materials over an intermodal transportation network, where the facility disruption is also considered. Recently, Hu et al (2019) develop a multi-objective location-routing model for hazmat logistics with consideration of inter-city traffic restriction constraints. The objective is to minimize the total risk and total cost. A few studies (e.g., Dadkar et al, 2010; Reilly et al, 2012) investigated how to determine which facilities to restrict the hazmat movement and for which time. The models are based on a three-party game that involves government, carrier, and terrorists.

2.2 hazmat routing (and scheduling)

Another line of research is associated with vehicle schedule problems. Verma et al. (2011) developed a tactical planning model for railroad hazmat transportation. Assadipour et al. (2015) incorporated the congestion effect at intermodal yards and equipment capacity decisions into intermodal hazmat transportation. Hosseini and Verma (2017) used a value-at-risk approach to route hazmat shipment, considering the low probability-high consequence nature of hazmat incidents. Later, Hosseini and Verma (2018) presented a conditional value-at-risk methodology for routing rail hazmat shipment, which shows that such a method outperforms other risk assessments. Kumar et al. (2018) developed an integrated optimization model for mixed hazmat fleet and routing, considering the monetary loss due to en route trunk type-dependent stoppages. Fan et al. (2015) formulated a hazmat routing optimization model subject to road closure. Bronfman et al. (2016) introduced a hazmat routing optimization problem that minimizes the risk to vulnerable centers. Szeto et al. (2017) addressed the hazmat routing problem given unknown or inaccurate incident probabilities, which was achieved by a multi-demon formulation. Mohri et al. (2020) addressed a hazmat routing-scheduling problem considering fairness among dispatchers.

2.3 Research gaps, objectives, and contributions

A comparison between this study and the available representative literature on hazmat transportation problems can be found in Table 1. A few gaps are identified from the existing literature. First, although substantial research effort has been made into hazmat network design, the integration of location and inventory decisions in hazmat logistics has not been fully addressed in the existing literature. Addressing these decisions separately may underestimate not only the system efficacy but also notable risk in managing the hazmat. Second, traditional research on hazmat logistics still lacks practicability. A critical issue is that existing studies primarily focus on a single type of merchandise, and more importantly, the superimposed risks among multiple hazardous materials have not yet been considered. Third, the impact of time-dependent traffic restriction policy on the long-term planning of hazmat networks has not been fully addressed.

Table 1 Overview of representative hazmat transportation problems

	Problem variant					
References	Routing	Location	Inventory	Mode	Objectives	Characteristics
Fontaine and Minner (2018)	✓			T	R	S
Ghaderi and Burdett (2019)	✓	✓		I	R,C	S
Verma et al. (2011)	✓			RR	R	S

Assadipour et al. (2015)	✓			I	R,C	S
Hosseini and Verma (2017)	✓			T	R	S
Kumar et al. (2018)	✓			T	R,C	S
Bronfman et al. (2016)	✓			T	R	S
Bronfman et al. (2019)	✓	✓		T	R	M
Kheirkhah et al. (2015)	✓			T	R,C	S
Paredes-Belmar et al. (2017)	✓			T	R,C	M
Mahmoudsoltani et al. (2018)	✓			T	R,C	S
Fan et al. (2015)	✓			T	R	S,RC
Xie et al. (2012)	✓	✓		I	R,C	S
Szeto et al. (2017)	✓			T	R	S
Beneventti G et al. (2019)	✓			T	R	M
Romero et al. (2016)	✓	✓		T	E	S
Mohri et al. (2020)	✓			T	E	S
Fontaine et al. (2020)	✓			I	C	S
This study	✓	✓	✓	T	R,C	M,SR,RC

Mode: **RR:** Railroad, **T:** Trunk, **I:** Intermodal

Objectives: **R:** Min. Risk, **E:** Max. Risk equity, **C:** Min. Cost

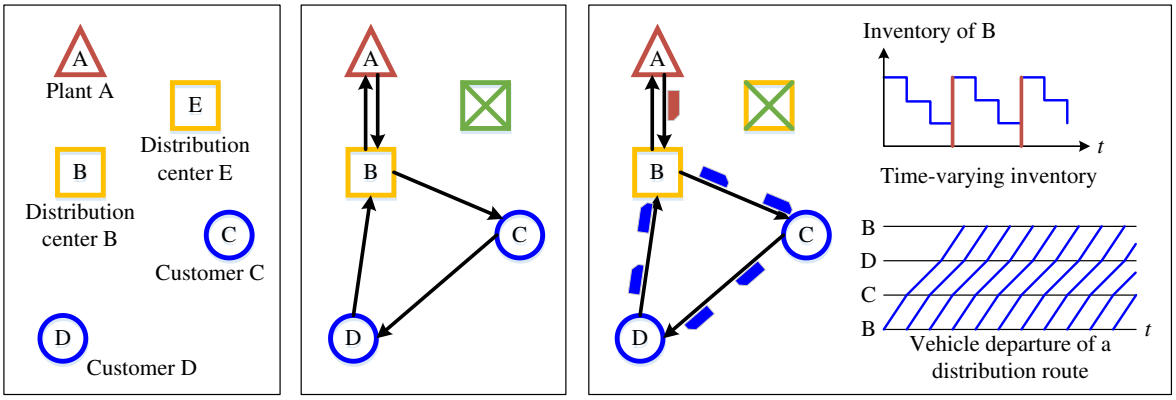
Characteristics: **S:** Single product, **M:** Multiple products, **SR:** Superimposed risk; **RC:** Time-dependent road closure

As a remedy, in this study, we develop a multi-objective model to optimally determine the locating, vehicle routing, and inventory level in a multi-echelon hazmat supply chain, to reduce the system total costs and risks comprehensively. The contributions of this paper are described as follows: (1) we introduce a new multi-class hazmat distribution network design problem with inventory and superimposed risk in a multi-echelon supply chain, where the planning of locations, inventory, and routes are made together. The long-term cyclic time windows penalty cost with a time-dependent (periodic) route closure policy is explicitly formulated. Our model is scalable for multiple types of hazmat, which has significant implications in the realm of hazmat logistics; (2) we propose a new population-based risk definition that evaluates the risk for the population at any location and any time with respect to its multi-class hazmat logistics system. In particular, we introduce risk superposition coefficients for multi-class hazmat, which allows for capturing possible superimposed risks (e.g., chemical reaction) among different hazmat and accommodates a general system with more than two hazmat types; (3) we develop a knowledge-based NSGA-II algorithm with cyclic dissimilarity-based elitist selection (NSGA-II-CD) to solve the problem. The cyclic dissimilarity-based elitist selection (CD) operator can tackle the issue of speeding proliferation, which greatly improves the solution quality; (4) we present a metropolitan-wide real-world case study and provide managerial insights. As far as we are aware, this is the first attempt in multi-class hazmat distribution network design with inventory and superimposed risk.

The rest of this paper is organized as follows. In Section 3, the model is presented, followed by the solution methodologies in Section 4. In Section 5, a metropolitan-wide real-world application is conducted and the managerial insights are provided. At last, the concluding remarks are given.

3 Model development

3.1 Problem description



(a) The long-term data (b) The scheme of a three-level supply chain (c) The long-running system with transportation and inventory strategy

Fig. 1 Planning framework of the hazmat distribution network of a three-level supply chain

Fig. 1 sketches the planning framework of a multi-class hazmat distribution network of a three-level supply chain that outputs deployment of DC locations, the distribution routes, and the corresponding order cycles. Two criteria are considered: (a) minimizing the total cost, including the opening costs for DCs, inventory cost transportation cost; (b) minimizing the total risk, which includes inventory risk and transportation risk associated with the population exposure, plus the superimposed risk affected by the chemical reactions. As the costs and risks vary over time in this system, the objective functions should concern the average values under long-term running. Under these evaluation criteria, as shown in Fig. 1(a), the first step is to acquire the long-term data including multiple plants, candidate capacitated distribution centers (DC), customer points, customer demand, and the restricted area. As shown in Fig. 1(b), the second step is to set up a hazmat distribution network by determining the optimal location of DCs, distribution routes, and the corresponding order cycles (departure headways of vehicles) given the long-term data and inventory policy. In the third step, as shown in Fig. 1(c), once the distribution network has been set up, the system will run for a long time, the inventories will change over time, and the vehicles depart periodically under the departure headways.

3.2 Assumptions and notations

To ease the model development, the basic assumptions or requirements are made for the MHND as follows:

(A1) Customer demands are independent and identically distributed, and follow a normal distribution.

(A2) Each plant center only produces one type of hazardous material, as with each distribution center and each customer. This is reasonable in practice due to the incompatibility between different hazardous materials.

(A3) The production capacity of the plants is sufficient to meet the demand.

(A4) The vehicle type is homogeneous.

(A5) To ensure transportation safety, a truck cannot simultaneously carry different hazardous materials.

(A6) Each customer is only served by one distribution center. This is because each distribution center and each customer only handles one type of hazardous material (A2).

(A7) Storage capacities exist for both distribution centers and customer points.

(A8) In this study, we investigate MHND specifically tailored for the economic-order-quantity (EOQ) policy, where customers determine their optimal order quantity based on the principle of economic order quantity. The reasons are two-fold: First, due to high-security standards, such a customer-oriented ordering strategy can reduce the terrorist threat of attacks by limiting inventory information sharing. This is supported by Reilly et al. (2012) who noted that hazmat risk includes not only natural risk but also induced risk generated by potential terrorist activities and that countermeasures should be taken to avoid as much potential attack as possible. Second, due to the generalized complexity of inventory-routing categories, a few computationally efficient policies have been introduced in the literature, such as order-up-to policy (Bartazzi et al., 2002), (Q, R) policy (Javid and Azad, 2010) and EOQ policy (Chen et al., 2017). Similar to Chen et al (2017), we adopt EOQ policy as the long-term replenishment strategy, which facilitates deriving the long-run average system-wide costs and risks associated with location, inventory rules, and routing patterns.

Table 2 lists the notations and their definitions used in our model.

Table 2 Primary notations used in this paper

Sets	
L	The set of hazardous material types, $L = \{l \mid l = 1,2, \dots, \lvert L \rvert\}$, which is also adopted as the set of corresponding plants (one plant for one type)
M	The set of candidate distribution centers, $M = \{m \mid m = \lvert L \rvert + 1, \lvert L \rvert + 2, \dots, \lvert L \rvert + \lvert M \rvert\}$
N	The set of customers, $N = \{n \mid n = \lvert L \rvert + \lvert M \rvert + 1, \lvert L \rvert + \lvert M \rvert + 2, \dots, \lvert L \rvert + \lvert M \rvert + \lvert N \rvert\}$
K	The set of distribution routes, $K = \{k \mid k = 1,2, \dots \lvert N \rvert\}$
Parameters	
$type_i^l$	1 if hazmat type l locates at node i; 0 otherwise. *i.e.*, $D_n type_i^l$ indicates that the demand of hazmat type l for customer n is D_n
$Dis_{i,j}$	Euclidean distance between node i and node j
vel	Average vehicle speed
D_n	The random consumption rate of customer n, $D_n \sim N(\mu_n, \sigma_n^2)$
S_i	The storage capacity of node i, $i \in M \cup N$
S^l	The vehicle capacity for hazmat type l
R_{max}	Maximum allowable risk
c_m	The fixed cost of opening distribution center m per unit time
F_i	The ordering cost of node i, $i \in M \cup N$
u	The fixed cost of vehicle departure, loading, and unloading
c_f	The fuel price
h_i	Holding cost for each unit of hazmat at node i per unit time, $i \in M \cup N$
Pw_n	The period of time window penalty cost function of customer n

tw_n	The desired arrival time at the first period of time window penalty cost function of customer n
mw_n	The maximum value of time window penalty cost function of customer n
r_l	Risk coefficient or accident probability of hazmat type l
τ_l	Attenuation coefficient of hazmat of type l
β	The critical ratio of accident impacts on a region to that on the whole plane
λ_l	The accident impact radius of hazmat of type l
ρ	Population density
$\delta_{ll'}$	The risk superposition coefficient between type l and type l'
$FT_{i,j}$	The restricted time set of the segment from i to j
Auxiliary variables	
t_i^k	The vehicle travel time from the distribution center to node i on route k
q_i^k	The vehicle load when arriving at node i on route k
Q_m	The order quantity for distribution center m
$\tilde{Q}_k$	The order quantity for distribution route k
St_m	The safety stock at distribution center m
P_m	The optimal order cycle for distribution center m
$\tilde{P}_k$	The order cycle or departure headway for distribution route k
U_m	The variable transportation cost of replenishment route from the plant to distribution center m
$\tilde{U}_k$	The variable transportation cost of distribution route k
$rs_{i,j,k}^l$	The average risk on the segment between node i and node j of route k (type l)
cw_n^k	The detour-modified time window penalty cost of customer n
Δc_k	The total detour cost on route k per unit time
Δr_k^l	The total detour risk on route k per unit time (type l)
Δt_i^k	The vehicle's detour time when traveling from the distribution center to node i on route k
A_i^l	The affected region by hazmat of type l at node i, $i \in L \cup M \cup N$
$\tilde{A}_{i,j}^l$	The affected region by hazmat of type l between node i and j, $i,j \in L \cup M \cup N$
$G(A)$	The population of region A
CL	The total location cost of opening distribution centers per unit time
CM	The total inventory cost per unit time
CR	The total cost of replenishment per unit time (from plants to distribution centers)
CD	The total cost of distribution per unit time (from distribution centers to customers)
RM	The total inventory risk per unit time
RS	The average superimposed risk
RR	Total replenishment risk per unit time (from plants to distribution centers)
RD	Total distribution cost per unit time (from distribution centers to customers)
Decision variables	
W_m	1 if distribution center m is opened, 0 otherwise
$X_{i,k}$	1 if node i is a visiting point of route k, 0 otherwise, $i \in M \cup N$

$Y_{i,j,k}$ 1 if the segment from node i to node j belongs to route k, 0 otherwise, $i \in M \cup N$

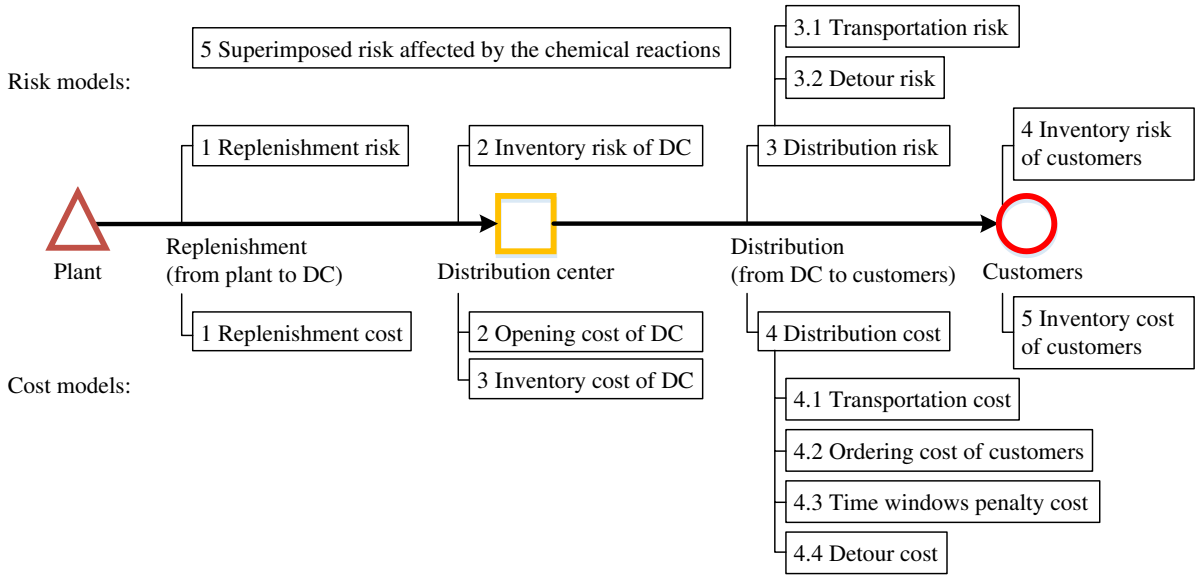


Fig. 2 Key cost and risk components for MHND

The overall three-level hazmat logistics system generates costs and risks in four aspects: location, inventory, routing, and time. The key cost and risk components for MHND are shown in Fig. 2. In the following sections, we derive analytical formulations for each cost and risk component.

3.3 Inventory models

3.3.1 The replenishment strategy for customers

As mentioned previously (A8), in the hazmat supply chain, the customers are assumed to determine their optimal order quantity based on the principle of economic order quantity to reduce the terrorist threat of attacks by limiting inventory information sharing. Following this principle, the reliable optimal delivery quantity of route k takes the following form (see detailed derivation in Appendix A).

$$\tilde{Q}_k = \frac{2\sum_{n\in N} X_{n,k}F_n}{\sum_{n\in N} X_{n,k}h_n\mu_n}\sum_{n\in N} X_{n,k}(\mu_n + \sigma_n z_\alpha) \tag{1}$$

$$\mu_k = \tilde{P}_k \sum_{n\in N} X_{n,k}\,\mu_n \tag{2}$$

$$\tilde{P}_k = \sqrt{\frac{2\sum_{n\in N} X_{n,k}F_n}{\sum_{n\in N} X_{n,k}h_n\mu_n}} \tag{3}$$

where z_α represents the quantile statistics with a confidence level of α. A higher value of α indicates a better level of service. In this way, the random variables due to uncertain demand have been transformed into deterministic values.

3.3.2 The replenishment strategy for distribution centers

The replenishment for distribution centers depends on the delivery quantity for customers. The expected hazmat consumption of route k is $\sum_{n\in N} X_{n,k}\mu_n$, then the expected total demand served by distribution center m is

$$\bar{D}_m = \sum_{n\in N}\sum_{k\in K} X_{n,k}X_{m,k}\mu_n \tag{4}$$

The optimal order quantity can be obtained by minimizing the total inventory cost, which is the summation of the ordering cost and inventory holding cost with safety stock. The ordering cost equals the ordering times (total demand divided by the ordering quantity) multiplied by the ordering cost. The inventory cost equals the unit holding cost multiplied by the average stock level, where the average stock level is half of the order quantity. The optimal order quantity Q_m for distribution center m is obtained applying first-order condition, which yields:

$$\frac{\bar{D}_m}{Q_m}\cdot F_m + h_m\left(\frac{Q_m}{2} + St_m\right) \Rightarrow Q_m = \sqrt{\frac{2\bar{D}_m F_m}{h_m}} \tag{5}$$

The lead time equals the order cycle of the distribution center, and the demand variation is $\sqrt{\sum_{n\in N}\sum_{k\in K} X_{m,k}X_{n,k}\sigma_n^2}$. Given the confidence level α and safety coefficient z_α, the safety stock can be determined as follows:

$$St_m = z_\alpha\sqrt{P_m\sum_{n\in N}\sum_{k\in K} X_{m,k}X_{n,k}\sigma_n^2} \tag{6}$$

Therefore, the order cycle for distribution center m can be determined as follows:

$$P_m = \frac{Q_m}{\bar{D}_m}\sqrt{\frac{2F_m}{h_m\sum_{n\in N}\sum_{k\in K} X_{m,k}X_{n,k}\mu_n}} \tag{7}$$

3.3.3 Location cost

The opening of a distribution center consumes a fixed cost. Hence, the fixed cost of distribution centers depends on the number of opened distribution centers:

$$CL = \sum_{m\in M} W_m c_m \tag{8}$$

3.3.4 Inventory cost

The average inventory cost is related to the unit holding cost, the average stock level, and the order cycle. As the average stock level is half of the optimal order quantity, the average inventory cost is $\frac{1}{2}\mu_n h_n$. Therefore, the total inventory costs of customers per unit time are $\frac{1}{2}\sum_{n\in N} h_n\mu_n$.

Similar to the inventory costs for customers, the average inventory cost for distribution centers is $\frac{1}{2}\sum_{m\in M}\sum_{k\in K}\sum_{n\in N} X_{m,k}h_m\mu_k$, and the inventory cost of the safety stock is $\frac{1}{2}\sum_{m\in M} W_m h_m St_m$. The total inventory cost for distribution centers is the summation of these two parts. As a result, the total inventory cost is the addition of those of customers and distribution centers.

$$CM = \frac{1}{2}\sum_{m\in M}\sum_{k\in K} X_{m,k}h_m\tilde{P}_k\mu_k + \frac{1}{2}\sum_{m\in M} W_m h_m St_m + \frac{1}{2}\sum_{n\in N} h_n\mu_n \tag{9}$$

3.4 Transportation models

3.4.1 Cyclic time windows penalty cost

In the supply chain, customers often have a requirement for the arrival time of goods due to factors such as production planning, loading and unloading storage costs, and natural conditions, and this requirement is usually repeated in natural cycles such as days, weeks, months, and years (e.g., daily 8 a.m., every Monday, 1st of each month, after the Spring Festival), particularly for the JIT systems. As an industrial raw material, the JIT delivery of hazmat is essential to production efficiency, customer experience, and satisfaction. Time windows penalty cost raises upon arrival at the undesirable time, such as the early and late arrival. Given this fact, the vehicle routing is conducted on a periodical basis under recycling supply, and the time windows violation penalty cost becomes a periodic function in the long run. For this reason, a multi-period time windows model in response to recycling supply, instead of single-period time windows, is adopted and incorporated into the system total cost. Prior to deriving the time windows violation penalty cost, the following properties are provided.

Proposition 1: As the independent variable of an infinite arithmetic progression $\{T_1 t_n + \varphi_1\}$, when the tolerance T_1 is not equal to the integral times of the trigonometric function's period T_2, the mean value of substituting the sequence elements into the function one by one is 0. When T_1 is equal to the integral times of T_2, the mean of substituting the sequence elements into the function one by one is a constant:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}\cos\left[\frac{2\pi}{T_2}(T_1 t+\varphi_1)+\varphi_2\right]=\begin{cases}0 & ,\forall\kappa\in N^+, T_1\neq\kappa T_2\\ \cos\left(\dfrac{2\pi}{T_2}\varphi_1+\varphi_2\right) & ,\exists\kappa\in N^+, T_1=\kappa T_2\end{cases} \tag{10}$$

Proof: A formal proof of this proposition can be found in Appendix B.

Corollary 1: For any periodic function with a period of T_2, when T_1 is not equal to the integral times of T_2, the mean function value of substituting sequence elements into the function one by one is equal to the mean value of this function in one period. When T_1 is equal to the integral times of T_2, the mean of substituting the sequence elements into functions one by one is a constant:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}f(T_1 t+\varphi_1)=\begin{cases}\dfrac{1}{T_2}\displaystyle\int_0^{T_2}f(t)\mathrm{d}t & ,\forall\kappa\in N^+, T_1\neq\kappa T_2\\ f(\varphi_1) & ,\exists\kappa\in N^+, T_1=\kappa T_2\end{cases} \tag{11}$$

Proof: A formal proof of this corollary can be found in Appendix C.

Generally, there are three models of cyclic time windows, as shown in Fig. 3(a)-(c), where the penalty cost changes periodically over time.

In each distribution cycle, the travel time t_i^k from the distribution center to-visited node i, is the sum of travel time of prior segments of route k:

$$t_i^k=t_j^k+\frac{Dis_{j,i}}{vel}\quad ,Y_{j,i,k}=1,\forall i,j\in M\cup N,\forall k\in K \tag{12}$$

Let cyc be the period index, then the vehicle arrival time of distribution in the long run is an infinite arithmetic progression $\{cyc\cdot\tilde{P}_k+t_i^k\}$, as shown in Fig. 3(d). In each distribution cycle, a vehicle will arrive at the customer n at a certain time and induce a certain penalty cost according to the arrival time

and the cyclic time windows model. When the arrival time happens to be the expected time point, such as t_0 in Fig. 3(a), or within the time window $[t_1, t_2]$ in Fig. 3(b), there will be no penalty cost for this distribution. The penalty cost at t_3 in Fig. 3(c) is positive infinity, which means that vehicle arrival at that time is not allowed.

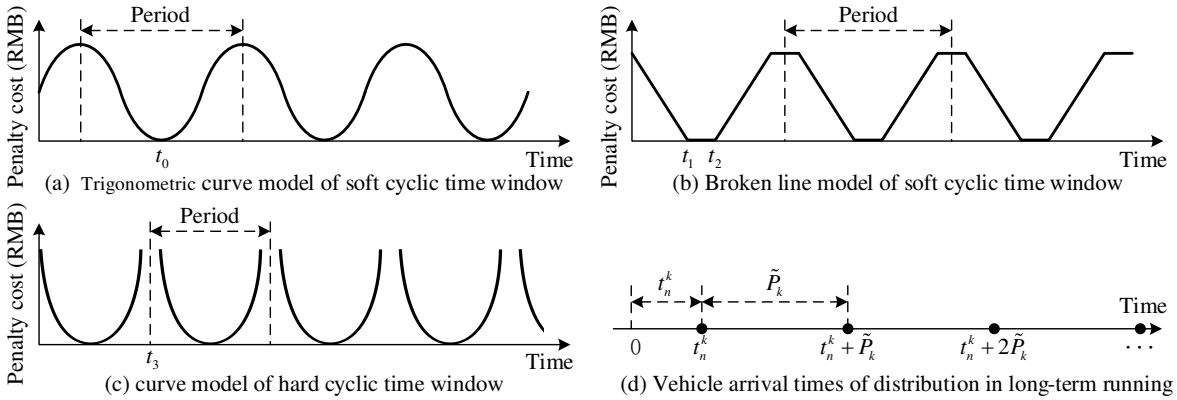


Fig. 3 Cyclic time windows model

Let Pw_n be the period of the penalty cost function. According to **Corollary 1**, for any soft cyclic time windows model $f_n(t)$, when the distribution period $\tilde{P}_k$ is not equal to the integral times of Pw_n, the mean penalty cost of the distribution period is equal to the mean value of $f_n(t)$ in a time window period; When $\tilde{P}_k$ is equal to the integral times of Pw_n, the mean penalty cost during the distribution period is equal to the penalty cost in the first distribution period:

$$\lim_{CYC\to+\infty}\frac{1}{CYC}\sum_{cyc=1}^{CYC} f_n\left(cyc\tilde{P}_k + t_n^k\right) = \begin{cases} \dfrac{1}{Pw_n}\displaystyle\int_0^{Pw_n} f_n(t)\,\mathrm{d}t & ,\forall\kappa\in N^+, \tilde{P}_k \neq \kappa Pw_n \\ f_n(t_n^k) & ,\exists\kappa\in N^+, \tilde{P}_k = \kappa Pw_n \end{cases}, X_{n,k}=1 \tag{13}$$

According to the Fourier transform theorem, any periodic functions with cycle Pw_n can be decomposed into a couple of trigonometric functions:

$$f_n(t) = \frac{1}{Pw_n}\int_0^{Pw_n} f_n(t)\,\mathrm{d}t + \sum_{\varepsilon=1}^{+\infty} A_\varepsilon \sin\left(\frac{2\pi}{Pw_n}\varepsilon t + \varphi_\varepsilon\right) \tag{14}$$

Since any periodic functions $f_n(t)$ can be decomposed into trigonometric functions with elegant periodic properties, we choose the sine/cosine curve in Fig. 3(a) to approximate the time window penalty, which takes the following form:

$$f_n(t) = mw_n sin\left[\frac{2\pi}{Pw_n}(t - tw_n) - \frac{\pi}{2}\right] + mw_n \tag{15}$$

where mw_n is the maximum value of penalty standing for the amplitude of the circular function. tw_n is the desired arrival time of the first period, which represents the initial phase of the circular function.

As the arrival time at customer n of each period is $t = cyc\cdot\tilde{P}_k + t_n^k$, $X_{n,k}=1$, using the trigonometric function induction formula, the sine function in Eq. (15) can be converted into the cosine function as follows:

$$\sin\left[\frac{2\pi}{Pw_n}\left(cyc\tilde{P}_k + t_n^k - tw_n\right) - \frac{\pi}{2}\right] = -\cos\left[\frac{2\pi}{Pw_n}\left(cyc\tilde{P}_k + t_n^k - tw_n\right)\right] \tag{16}$$

By summing up the above equations with constant terms over multiple periods, the average penalty cost for an infinite number of delivery cycles can be derived as follows:

$$f_n\left(cyc\tilde{P}_k + t_n^k\right) = \lim_{CYC\to+\infty} \frac{1}{CYC} \sum_{cyc=1}^{CYC} \left\{-mw_n \cos\left[\frac{2\pi}{Pw_n}\left(cyc\tilde{P}_k + t_n^k - tw_n\right)\right] + mw_n\right\} \tag{17}$$

According to **Proposition 1**, the intermediate-term in Eq. (17) can be rewritten as follows:

$$\lim_{CYC\to+\infty} \frac{1}{CYC} \sum_{cyc=1}^{CYC} \cos\left[\frac{2\pi}{Pw_n}\left(cyc\tilde{P}_k + t_n^k - tw_n\right)\right] = \begin{cases} 0 & ,\forall \kappa \in N^+, \tilde{P}_k \neq \kappa Pw_n \\ \cos\left[\frac{2\pi}{Pw_n}\left(t_n^k - tw_n\right)\right] & ,\exists \kappa \in N^+, \tilde{P}_k = \kappa Pw_n \end{cases} \tag{18}$$

Up to now, the penalty cost of time windows for customer n of each order cycle is identical. Therefore, the penalty cost per unit time for customer n of route k can be determined via dividing the average penalty cost for an infinite number of order cycles by the period length.

$$cw_n^k = \frac{f_n(t)}{\tilde{P}_k} = \begin{cases} \dfrac{mw_n}{2\tilde{P}_k} & ,\forall \kappa \in N^+, \tilde{P}_k \neq \kappa Pw_n \\ \dfrac{mw_n}{2\tilde{P}_k}\left\{1 - \cos\dfrac{2\pi}{Pw_n}\left[\left(t_n^k + \Delta t_n^k\right) - tw_n\right]\right\} & ,\exists \kappa \in N^+, \tilde{P}_k = \kappa Pw_n \end{cases}, X_{n,k} = 1 \tag{19}$$

where Δt_n^k represents the additional travel time caused by a detour from the distribution center to customer n on route k, whose specifics are described in Section 3.4.2.

3.4.2 Detour cost/risk/time of traffic restriction

In practice, trucks are not allowed to enter the urban area, especially during the daytime. Such a traffic restriction policy will induce detour cost/risk. Given the travel time from the distribution center to visited node i on route k, t_i^k (see Eq. (12)), the arrival time at visited node i of each vehicle is its travel time plus the cycles that have passed, that is, $t_i^k + cyc \cdot \tilde{T}_k$. To ensure that vehicles can run through the road sections before the starting restricted time, the vehicles need to enter the restricted road section t_0 in advance, i.e., $t_0 = Dis_{i,j}/vel$. Therefore, let the restricted time set of the segment from i to j be $FT_{i,j}$, and hazmat vehicles are restricted from t_1 to t_2 each day in this segment, then this set can be expressed as follows:

$$FT_{i,j} = [t_1 + t_0 : t_2) \cup \ldots \cup [t_1 + t_0 + 24\kappa : t_2 + 24\kappa) \quad ,\forall \kappa \in N^+, \kappa \to +\infty \tag{20}$$

Suppose the vehicle arrival time is t, when $t \in FT_{i,j}$, there will be a detour cost $\Delta c_{i,j}$. If no path is available, the detour cost can be set as infinity and the routing scheme is infeasible. Likewise, this will also result in extra detour risk $\Delta r_{i,j}^l$ and detour time $\Delta t_{i,j}$. ($\Delta t_{i,j}$ and $\Delta c_{i,j}$ can be obtained from the length of the shortest path between node i and node j without restricted sections. The corresponding detour risk $\Delta r_{i,j}^l$ can be obtained from the population along the above path based on Eq. (37) in Section 3.5.3.) Then the resulting detour cost $c(t)$ can be treated as the following periodic function:

$$c(t) = \begin{cases} \Delta c_{i,j} & ,t \in FT_{i,j} \\ 0 & ,t \notin FT_{i,j} \end{cases} \tag{21}$$

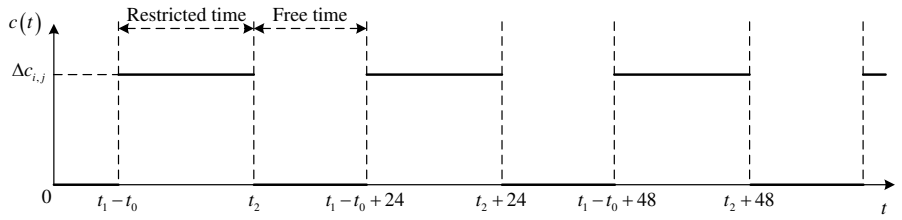


Fig. 4 Diagram of the periodic detour cost function

Fig. 4 shows the graph of $c(t)$ overtime. The type of this graph is termed as the rectangle wave in the physics domain, where the high level takes the larger function value. The proportion of high levels in a waveform cycle is termed the duty cycle. According to the Fourier transform, the rectangular wave function $c(t)$ with a period of 24 hours and a duty cycle of $\frac{t_2-t_1+t_0}{24}$ can be formulated as:

$$c(t) = \frac{t_2 - t_1 + t_0}{24}\Delta c_{i,j} + \sum_{\varepsilon=1}^{+\infty} A_\varepsilon \sin\left(\frac{2\pi}{24}\varepsilon t + \varphi_\varepsilon\right) \tag{22}$$

where the first (constant) term is the average value of the detour cost function in one day.

According to **Corollary 1**, for any periodic function with a period of T_2, when the independent variable is an infinite arithmetic progression and its tolerance is not equal to the integral times of function period (i.e., the distribution period $\tilde{P}_k$ is not equal to the integral times of 24 hours), the following equation holds:

$$\lim_{CYC\to+\infty} \frac{1}{CYC}\sum_{cyc=1}^{CYC} c\left(cyc\tilde{P}_k + t_i^k\right) = \frac{t_2 - t_1 + t_0}{24}\Delta c_{i,j} + 0 \quad , X_{i,k} = 1 \tag{23}$$

On the other hand, when the distribution period $\tilde{P}_k$ equals to the integral times of 24 hours, vehicles will encounter the same traffic restriction condition in each cycle such that the detour cost is identical. If the segment is not restricted, the restricted time set $FT_{i,j}$ is empty and the detour cost is zero. As a result, the total detour cost per unit of time on route k is the sum of average detour cost in each segment per distribution period divided by the period length.

$$\Delta c_k = \frac{1}{\tilde{P}_k}\sum_{i\in M\cup N}\sum_{j\in M\cup N}\sum_{k\in K} Y_{i,j,k} \times \begin{cases} \dfrac{t_2 - t_1 + t_0}{24}\Delta c_{i,j} & , \forall \kappa \in N^+, \tilde{P}_k \neq 24\kappa \\ 0 & , \exists \kappa \in N^+, \tilde{P}_k = 24\kappa, t_i^k \notin FT_{i,j} \\ \dfrac{\Delta c_{i,j}}{\kappa} & , \exists \kappa \in N^+, \tilde{P}_k = 24\kappa, t_i^k \in FT_{i,j} \end{cases} \tag{24}$$

The total detour risk (change caused by detour) per unit of time on route k, Δr_k^l can be calculated from $\Delta r_{i,j}^l$ in the same way. Note that the risk is associated with type l while the cost is not. Again, the cumulative detour time Δt_i^k (change of travel time caused by detour) when vehicles travel from the distribution center to node i on route k is:

$$\Delta t_i^k = \Delta t_j^k + \Delta t_{j,i} \quad , Y_{j,i,k} = 1, \forall k \in K \tag{25}$$

3.4.3 Transportation cost from distribution centers to customers

The transportation cost includes the fixed cost and variable cost. The former refers to the nonrecurring costs due to vehicle start-ups, such as vehicle delivery management and driver scheduling, while the latter is associated with the mileage. To better reflect the reality, in our model the vehicle fuel consumption cost is incorporated into the variable transportation cost and total cost. Herein, we adopt the Comprehensive Model Emission Model (CMEM) proposed by Barth et al. (2005) to calculate the vehicle fuel consumption. In the CMEM model, the vehicle fuel consumption F (unit: Litre) is related to the driving speed vel, driving range d and vehicle load q. Similar to Huang et al. (2007), the fuel computation formula is expressed as follows:

$$F = 0.28\frac{d}{vel} + 1.529 \times 10^{-3} dvel^2 + 8.403 \times 10^{-3} d(3.85 + q) \tag{26}$$

The travel time between customers i and j is $Dis_{i,j}/vel$, and the corresponding load is:

$$q_j^k = q_i^k - \mu_i \quad , Y_{i,j,k} = 1, \forall i, j \in N, \forall k \in K \tag{27}$$

Then the variable transportation cost of route k equals the route-specific total fuel consumption multiplied by the fuel price c_f.

$$\widetilde{U}_k = c_f \sum_{i \in M \cup N} \sum_{j \in M \cup N} Y_{i,j,k} \left[0.28\frac{Dis_{i,j}}{vel} + 1.529 \times 10^{-3} Dis_{i,j} vel^2 + 8.403 \times 10^{-3} Dis_{i,j} \left(3.85 + q_j^k\right) \right] \tag{28}$$

Consequently, the total transportation cost per unit time equals the transportation cost of a delivery divided by the order cycle plus the time windows penalty cost and the cost changed by detour.

$$CD = \sum_{n \in N} \sum_{k \in K} X_{n,k} \left(\frac{u + \widetilde{U}_k + F_n}{\tilde{P}_k} + cw_n^k + \Delta c_k \right) \tag{29}$$

3.4.4 Replenishment cost from plants to distribution centers

The replenishment cost includes the ordering cost and transportation cost. The distance from a plant to a distribution center and the load are $Dis_{l,m}$ and Q_m, respectively. Alternately, the detour journey is empty with a distance of $Dis_{m,l}$. Similar to Section 3.4.3, the variable transportation cost can be derived as:

$$\begin{aligned} U_{m,l} = {} & 0.28\frac{Dis_{l,m} + Dis_{m,l}}{vel} + 1.529 \times 10^{-3} \left(Dis_{l,m} + Dis_{m,l}\right) vel^2 \\ & + 8.403 \times 10^{-3} \left[Dis_{l,m} Q_m + 3.85 \left(Dis_{l,m} + Dis_{m,l}\right) \right] \end{aligned} \tag{30}$$

The replenishment cost per unit time is the cost of a replenishment divided by the order cycle, thus the total replenishment cost of all distribution centers per unit time takes the following form:

$$CR = \sum_{l \in L} \sum_{m \in M} type_m^l \frac{u + U_{m,l} + F_m}{P_m} \tag{31}$$

3.5 Risk assessment models

The storage of hazmat forms a dangerous area, and their transportation could be viewed as a movement of this area along a route. The severity of an explosion accident should not only account for the impact area but also the exposed receptors in the vicinity. Therefore, in this study, the risk is defined as the product of

accident frequency, exposed population, and economic loss per capita. In what follows, the node risk and arc risk are first derived, based on which the systematic inventory risk and transportation risk are derived.

3.5.1 Exposed population

The residents in the vicinity are considered as the exposed receptors. Given the affected region A, and the population density $\rho(x,y)$ of any point (x,y) on the plane, then the exposed population $G(A)$ can be calculated as the double integrating $\rho(x,y)$ in region A, or the average value of function $\rho(x,y)$ in region A multiplied by the corresponding area $S(A)$. In practice, $\rho(x,y)$ is difficult to acquire in a straightforward, but it can be approximated by randomly or equidistantly sampling $\hat{a}$ points in region A:

$$G(A) = \iint_A \rho(x,y)\mathrm{d}x\mathrm{d}y = S(A) \times \rho_A \approx S(A)\frac{1}{\hat{a}}\sum_{a=1}^{\hat{a}} \rho(x_a, y_a) \quad ,(x_a, y_a) \in A \tag{32}$$

3.5.2 Node(inventory) risk

A circular explosive field is used to represent the range and intensity of the impact of hazmat accidents that attenuate from the inside out. Since the impact intensity decreases quickly as the distance increases, given the location of the hazmat (x_0, y_0) and the attenuation coefficient τ_l of hazmat type l (representing the rate at which the impact intensity decreases with the distance), the impact intensity at the location (x,y) can be formulated in an exponential way as (Carotenuto et al., 2007):

$$e^{-\tau_l[(x-x_0)^2+(y-y_0)^2]} \tag{33}$$

Therefore, taking the location (x,y) as the origin, and taking the anomalous integral of Eq. (33) yields the impact of hazmat accident on the whole plane as follows:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\tau_l[(x-x_0)^2+(y-y_0)^2]}\mathrm{d}x\mathrm{d}y = \int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty} e^{-\tau_l r^2} r\mathrm{d}r = \frac{\pi}{\tau_l} \tag{34}$$

where θ is an auxiliary variable used for the surface integral.

As shown in Fig. 5, assume that the affected region by hazmat accident is a circle. When the ratio of accident impact on a circle to that on the whole plane is higher than a critical value β (e.g., 90%), it can be recognized that the accident effects are distributed within the circle, then its impact radius λ_l can be determined to be

$$\int_0^{2\pi}\mathrm{d}\theta\int_0^{\lambda_l} e^{-\tau_l r^2} r\mathrm{d}r = \beta\int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty} e^{-\tau_l r^2} r\mathrm{d}r \Rightarrow \frac{\pi}{\tau_l}\left(1 - e^{-\tau_l\lambda_l^2}\right) = \beta\frac{\pi}{\tau_l} \Rightarrow \lambda_l = \sqrt{\frac{-\ln(1-\beta)}{\tau_l}} \tag{35}$$

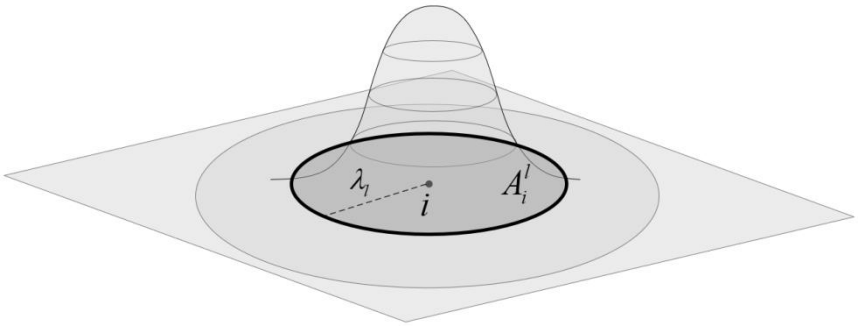


Fig. 5 Affected circle A_i^l and the determination of its radius

Let $G(A_i^l)$ be the exposed population and risk coefficient r_l be the product of accident probability and economic loss per capita of hazmat type l, then the inventory risk at node i can be calculated as follows:

$$r_l G(A_i^l) = r_l \iint_{A_i^l} \rho(x,y)\mathrm{d}x\mathrm{d}y = r_l \pi \lambda_l^2 \rho_{A_i^l} \approx \frac{r_l \pi \lambda_l^2}{\hat{a}} \sum_{a=1}^{\hat{a}} \rho(x_a, y_a) \quad ,(x_a, y_a) \in A_i^l \tag{36}$$

3.5.3 Path(transportation) risk

The movement of hazardous materials would impose risk on the surroundings along a route. As shown in Fig. 6, when the hazmat is moved along the arc (i,j), the potentially affected region forms a ribbon, which is the combination of two semicircles and a rectangle (Batta and Chiu, 1988).

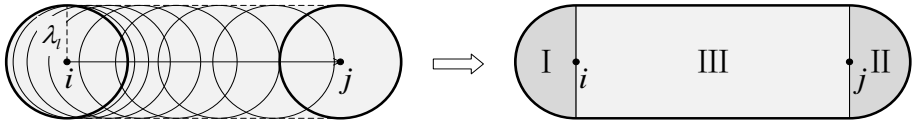
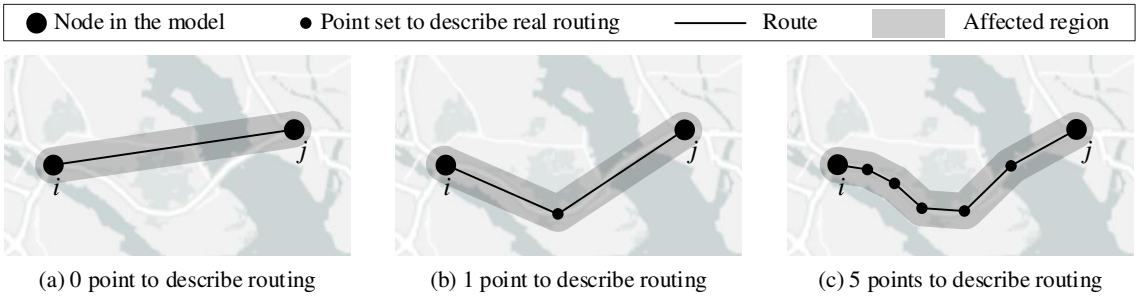


Fig. 6 The form of an affected ribbon

By integrating the density function in the affected region $\tilde{A}_{i,j}^l$, i.e., the ribbon, the exposed population $G(\tilde{A}_{i,j}^l)$ can be approximated as the average population density, $\bar{\rho}_{\tilde{A}_{i,j}^l}$, of region $\tilde{A}_{i,j}^l$, multiplied by its corresponding area $S(\tilde{A}_{i,j}^l)$.

$$G(\tilde{A}_{i,j}^l) = \iint_{\tilde{A}_{i,j}^l} \rho(x,y)\mathrm{d}x\mathrm{d}y = \bar{\rho}_{\tilde{A}_{i,j}^l} S(\tilde{A}_{i,j}^l) = \bar{\rho}_{\tilde{A}_{i,j}^l} [S(\mathrm{I}) + (\mathrm{II}) + (\mathrm{III})] \tag{37}$$

In a real situation, the path or road is often not straight. To facilitate the calculation of exposed population between two visited points, a curved road can be approximated as a set of nodes and links, where the nodes can be the road turnings, as shown in Fig. 7. Naturally, a larger number of inserted nodes indicate a higher accuracy of road network extraction at the expense of greater computation burden. Nevertheless, this can be computed in a pre-processing phase before running the model.



(a) 0 point to describe routing (b) 1 point to describe routing (c) 5 points to describe routing

Fig. 7 Schematic of the road network extraction

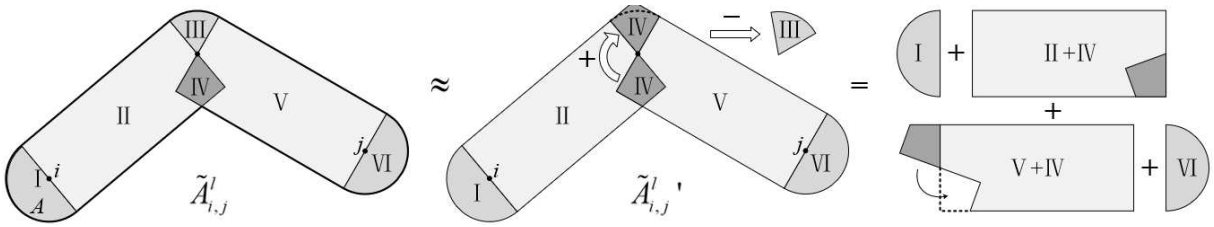


Fig. 8 The integration of exposed ribbon

With this in mind, we proceed to discuss the calculation of the population along a curved road. As shown in Fig. 8, suppose that the road has only a turning between node i and node j, then the affected ribbon can be partitioned into six mutually disjoint regions, including two semicircles at the nodes (I and VI), two angle-deficient rectangles in between (II and V), a quadrilateral (IV) and a sector (III) at the junction. If we remove the sector and add a quadrilateral, the affected ribbon can be approximated as the combination of two semicircles and two rectangles.

$$G\tilde{A}_{i,j}^{l} = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI} \approx \tilde{A}_{i,j}^{l}{}' = \mathrm{I} + (\mathrm{II} + \mathrm{IV}) + (\mathrm{IV} + \mathrm{V}) + \mathrm{VI} \tag{38}$$

In this vein, a curved road with a number of turning points can be simplified into a combination of two semicircles and a couple of rectangles. As a result, the total exposed population $G(\tilde{A}_{i,j}^{l})$ along a curved road can be obtained by summing up the area of each region multiplied by the corresponding population density (or its sample average), then the transportation risk is:

$$r_l G(\tilde{A}_{i,j}^{l}) = r_l \iint_{\tilde{A}_{i,j}^{l}} \rho(x,y)\mathrm{d}x\mathrm{d}y \approx r_l S(\tilde{A}_{i,j}^{l}{}') \frac{1}{\hat{a}} \sum_{a=1}^{\hat{a}} \rho(x_a, y_a) \quad ,(x_a, y_a) \in \tilde{A}_{i,j}^{l}{}' \tag{39}$$

3.5.4 Systematic risk per unit time

The purpose of this section is to calculate the expected risk involved in the hazmat distribution network. According to the average inventory level derived from Section 3.3, and the inventory risk (Eq. (36)), the network-wide total inventory risk can be calculated.

Since the average inventory level for a retailer is μ_n, the inventory risk for customers per unit time equals $\sum_{l\in L}\sum_{n\in N} type_n^l \mu_n r_l G(A_n^l)$. Similarly, the average inventory level at the distribution center m is its safety stock plus the customers' demand, thus the inventory risk for distribution centers per unit time is $\sum_{l\in L}\sum_{m\in M} type_m^l r_l G(A_m^l)\left(\sum_{k\in K}\sum_{n\in N} X_{m,k}X_{n,k}\,\mu_n + St_m\right)$. As a result, the overall inventory risk is the summation of those of customers and distribution centers.

$$RM = \sum_{l\in L}\sum_{n\epsilon N} type_n^l \mu_n r_l G(A_n^l) + \sum_{l\in L}\sum_{m\in M} type_m^l r_l G(A_m^l)\left(\sum_{k\in K}\sum_{n\in N} X_{m,k}X_{n,k}\mu_n + St_m\right) \tag{40}$$

The average risk on the segment between node i and j is associated with the arc-specific vehicle load q_j^k, exposed population $G(\tilde{A}_{i,j}^{l})$, and the order cycle $\tilde{P}_k$.

$$rs_{i,j,k}^{l} = \sum_{i\in M\cup N}\sum_{j\in M\cup N} Y_{i,j,k} \frac{G(\tilde{A}_{i,j}^{l}) r_l}{\tilde{P}_k} q_j^k \tag{41}$$

Despite their in-transit and storage incompatibility, the superimposed risks among different materials, which results from possible chemical reaction once accidents (e.g., leakage, explosion) happen, further complicate the comprehensive plans. Specifically, when a location is within the affected range of multiple hazmat classes at the same time, the combined impact of the hazmat incident cannot be a "simple sum". For example, some chemical energy that can produce a chemical reaction would result in a consequence much greater than the sum. To capture the compound effect, we introduce a new variable $\delta_{ll'}$ that represents the

risk superposition coefficient between hazmat type l and l'. When the route k (of type l) overlaps route k' (of type l') at a specific segment, there will be an additional superimposed risk $\delta_{ll'}\sqrt{rs_{i,j,k}^{l}rs_{i,j,k'}^{l'}}$ on the segment. $\delta_{ll'} > 0$ indicates that the coexistence of l and l' will cause the additional superimposed risk due to chemical reaction; $\delta_{ll'} = 0$ indicates that l and l' can coexist without any chemical reaction (or they are the same type of hazmat, i.e., $l = l'$). $\delta_{ll'} < 0$ indicates that the coexistence of l and l' can reduce the total risk. In practice, the value of $\delta_{ll'}$ can be tuned by fine-tuned or estimated by rule-of-thumb by chemical engineers. Then, the additional total superimposed risk SR can be calculated as the product of $\delta_{ll'}\sqrt{rs_{i,j,k}^{l}rs_{i,j,k'}^{l'}}$, routing decision variables and hazmat type parameters of each segment of each route. Since every two segments are multiplied twice, the additional total superimposed risk SR is the sum of all products divided by 2.

$$SR = \frac{1}{2}\sum_{i\in M\cup N}\sum_{j\in M\cup N}\sum_{l\in L}\sum_{m\in M}\sum_{k\in K}\sum_{l'\in L}\sum_{m'\in M}\sum_{k'\in K} Y_{i,j,k}\delta_{ll'}\sqrt{type_m^l X_{m,k} rs_{i,j,k}^l \times type_{m'}^{l'} Y_{m',k'} rs_{i,j,k'}^{l'}} \tag{42}$$

As a result, the total distribution risk per unit time takes the following form, where the average risk change Δr_k^l caused by the detour of route k per unit time is calculated by the same way of Eq. (24).

$$RD = \sum_{l\in L}\sum_{m\in M}\sum_{k\in K} type_m^l X_{m,k}\left(\Delta r_k^l + \sum_{i\in m\cup N}\sum_{j\in m\cup N} rs_{i,j,k}^l\right) + SR \tag{43}$$

Given the order quantity of the distribution center m (Eq. (5)), the transportation risk from the plants to distribution centers per unit time can be determined as:

$$RR = \sum_{l\in L}\sum_{m\in M} type_m^l \frac{G(\tilde{A}_{i,j}^l)r_l}{P_m}\sqrt{\frac{2F_m}{h_m}\sum_{n\in N}\sum_{k\in K} X_{m,k}X_{n,k}\mu_n} \tag{43}$$

3.6 Spatio-temporal risk distribution

In the hazmat logistics system, the inventory level of distribution centers and demand points, and the vehicle load over the network will change over time. Due to the radiation and compounded effects of the hazmat risks, excessive risk threats may be imposed to a specific location at a specific time in the operation. This point could be either a node in the network or any location outside the road network. Therefore, it is imperative to deduce the spatio-temporal risk distribution and introduce the spatio-temporal risk constraint in the optimization model (see constraints (71)). The inventory risk depends on the inventory volume and the population exposure. With the decision variables and mathematical formulations, the location and load of vehicles, and the inventory level over time can be calculated. Based on this, the risk value can be derived as a function of the location and time, from which the spatio-temporal risk distribution is obtained.

3.6.1 Inventory risk distribution for customers

The customer n is served once within each order cycle $\tilde{P}_k$, and the corresponding demand is $D_n\tilde{P}_k$. The travel time from the distribution center to customer n on route k is the travel time t_n^k (Eq. (12))

plus the time change caused by detours between each pair of visited points before this customer Δt_n^k (Eq. (25)).

In other words, the vehicle arrival time at the customer n that departs from the distribution center at time is $t_n^k + \Delta t_n^k$. The customer n will be served every order cycle $\tilde{P}_k$ after the first delivery, and the corresponding delivery quantity is $D_n \tilde{P}_k$. Therefore, given any time t, the difference between this time point and that of preceding delivery equals the remainder of the difference between this time point and that of first delivery after dividing the order cycle:

$$\mathrm{mod}\left(t - t_n^k - \Delta t_n^k, \tilde{P}_k\right) \tag{45}$$

where mod is the modulus operator.

The demand during the time interval equals the product of the consumption rate and the interval:

$$D_n \mathrm{mod}\left[t - \sum_{k \in K} X_{n,k}(t_n^k + \Delta t_n^k), \tilde{P}_k\right] \tag{46}$$

As a result, the inventory level for customer n at time t is the delivery quantity minus the consumption in the current cycle:

$$I_{l,n}^{(1)}(t) = type_n^l D_n \left\{\tilde{P}_k - \mathrm{mod}\left[t - \sum_{k \in K} X_{n,k}(t_n^k + \Delta t_n^k), \tilde{P}_k\right]\right\} \tag{47}$$

Taken together, the risk of any coordinate (x, y) affected by the inventory of customer n at time t can be expressed as follows:

$$R_{l,n}^{(1)}(t, x, y) = r_l \rho(x, y) I_{l,n}^{(1)}(t) e^{-\tau_l[(x-x_n)^2 + (y-y_n)^2]} \tag{48}$$

3.6.2 Inventory risk distribution for distribution centers

The distribution center m will be replenished from the plants every order cycle P_m, and the average replenishment quantity is Q_m. On the other hand, the customers will be replenished from the distribution center m through distribution route k. At time t, the replenished quantity Q_m delivered to the distribution center m at the beginning of each order cycle has been consumed for a time of $\mathrm{mod}(t, P_m)$, thus the consumption of distribution center m at time t is:

$$Q_m \frac{\mathrm{mod}(t, P_m)}{P_m} \tag{49}$$

Considering the safety stock St_m, the inventory level for distribution center m at time t can be calculated as the sum of safety stock and replenished quantity, minus the consumption in the current cycle.

$$I_{l,m}^{(2)}(t) = W_m type_m^l \left\{St_m + Q_m \left[1 - \frac{\mathrm{mod}(t, P_m)}{P_m}\right]\right\} \tag{50}$$

Therefore, the risk of any coordinate (x, y) affected by the inventory of distribution center m takes the following form:

$$R_{l,m}^{(2)}(t, x, y) = r_l \rho(x, y) I_{l,m}^{(2)}(t) e^{-\tau_l[(x-x_m)^2 + (y-y_m)^2]} \tag{51}$$

3.6.3 *Transportation risk distribution from distribution centers to customers*

The en-route vehicles will bring about transportation risks. The risk level is associated with the number of en-route vehicles, which vary with the difference between the order cycle $\tilde{P}_k$ and travel time of a distribution route $(t_m^k + \Delta t_m^k)$, and could be 0 when $\tilde{P}_k > t_m^k + \Delta t_m^k \quad , X_{m,k} = 1$.

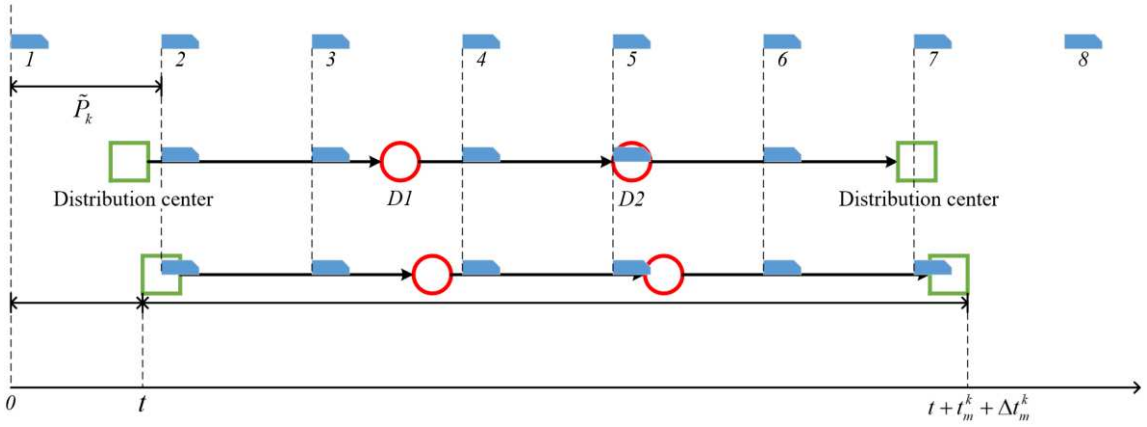


Fig. 9 The number of vehicles on a distribution route

As shown in Fig. 9, let the set of vehicles driving on route k at time t be $\tilde{V}_k^t = \{1,2,\dots,\lfloor (t + t_m^k + \Delta t_m^k)/\tilde{P}_k \rfloor - \lfloor t/\tilde{P}_k \rfloor\}$, then the accumulated travel time of v^{th} vehicle in this set is $\mathrm{mod}(t, \tilde{P}_k) + (v-1)\tilde{P}_k \quad , v \in \tilde{V}_k^t$. Based on this, we introduce the following state variable:

$$\Theta(k,i,j,v,t) = \begin{cases} 1, & [\mathrm{mod}(t,\tilde{P}_k) + (v-1)\tilde{P}_k - t_i^k][\mathrm{mod}(t,\tilde{P}_k) + (v-1)\tilde{P}_k - t_j^k] \le 0 \\ 0, & [\mathrm{mod}(t,\tilde{P}_k) + (v-1)\tilde{P}_k - t_i^k][\mathrm{mod}(t,\tilde{P}_k) + (v-1)\tilde{P}_k - t_j^k] > 0 \end{cases} \quad , v \in \tilde{V}_k^t \quad (52)$$

where $\Theta = 1$ indicates that v^{th} vehicle is running between node i and node j on route k at time t.

According to Eq. (27), the load of v^{th} vehicle at time t is:

$$I_{l,k,v}^{(3)}(t) = \sum_{m \in M} type_m^l X_{m,k} \sum_{i \in M \cup N} \sum_{j \in M \cup N} Y_{i,j,k} \Theta(k,i,j,v,t) q_j^k \quad (53)$$

The coordinate of each vehicle in the set of $\tilde{V}_k^t$ can be calculated based on the current time t and the coordinate of node i,j, thereafter the set of vehicle coordinates $\widetilde{Co}_k^t$ can be built by mapping from $\tilde{V}_k^t$. More specifically, the segment length can be known based on the coordinate of node i,j. The ratios between the mileages that vehicles have traveled in this segment and the total length of this segment at time t can be calculated. Then, the coordinate of each vehicle in the set of $\tilde{V}_k^t$ can be calculated given the node coordinates. Consequently, the risk of any coordinate (x,y) affected by vehicles on the route k at time t can be written as follows:

$$R_{l,k}^{(3)}(t,x,y) = \sum_{v \in \tilde{V}_k^t} r_l \rho(x,y) I_{l,k,v}^{(3)}(t) e^{-\tau_l[(x-x_v)^2 + (y-y_v)^2]} \quad (54)$$

where (x_v, y_v) denotes the v^{th} element of the set $\widetilde{Co}_k^t$.

3.6.4 *Transportation risk distribution from plants to distribution centers*

In the transportation from plants to distribution centers, the number of vehicles en-route can be calculated similarly to Section 3.6.3. Let the set of vehicles driving to distribution center m at time t be $V_m^t = \{1,2,\cdots,\lfloor (t + Dis_{l,m})/P_m \rfloor - \lfloor t/P_m \rfloor\}$. At this moment, the load of each vehicle takes the following form. Note that the vehicles running from the plant (corresponding to hazmat type l) to a distribution center m are identical since there will be no unloading process en-route.

$$I_{l,m}^{(4)}(t) = Q_m = W_m type_m^l \sqrt{\frac{2F_m}{h_m} \sum_{n\in N} \sum_{k\in K} X_{m,k} X_{n,k} \mu_n} \tag{55}$$

The coordinate of each vehicle in the set of V_m^t can be calculated based on the current time t and the coordinates of plants and distribution centers, then the set of vehicle coordinates Co_m^t can be built by mapping from V_m^t. Specifically, according to the replenishment period and route length, the ratios between the mileages that vehicles have traveled and the total route length at time t can be calculated. Then, given the node coordinates of plants and distribution centers as well as the associated routes, the coordinate of each vehicle in the set of V_m^t can be calculated. Therefore, the risk of any coordinate (x, y) affected by vehicles on the route from the plant (corresponding to hazmat type l) to distribution center m at time t can be written as follows:

$$R_{l,m}^{(4)}(t,x,y) = \sum_{v\in V_m^t} r_l \rho(x,y) I_{l,m}^{(4)}(t) e^{-\tau_l[(x-x_v)^2+(y-y_v)^2]} \tag{56}$$

where (x_v, y_v) denotes the v^{th} element of the set Co_m^t.

3.6.5 Superimposed risk

Given the additional superimposed risk affected by the possible chemical reactions between every two risks, that is, the product $R_l^{(\ell)} R_{l'}^{(\hbar)} \delta_{ll'}, \ell, \hbar \in \{1,2,3,4\}$ (1~4 corresponding to the four risks in Section 3.6.1~3.6.4), the total superimposed risk of a point (x, y) is the sum of all products divided by 2:

$$RS(t,x,y) = \frac{1}{2} \sum_{l\in L} \sum_{l'\in L} \sum_{\ell=1}^{4} \sum_{\hbar=1}^{4} R_l^{(\ell)}(t,x,y) R_{l'}^{(\hbar)}(t,x,y) \delta_{ll'} \tag{57}$$

3.6.6 The consolidated formulation for transportation risk distribution

By combining the inventory risk with transportation risk, the total risk can be summarized as follows:

$$R(t,x,y) = \sum_{l\in L} \left[\sum_{n\in N} R_{l,n}^{(1)} + \sum_{m\in M} \left(R_{l,m}^{(2)} + R_{l,m}^{(4)} \right) + \sum_{k\in K} R_{l,k}^{(3)} \right] + RS \tag{58}$$

Given the location (x, y) and time t, the risk value can be calculated directly by Eq. (58), from which the location with excessive risk can be identified.

3.6.7 Optimization model

The objective of this paper is two-fold: one aims to find out the optimal location, inventory, and routing schedule simultaneously that minimizes the total cost from the system perspective, while the other aims to minimize the total risk exposing the population of the regional area. This gives rise to a bi-objective

optimization problem. In addition, to prevent any site from being exposed to excessive risks, we enhance the solution approach by additionally incorporating the risk tolerance.

With the costs and risk distribution discussed above, the optimization model for the multi-class hazmat distribution network design problem could be formulated as follows:

$$\min CL + CM + CR + CD \tag{59}$$

$$\min RM + RR + RD \tag{60}$$

s.t.

$$\sum_{m\in M} W_m type_m^l \geq 1 \quad ,\forall l \in L \tag{61}$$

$$\sum_{k\in K} X_{m,k} \begin{cases} = 0 & ,W_m = 0 \\ \geq 1 & ,W_m = 1 \end{cases} \quad ,\forall m \in M \tag{62}$$

$$\sum_{m\in M} X_{m,k} \begin{cases} = 0 & ,\sum_{n\in N} X_{n,k} = 0 \\ = 1 & ,\sum_{n\in N} X_{n,k} \geq 1 \end{cases} \quad ,\forall k \in k \tag{63}$$

$$\sum_{n\in N} X_{n,k} \begin{cases} = 0 & ,\sum_{m\in M} X_{m,k} = 0 \\ \geq 1 & ,\sum_{m\in M} X_{m,k} = 1 \end{cases} \quad ,\forall k \in k \tag{64}$$

$$\sum_{k\in K} X_{n,k} = 1 \quad ,\forall n \in N \tag{65}$$

$$\sum_{j\in M\cup N, j\neq i} Y_{i,j,k} = \sum_{j\in M\cup N, j\neq i} Y_{j,i,k} = X_{i,k} \quad ,\forall i \in M \cup N, \forall k \in K \tag{66}$$

$$\sum_{l\in L} \prod_{i\in M\cup N} type_i^l = \sum_{m\in M} X_{m,k} \quad ,\forall k \in K \tag{67}$$

$$\sum_{l\in L} type_m^l Q_m + St_m \leq S_m \quad ,\forall m \in M \tag{68}$$

$$\sum_{l\in L}\sum_{k\in K} type_n^l X_{n,k} \tilde{P}_k D_n \leq S_n \quad ,\forall n \in N \tag{69}$$

$$\sum_{m\in M} type_m^l X_{m,k} \tilde{Q}_k \leq S^l \quad ,\forall l \in L, \forall k \in K \tag{70}$$

$$R(t,x,y) \leq R_{max} \quad ,\forall t,x,y \tag{71}$$

$$W_m \in \{0,1\} \quad ,\forall m \in M \tag{72}$$

$$Y_{i,j,k} \in \{0,1\} \quad ,\forall i,j \in M \cup N, \forall k \in K \tag{73}$$

$$X_{i,k} = \begin{cases} 0 & , \sum_{j \in M \cup N} Y_{i,j,k} = 0 \\ 1 & , \sum_{j \in M \cup N} Y_{i,j,k} = 1 \end{cases} \quad , \forall i \in M \cup N, \forall k \in K \tag{74}$$

The first objective function Eq. (59) represents the summation of the opening costs, inventory cost of distribution centers and customers, and transportation cost of distribution and replenishment per unit time (see the derivation in Sections 3.3.3, 3.3.4, 3.4.3, and 3.4.4). The second objective function Eq. (60) represents the summation of inventory risk of distribution centers and customers and transportation risk of distribution and replenishment per unit time (see the derivation in Section 3.5.4). Constraints (61) ensure that at least one distribution center is opened for each type of hazmat. Constraints (62) ensure that an opened distribution center has at least one distribution route. Constraints (63) ensure that there is only one distribution center on each route. Constraints (64) ensure that there will be customers on each distribution route. Constraints (65) ensure that each customer is served. Constraints (66) ensure that each customer is allowed to be served by at most one distribution route, with only one delivery per order cycle, and that each distribution route is closed-loop. Constraints (67) ensure that for each route the hazmat type is identical for both the distribution center and customers. Constraints (68) ensure that the inventory level of each distribution center should not exceed its maximum storage capacity. Constraints (69) ensure that the inventory level of each customer should not exceed its maximum storage capacity. Constraints (70) represent the vehicle capacity constraint. Constraints (71) ensure that the spatio-temporal risk should not exceed a specific threshold. Constraints (72) and (73) indicate the attribute of the decision variables. Constraints (74) state the relationship between decision variables.

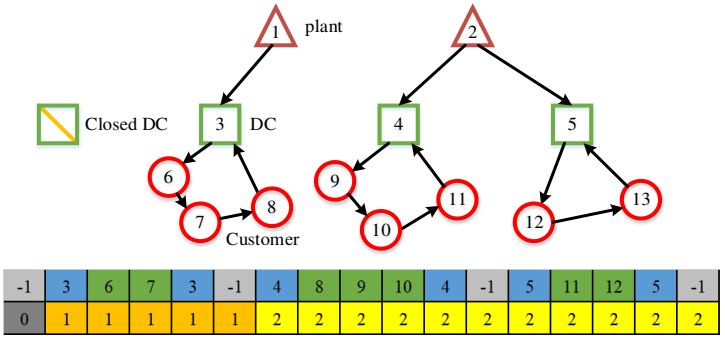
4 Solution method

MHND is an NP-hard issue as it concerns location allocation and inventory and route planning. The objective functions are nonconvex and discontinuous (i.e., nonlinear and non-smooth), and the constraints (e.g., constraints (71)) are difficult to be transformed into linear conditions. Since the decision variables are binary, the model optimization is a nonlinear integer programming problem, and the problem is not solvable by commercial solvers with exact solutions such as CPLEX and GUROBI. Besides, exact algorithms and solvers can only solve NP-hard problems for a limited size in a reasonable computational time. In comparison, metaheuristic algorithms can solve large-size problem instances within a reasonable computation time, which are often applied in addressing various combinatorial optimization problems. The NSGA-II (Non-dominated sorting Genetic Algorithm-II) algorithm is a fast non-dominated sorting genetic algorithm with an elite mechanism (Deb et al., 2002), which has been adopted in a variety of combinational optimization problems. To produce high-quality solutions, we propose a metaheuristic called knowledge-based NSGA-II algorithm with cyclic dissimilarity-based elitist selection operator (NSGA-II-CD) to solve the multi-objective optimization model and search a near-optimal solution, where the key components are redesigned based on the domain knowledge to obtain an efficient Pareto frontier for

minimizing the total cost and minimizing the total risk. In particular, we devise a cyclic dissimilarity-based elitist selection operator to tackle the issue of speeding proliferation, which greatly improves the solution quality. Then we create the benchmark networks and multiple artificial problem instances to test the proposed algorithm.

4.1 Encoding of chromosome

Representing the chromosome is not a trivial task in this problem. An effective coding scheme is devised to endow the chromosome with information about the location of distribution centers, hazmat types, and routing schedule. Fig. 10 shows a sample of the encoding procedure. The individual is a two-row matrix. The first row stands for the index of nodes, and the second row stands for the hazmat type of the corresponding node. The visited nodes of each distribution route are fed into the matrix, and different routes are separated by a separation mark (SM). The distribution center in the matrix is opened. Since each route has two identical DC nodes (starting and ending) and an SM, the length of matrices may be distinct for different numbers of routes. For example, the length of a matrix with two customers served by one route is 6: SM-DC-customer 1-customer 2-DC-SM. Alternately, the length of a matrix with two customers separately served by two routes is 9: SM-DC1-customer 1-DC1-SM-DC2-customer 2-DC2-SM.



-1	3	6	7	3	-1	4	8	9	10	4	-1	5	11	12	5	-1
0	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2

Fig. 10 Illustration of coding solution (Individual A, DC 3, 4, 5 are opened with a route respectively)

Fig. 10 illustrates an example with 2 hazmat types, 2 plants, 4 candidate DCs, and 7 customers. To begin with, an SM-1 is generated, and the corresponding index of hazmat type is set as 0, which indicates no hazmat type information. The visited nodes of route 1 (3-6-7-3) are filled into row 1, and the corresponding type number (1) is filled into row 2. Then, the next SM (-1) is generated, of which the type number is determined by the previous distribution route. The nodes and SM of route 2 (4-8-9-10-4-(-1)) are filled into row 1, and the type number (2) is filled into row 2. The nodes and SM of route 3 (4-8-9-10-4-(-1)), and the corresponding type number (2) are processed similarly.

4.2 Initial solution generation

We propose a constructive heuristic to generate an initial feasible solution. The initial population is generated by randomly assigning customers to the candidate distribution centers with identical hazmat types. The distribution center is opened once a customer is assigned, otherwise, it is closed. The customer is served only by one distribution route by default, and the delivery order is randomly determined. The feasible solution of an individual can be generated according to Section 4.1. This process can guarantee to produce a feasible solution in each run. According to the aforementioned coding solution and available

nodes in Fig. 10, Individual B is generated (Fig. 11), forming a simple population with Individual A. An initial population of any size can be obtained by repeating the above operations several times.

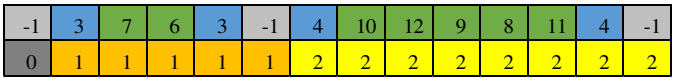


Fig. 11 Individual B (DC 3,4 are opened, having a route respectively)

4.3 Crossover

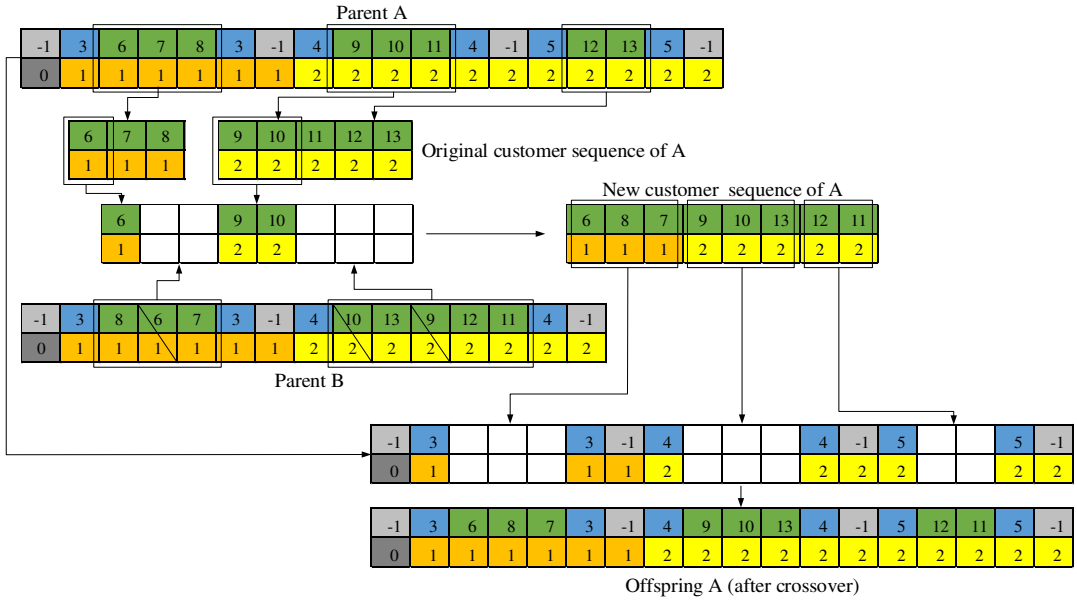


Fig. 12 Crossover operator

Due to the provision of various hazmat types, in the crossover the information exchange should be performed separately for each hazmat type. Since each route has two identical DC nodes (starting and ending) and an SM, the length of matrices may be distinct for different numbers of routes. Since the number of distribution routes may be different among individuals given a number of customers, the crossover operation should be used to change the customer order. Taking Fig. 12 as an example, the main steps are as follows:

Step 1: All customers are extracted from an individual to form a customer sequence.

Step 2: For each hazmat type, retain only half of the customers and remove the remaining ones, and reserve the slots left after the removal.

Step 3: Refill the customers removed before into the slots, according to the order they were arranged in another individual. In this way, a complete customer sequence of the next generation can be produced.

Step 4: Refill the new customer sequence into the individual matrix to produce a new generation.

As a result, a new population is formed by randomly pairing the individuals, and conducting crossover operations according to the above steps.

4.4 Mutation

To realize the change of the number of opened distribution centers, the number of routes for each distribution center, and the visiting order of each route for each hazmat type in the mutation, in this section we propose four mutation procedures. These four mutation methods are independent with respective

execution probability. According to these probabilities, each individual will perform mutation operations with different means and degrees. A new population can be obtained after each individual performs each mutation operation with a given probability. The diagram is shown in Fig. 13. The main steps are described as follows:

Fig. 13 Mutation operator

Mutation 1: Randomly select two customers from a route and exchange their positions.

Mutation 2: Randomly select a customer from a route and insert it into another route.

Mutation 3: Randomly delete a route and randomly assign its customers to other routes with the same hazmat type.

Mutation 4: Generate a new route by randomly drawing customers from other routers with the same hazmat type. If an original route has no customer after mutation, delete it.

4.5 Selection improvement and evaluation

In this section, we propose an improved selection strategy for the NSGA-II algorithm. To this end, we first propose a method to create the benchmark hazmat supply network for comparison, and evaluate the proposed strategy through instances of different sizes.

4.5.1 Instance generation method

The instance generation method includes the parameter settings and the supply network generation, while the latter is more critical. The proposed road network generation method can generate nodes and links with specific node attributes, hazmat type, and other relevant parameters for links, while satisfying the number and density of hazmat plants, candidate DCs, and customers. The main steps are described as follows:

Step 1: Specify the upper and lower bounds of the link length and the total number of nodes. Generate a node and a circle centered on this node with an upper bound as its radius. Randomly generate several nodes in this circle as the nodes of the second round. Then, draw the union of the circles centered on the nodes of the last round, and generate the nodes of the next round in it. Fig. 14 shows the example.

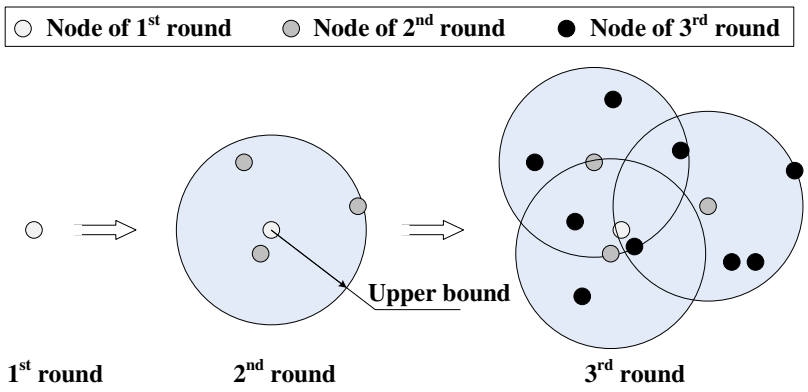
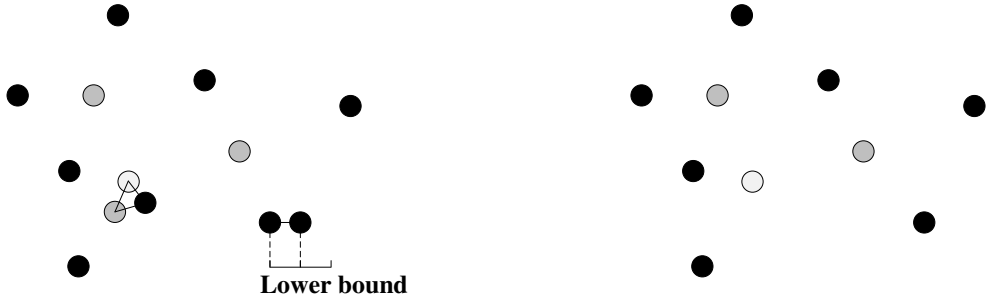


Fig. 14 Node generation based on the upper bound

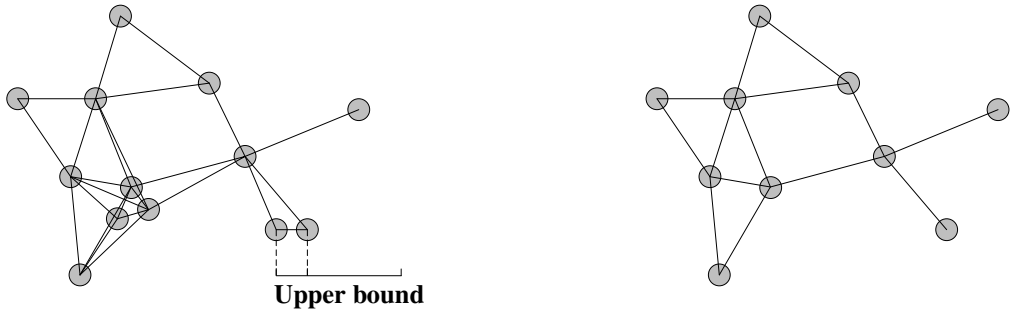
Step 2: Remove the nodes between which the distance is less than the lower bound. Note that this step is not necessarily compulsory. This step can avoid a high concentration of risks caused by the high density of nodes and links, whereas the absence of a lower limit enables the simulation of real-life risk accumulation such as industrial parks. Fig. 15 shows the example.



(a) Distance check (b) Too-close nodes removal

Fig. 15 Node removal based on the lower bound

Step 3: Generate a link between two nodes when their distance is not greater than the given upper bound. Fig. 16 shows the example.



(a) Road network without node removal (b) Road network after node removal

Fig. 16 Road network generation based on the upper bound

Suppose there are two hazmat types, three candidate DCs for hazmat type 1, and two candidate DCs for hazmat type 2. Choose a node that connects to three nodes as the plant of hazmat type 1, while the three nodes are treated as the candidate DCs. Similarly, choose another node that connects to the two nodes as the plant of hazmat type 2, while the two nodes are treated as the candidate DCs. The remaining nodes are treated as the customers. Randomly assigning the hazmat types to customers completes the road network

generation, as illustrated in Fig. 17. Fig. 18 presents the overall solution framework for network generation.

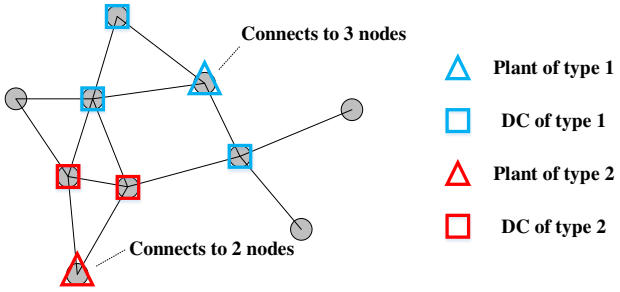


Fig. 17 Illustration of road network generation

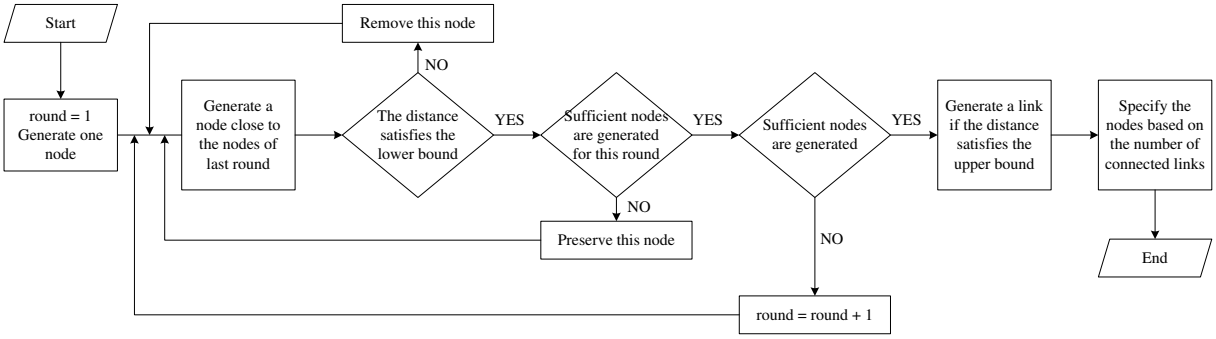


Fig. 18 Overall solution framework for network generation

4.5.2 The phenomenon of speeding proliferation

According to the road network generation method proposed in Section 4.5.1, a road network is generated to evaluate the algorithm. Fig. 19 depicts the generated network with three hazmat types, two candidate DCs for each hazmat type, and a total of 30 customers.

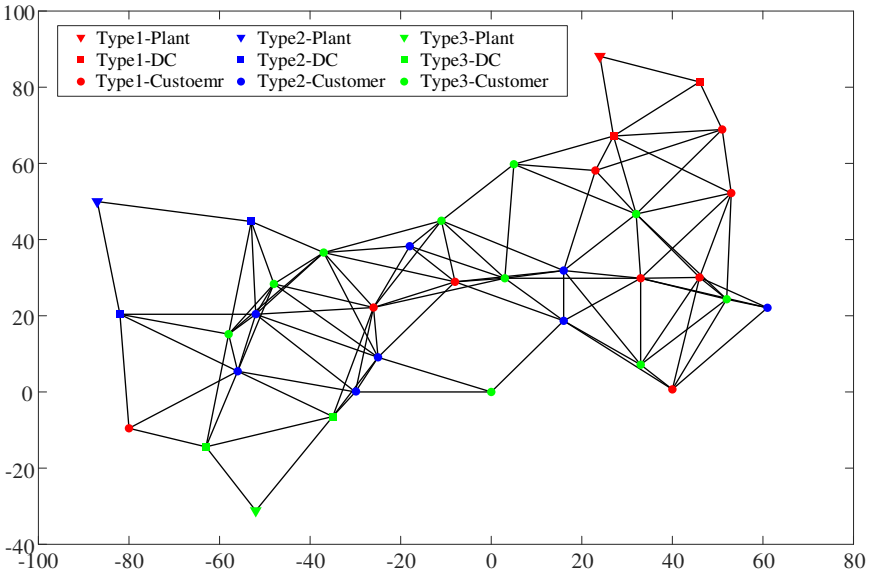
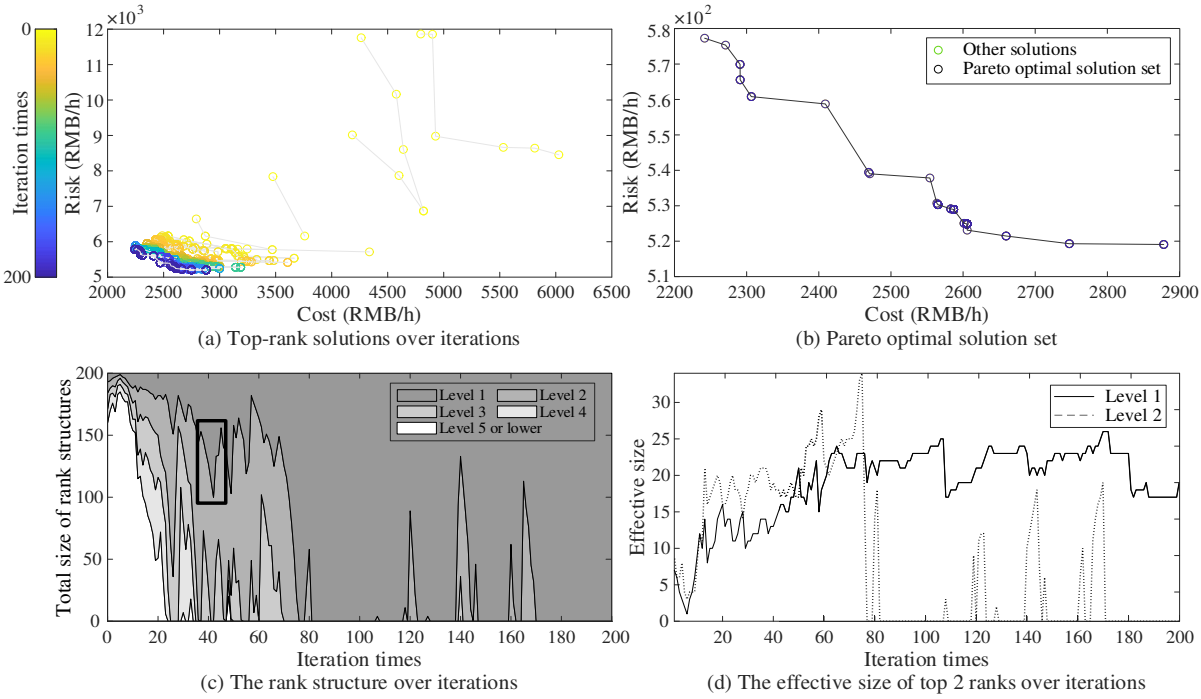


Fig. 19 Road network of the example



(a) Top-rank solutions over iterations

(b) Pareto optimal solution set

(c) The rank structure over iterations

(d) The effective size of top 2 ranks over iterations

Fig. 20 The iteration process and Pareto optimal solutions under classic NSGA-II

Based on the instance generation method, a medium-size road network is generated to evaluate the algorithm. The network consists of three hazmat types, two candidate DCs for each hazmat type, and totally 30 customers. Fig. 20(c) shows the total size (i.e., the total number of individuals) of non-dominated rank structures over iterations. 19 rapid proliferation, shaping like the steep slope in the box area of Fig. 20(c), takes place among the top-rank individuals. For the first 12 times, each rise of the size is followed by an immediate decline, while afterward the top-rank individuals occupy the entire population and fail to recover. However, the effective size (i.e., the number of distinct individuals) of the top-rank individuals in Fig. 20(d) has no significant change when the top-rank individuals proliferate abruptly. This indicates that a large number of top-rank individuals are trivial. In other words, the entire population becomes the replicas of only a couple of individuals. Instead of generating more new feasible solutions, the existence of these replicates will further increase the size of top-rank solutions over iterations, which does not make sense. Motivated by this observation, we term this phenomenon as “speeding proliferation”, a proliferative phenomenon that exceeds the normal scale of evolution.

Fig. 21 shows the mechanism of speeding proliferation. Suppose that the population size is 5. Both the parents and the first generation (F1) have an identical non-dominated solution, namely “individual 1”. Since individual 1 exists in both the parents and the F1, these individuals will be reserved in the second generation (F2) by the selection operator with the elitism strategy, which doubles the number of individual 1 in the new population. More replicas of individual 1 indicate a greater occurrence probability of itself in the next generation due to their self-crossover not being caught by possible mutation, which in turn increases the number of individual 1 after the next election. This leads to a vicious circle that intensifies the speeding proliferation of individual 1 and destroys the population diversity, ending up with only a couple of replicas that occupy the entire population. Due to the limitation of population size, this may also result in

the spillover phenomenon where some top-rank individuals cannot be selected, as represented by the election process from F2 to F4.

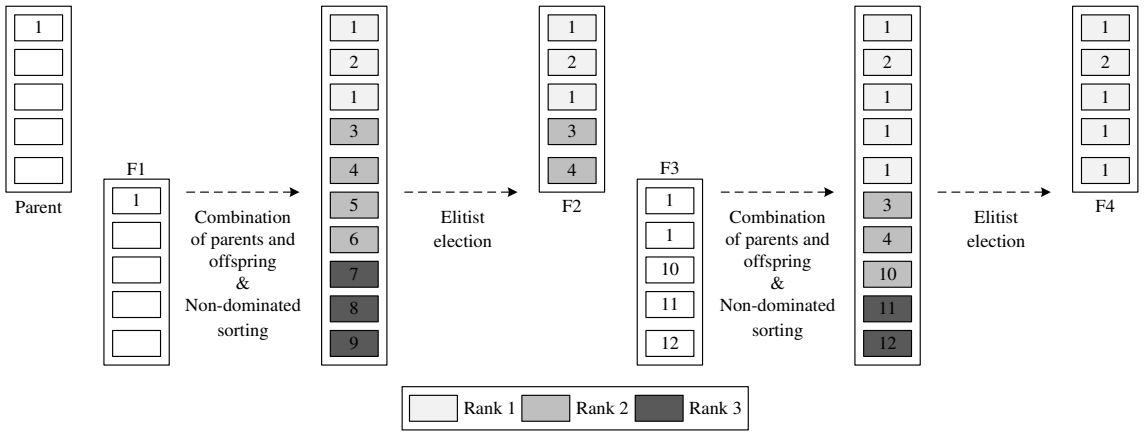


Fig. 21 Diagram of speeding proliferation mechanism

One way to hinder this proliferative process and rapidly normalize the population structure is the occurrence of a new individual that can dominate individual 1, which corresponds to the recovery of the speeding proliferation in Fig. 20(c). However, the increasing replicas in the population will decrease the diversity, narrow the search space, and reduce the occurrence probability of better individuals. Therefore, the issue of speeding proliferation will be usually inevitable and unstoppable once the same top-rank individuals in the parents and offspring emerge.

The speeding proliferation is a typical premature phenomenon of the NSGA-II algorithm. The elitism selection operator, which combines parents and offspring and sorts them in layers, can preserve the best individuals but may exponentially duplicate the local optimal individuals, quickly destroy the diversity, and therefore reduce the efficiency and the solution quality. This phenomenon often appears in integer programming like the MHND in this research. With a continuous domain and countless feasible solutions, the non-integer programming can have infinite solutions, such that the same individual rarely exists between the two generations after crossover and mutation. On the contrary, the definition domain of integer programming is discrete, and there may be only limited feasible solutions after considering constraints. In addition, in the study of real-world problems, the coding, decoding of individuals, and particularly the initialization often cannot guarantee the complete coverage of potential solution space. This further reduces the total number of feasible solutions and the same individual frequently results, thereby exacerbating the speeding proliferation.

4.5.3 Tackling the issue of speeding proliferation: Cyclic dissimilarity-based elitist selection (CD)

To address the issue of speeding proliferation, this paper proposes a new generic CD strategy, which is an abbreviation of "cyclic dissimilarity-based elitist selection". The principle of this strategy is that a cyclic selection strategy based on dissimilarity is added to the original elitist selection strategy to diversify the population to the greatest extent, such that the speeding proliferation is completely avoided, and the iterative efficiency and solution quality are improved.

The main process of CD is described as follows:

Step 1: Combine parents with offspring and perform non-dominated sorting.

Step 2: Extract the individuals that are different from each other from the population, and construct a set with the same order as Step 1.

Step 3: If the size of the set is now smaller than the required population size, duplicate this set and arrange it after the previous set. This process is repeated until the required size of the set reaches.

Step 4: Extract a certain number of individuals ranking in the front from the set in Step 3 to generate the next generation.

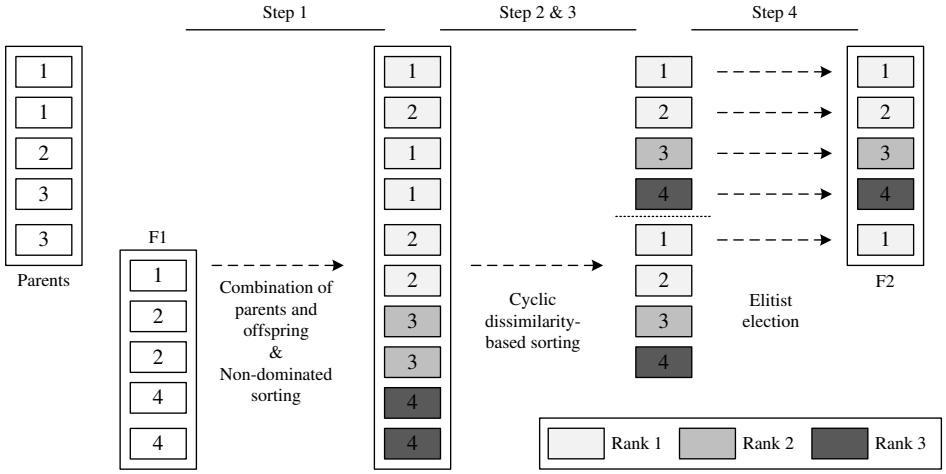


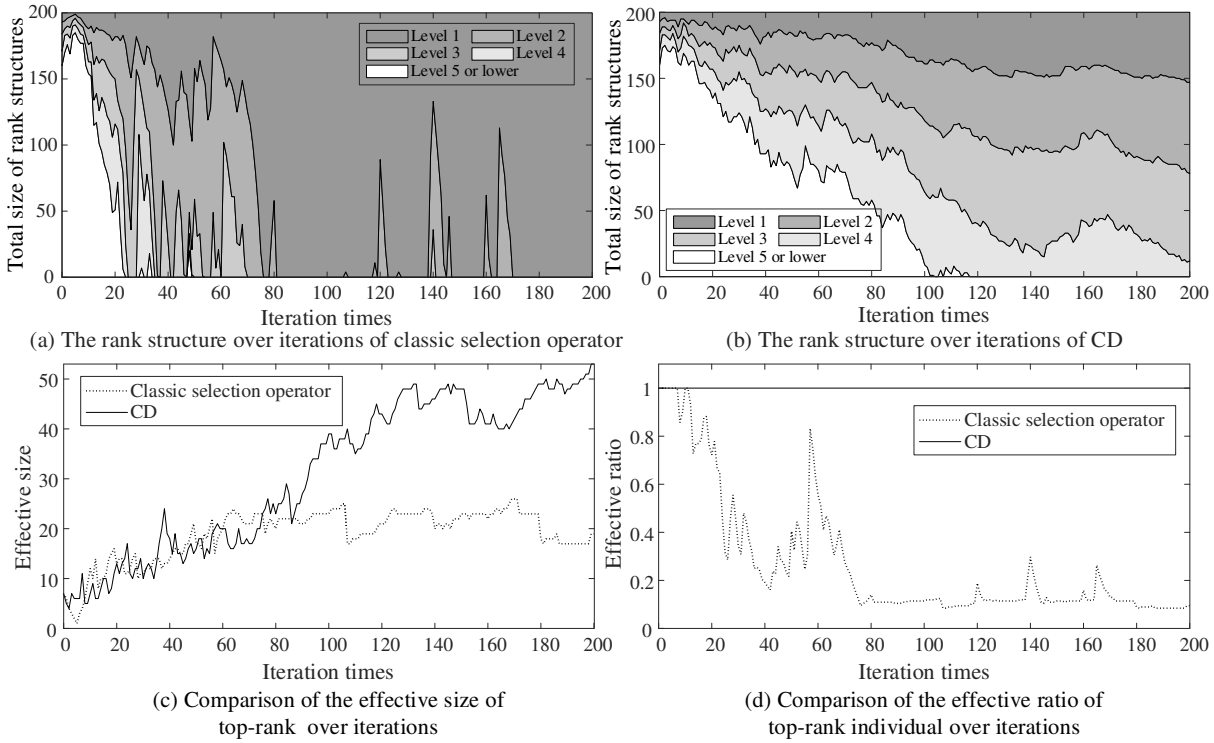
Fig. 22 Diagram of the CD process

Fig. 22 diagrams the CD process. As we can see, the CD can retain the diversity of the population to a great extent even if a number of identical individuals exist in the parent and offspring. Commendably, CD not only limits the replicas in the population to prevent speeding proliferation but also enables the high-rank individuals to enter the next generation with a greater probability, such that high-quality individuals are more likely to generate and better iterative efficiency and solution can be expected. It is worth mentioning that CD is more suitable for sophisticated practical problems. For example, in the well-known traveling salesman problem (TSP), the optimization can be realized only by visited node reordering, and the classical operator can obtain the optimal solution by neighborhood search from a few local optimal solutions. In this case, the performance of the CD may be not quite outstanding. However, CD takes more effect in complex problems like the MHND where the decision variables include location and routing options. Unlike the TSP, there may be some outstanding solutions in the neighborhood of an inferior solution, and the global optimization is difficult to achieve only by neighborhood search from a few local optimal solutions.

4.5.4 Algorithmic comparison under different sizes

In this section, we investigate the merit of the CD strategy by examining its performance in different instances. To this end, we implement the traditional selection operator and compare the performance of these two versions using the exemplified network in Section 4.5.2, and the results are shown in Fig. 23. Figs. 23(a) and (b) show the outcome of the rank structures for the two operators. We can see that the CD is effective in preventing speeding proliferation, and enables the size of the top rank to maintain a reasonable

and stable growth throughout the whole process, ending up with 52 Pareto optimal solutions and with an effective ratio of 1 (Fig. 23(d)).



(a) The rank structure over iterations of classic selection operator

(b) The rank structure over iterations of CD

(c) Comparison of the effective size of top-rank over iterations

(d) Comparison of the effective ratio of top-rank individual over iterations

Fig. 23 Comparison between classic selection operator and CD over iterations

Fig. 23(c) shows the effective number of top-rank individuals of the algorithms over iterations with the two operators, and Fig. 23(d) shows the corresponding effective ratios. The effective number refers to the number of distinct solutions after leaving repeated ones out in the Pareto solution set, while the effective ratio refers to the ratio of the effective number of Pareto solution set to the total size of the set. A larger number of Pareto solutions (higher effective ratio) indicate better diversity of solutions. As we can see, the effective size under CD continuously increases over iterations. After the 80th iteration, although the size of the top rank with classical operator is 200 (Fig. 23(a)), its effective size keeps only approximately 20, and the effective ratio drops to below 0.2.

Table 3 Comparison between the classic selection operator and CD

Instance	Operator	Effective size	Minimum cost	Minimum risk	CPU time (s)
2-2-20	Classic	16.40	1288.45	2677.48	520.6
	CD	23.60 (+44%)	1241.91(-4%)	2567.89 (-4%)	557.5 (+7%)
3-3-30	Classic	34.00	2278.18	5147.98	723.9
	CD	41.20 (+21%)	2178.68 (-4%)	5114.64 (-4%)	781.0 (+8%)
4-4-40	Classic	20.67	4061.71	4594.03	1164.1
	CD	33.00 (+60%)	3879.30 (-4%)	4479.57 (-2%)	1217.5 (+5%)
5-5-50	Classic	21.33	6326.22	9748.94	1725.1
	CD	30.17 (+41%)	5936.93 (-6%)	9651.41 (-1%)	1859.7 (+8%)
6-6-60	Classic	19.50	6254.16	9549.49	2532.2

CD	25.20 (+29%)	5841.23 (-7%)	9169.53 (-4%)	2752.5 (+9%)
Average improvement	39.08%	5.05%	2.44%	-7.21%

Table 3 shows the comparison under different sizes of instances generated by our instance generation method (10 runs for each instance). For example, 3-3-30 refers to the case of 3 hazmat types and 30 demand points, with 3 candidate DC for each hazmat type. The programs are implemented in Matlab 2019a on an Intel(R) Core(TM) i7-6500U CPU @ 2.50 GHz with an 8GB RAM PC. Compared to the classic selection operator, the CD can generate a larger number of effective Pareto solutions (39.08% improvement) and achieve better solutions in both minimum cost (5.05% improvement) and minimum risk (2.44% improvement). In terms of computational time, CD takes slightly more time than the classic algorithm by 7.21% on average, which is within the acceptable range. Therefore, the CD can obtain better solutions without significantly increasing the computational time.

4.5.5 Algorithmic comparison based on a benchmark function

To further investigate the generalization of CD in the multi-objective optimization, a standard test function for multi-objective optimization, ZDT-1 (Zitzler et al., 2000), is also taken as a benchmark. In the genetic algorithm with CD and ZDT-1, simulated binary crossover (SBX) and polynomial mutation (PM) are adopted.

Table 4 Comparative results based on the standard benchmark function

	Classic		CD	
Heredity ratio (%)	Effective size	CPU time (s)	Effective size	CPU time (s)
0	94.4	17.0	94.0 (-0.42%)	17.4 (-2.35%)
10	86.7	25.4	93.0 (7.27%)	25.8 (-1.57%)
20	84.7	24.6	90.2 (6.49%)	27.6 (-12.20%)
30	64.7	25.0	89.3 (38.02%)	25.8 (-3.2%)
40	55.6	25.6	85.8 (54.32%)	26.4 (-3.13%)
50	53.1	24.4	83.4 (57.06%)	25.6 (-4.92%)
60	35.8	25.2	83.7 (133.80%)	26.2 (-3.97%)
70	32.3	24.6	78.4 (142.73%)	25.8 (-4.88%)
80	15.0	24.4	77.3 (415.33%)	26.6 (-9.02%)

Since ZDT-1 is not originally for integer programming, an identical individual rarely exists both in the parents and offspring. To conduct a fair comparison, a direct-heredity ratio is arbitrarily set to ensure that a certain proportion of individuals can be directly inherited from the parents to the offspring so that the same individual phenomenon can be reproduced. To this end, the direct-heredity ratio is set as from 0% to 80%. The population size and iteration times are taken as 100. Again, the algorithm is run 10 times for each instance. Table 4 shows the effective number of the Pareto optimal solutions set and computational time of the two operators. As we can see, as the heredity ratio increases, the effective number of Pareto solutions of traditional selection decreases for both operators, which indicates that an increasing number of identical individuals come out in the population. Nevertheless, in comparison, the CD achieves a larger number of non-dominated solutions at the expense of slightly additional computational time. Specifically, when the

heredity ratio is 0, the effective sizes of the two algorithms are almost the same. However, when the heredity ratio is higher, the effective size of CD is consistently larger than that of the classic operator and the gap becomes larger. This result reinforces the message that our improved algorithm can produce better-spread non-dominated solutions with a limited computation burden.

4.5.6 Algorithmic comparison on the solution quality and optimality gap

While Sections 4.5.4 and 4.5.5 have verified that the proposed algorithm (NSGA-II-CD) is more efficient than its original version (NSGA-II), this section further verifies the effectiveness of the proposed algorithm by comparing its solutions with other heuristics in terms of solution quality and optimal gap. In addition to NSGA-II and NSGA-II-CD algorithms, the multi-objective simulated annealing (MOSA) and multi-objective particle swarm optimization (MOPSO) heuristic algorithms are selected for comparison. The mutation operator in NSGA-II is utilized to generate the new solution in MOSA.

To evaluate the solution quality and optimal gap of different algorithms, the ZDT-1 test function is taken as the benchmark bi-objective optimization problem, which has a theoretical non-convex (global optimal) Pareto front. Fig. 24 shows the comparison between the solutions and theoretical Pareto fronts of the four algorithms under 100 iterations. The algorithmic solution quality is higher (and thus the optimal gap is smaller) when the solutions are closer to the theoretical Pareto front. As we can see, NSGA-II based algorithms appear to perform better than MOSA and MOPSO. NSGA-II-CD outperforms other algorithms since its solutions are closer to the theoretical Pareto front.

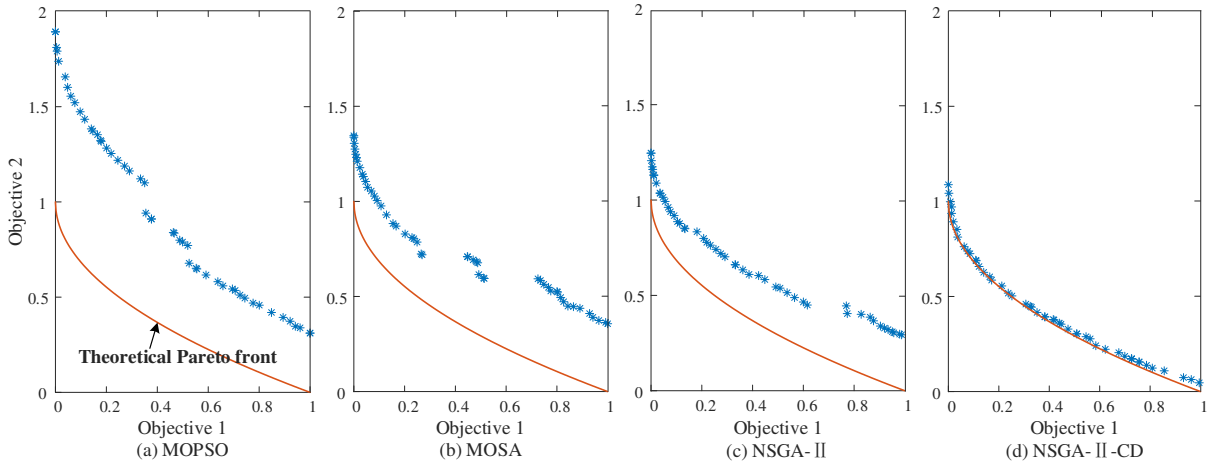


Fig. 24 Comparison of the optimal results and optimal Pareto fronts of the four algorithms

To further quantify their solution quality, two performance indicators are adopted to evaluate the convergence, universality, and uniformity of the solutions generated by different algorithms:

(1) Generational Distance (GD)

GD represents the average of the minimum distance from the optimal solution set obtained by the algorithm to the Pareto solution set, which is calculated as:

$$GD(\hat{P},\hat{P}^*)=\frac{\sqrt{\sum_{\dot{y}\in\hat{P}}\min_{\dot{x}\in\hat{P}^*}dis(\dot{x},\dot{y})^2}}{\left|\hat{P}\right|} \tag{75}$$

where $\hat{P}$ is the optimal solution set obtained by the algorithm, $\left|\hat{P}\right|$ is the solution quantity in the optimal

solution set, $\hat{P}^*$ is the theoretical Pareto solution set, and $dis(\dot{x},\dot{y})$ represents the Euclidean distance between solutions $\dot{x}$ and $\dot{y}$. In our test, 500 solutions are evenly selected on the Pareto front to construct the Pareto solution set. A smaller value of the GD indicator indicates the smaller gap between the optimal solution and the theoretical Pareto solution set, and thus better convergence of the algorithm and smaller optimality gap.

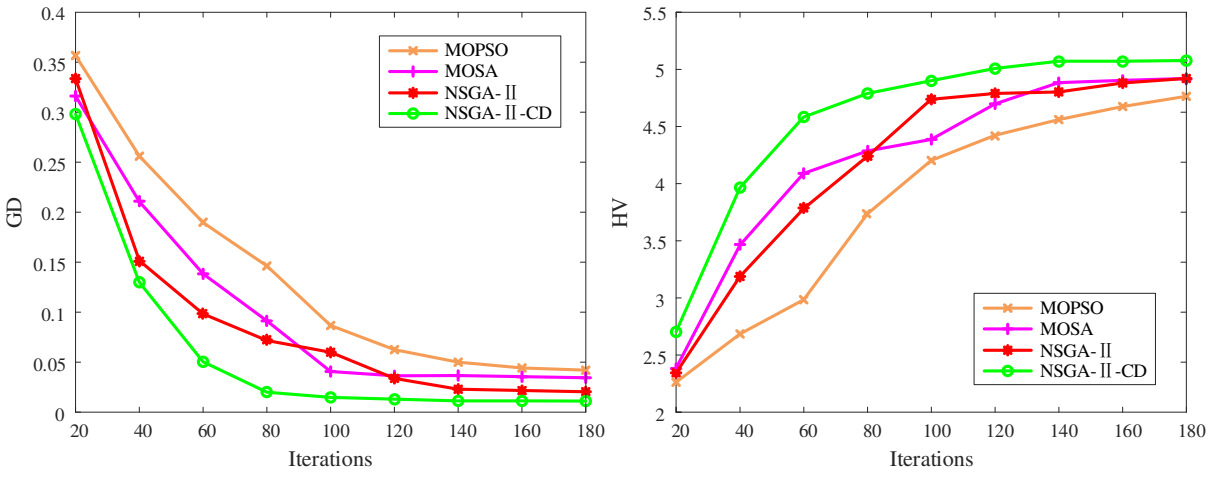
(2) Hyper Volume (HV)

HV represents the sum of Euclidean volume (the volume means the area in two-dimensional Euclidean space) formed between the optimal solution set and a reference point:

$$HV = \delta(\bigcup_{p=1}^{|\hat{P}|} \dot{x}_p) \tag{76}$$

where δ is the Lebesgue measure to calculate the volume. The HV indicator can be used to evaluate the universality and uniformity of the algorithms. A higher value of HV indicates the better the solutions perform in these two aspects.

We now calculate the two different indicators for each algorithm. To obtain less biased indicators, each experiment is repeated by 40 times, and the average value is taken. The results are shown in Fig. 25. As we can see, the values of GD are the lowest for the proposed algorithm under different iterations, which reinforces the message that the optimality gap of NSGA-II-CD is the smallest. Commendably, the values of HV are the highest for the proposed algorithm under different iterations. Therefore, it is concluded that the NSGA-II-CD can obtain near-optimal solutions and that with limited computational costs, NSGA-II-CD is more applicable to address the proposed bi-objective optimization problem than three others.



(a) GD (b) HV

Fig. 25 Convergence performance comparison of the four algorithms

5 Model application

5.1 Data setting

We assess the functionality of the proposed model and algorithm with a real-world case study in Guangzhou, which is the capital of Guangzhou Province and has one of the highest population densities in

China. The petrochemical industry is one of the pillar industries in Guangdong Province, while Guangdong's ethylene production ranks first in the country. The products produced in the petroleum and ethylene refining industry can provide a variety of raw materials for the chemical industry. For example, ortho-xylene is one of the raw materials for the production of pesticides, and ethylene plays an important role in the petrochemical industry chain. Benzene is an important raw material for fine chemicals. Given this fact, we choose these 3 materials in Guangzhou for the case study. Each material involves 1 factory and a couple of candidate distribution centers, and there are totally 37 respective industrial enterprises as the demand points. The classification of their chemical index in GHS (Globally Harmonized System of Classification and Labelling of Chemicals) is presented in Table 5.

Table 5 GHS classification of each material

Hazard type	Hazard category
Ortho-xylene	
Flammable liquids	3
Skin corrosion/irritation	2
Hazardous to the aquatic environment - short term (acute)	Acute 2
Ethylene	
Flammable gases	1
Gases under pressure	Compressed gas
Specific target organ - single exposure	3
Benzene	
Flammable liquids	2
Skin corrosion/irritation	2
Serious eye damage/eye irritation	2
Germ cell mutagenicity	1B
Carcinogenicity	1A
Specific target organ -repeated exposure	1
Aspiration hazard	1
Hazardous to the aquatic environment -short term (acute)	Acute 2
Hazardous to the aquatic environment -long term (chronic)	Chronic 3

In 2015, the total volume of road transportation of hazmat in China was 300 million tons, with a total of 196.4 thousand vehicles. Assume that the average vehicle capacity is 10 tons, and then the number of vehicle departures is 30 million. In this year, there were 50 accidents about hazmat transportation reported in China, thus the probability of hazmat transportation accidents is approximately 1.67 times per million departures. Multiplying the probability by the economic loss per capita caused by the corresponding hazmat type yields the risk multiplier r_l of type l. Following Nicolet-Monnier and Gherghe (1995), the impacted radius of a hazmat accident is approximated as 0.8 km. On this basis, the attenuation coefficient

τ_l can be estimated by Eq. (35), which equals 3.6.

After communication with chemical engineers, it is confirmed that no reaction exists among the three substances in this study, thus the risk supposed coefficients are 0. In fact, most inflammable and explosive hazmat will not produce nonlinear superimposed risk even when coexisted. If some special hazmat will interact, the superimposed risk can be simply calculated by Eq. (42) and Eq. (57), which will not affect the generalization of the framework.

Table 6 Information of hazardous materials and their plants

		Ortho-xylene	Ethylene	Benzene
Node number (i.e., type number) l		1	2	3
Risk coefficient r_l (RMB/capita)		0.00117	0.00217	0.00254
Attenuation coefficient τ_l		4.70	3.60	2.84
Plant's longitude		113.8388	113.1333	113.4210
Plant's latitude		23.1420	23.3677	23.4456
Risk supposed coefficient $\delta_{ll'}$	Ortho-xylene	0	0	0
	Ethylene	0	0	0
	Benzene	0	0	0

The information of each hazardous material and the corresponding plant are listed in Table 6. The information of each candidate distribution center is shown in Table 7. The information of customers (i.e., the companies that call for hazardous materials) is shown in Table 8. The other default model parameters are provided in Table 9, except where they are the subject of a test. In addition, the peak risk constraints (71) are not considered in the base case, which is equivalent to the infinite peak risk limit, whereas the effect of peak risk constraint will be investigated in Section 5.2.3.

Table 7 Information of candidate distribution centers

	ortho-xylene		ethylene			benzene	
Node number m	4	5	6	7	8	9	10
Longitude	113.70	113.66	113.22	113.20	113.17	113.27	113.42
Latitude	23.11	23.16	23.40	23.34	23.31	23.40	23.38
Opening cost c_m (RMB)	2076701	1984053	2194922	2278336	1782026	1806619	2924069
Service life (year)	15	15	15	15	15	15	15
Holding cost h_m (RMB/h)	32.20	45.50	40.21	23.15	45.08	23.58	37.62

Table 8 Information of customers (part 1)

Node number n & Name	Long.	Lat.	Node number n & Name	Long.	Lat.
agricultural chemical industry	**113+**	**22+**	**petrochemical industry-continued**	**113+**	**22+**
11 Donghong Chemical Plant	0.32	1.17	30 Jinhong Petrochemical Co., Ltd	0.62	1.29
12 Ruishang Chemical Co., Ltd	0.24	0.98	31 Shengxin Petrochemical Co., Ltd	0.65	1.37
13 Guangzhou Pesticide Factory	0.32	0.86	32 Senming Petrochemical Co., Ltd	0.45	1.11
14 Gangzhou Liwan Chemical Plant	0.27	1.20	33 Guanshi Petrochemical Co., Ltd	0.37	0.99
15 Tongbao Chemical Co., Ltd	0.84	1.15	34 LL Petrochemical-tech Co., Ltd	0.26	1.29

16 Jianhong Chemical Co., Ltd	0.39	1.12	35 Kaiguang Petrochemical Co., Ltd	0.50	1.13
17 Shidai Chemical Co., Ltd	0.41	0.90	**Fine chemical industry**	**113+**	**22+**
18 Litu Chemical Co., Ltd	0.48	1.13	36 Gao Tiya Fine Chemical Co., Ltd	0.27	1.40
19 Tengfu Chemical Co., Ltd	0.42	1.12	37 Huayu Fine Chemical Co., Ltd	0.50	1.40
20 Yuli Chemical Co., Ltd	0.47	1.10	38 DeRMB Fine Chemical Co., Ltd	0.57	1.19
21 Hualong Chemical Co., Ltd	0.47	1.10	39 Xiangying Fine Chemical Co., Ltd	0.42	1.38
22 Nanchaung Chemical Co., Ltd	0.62	0.87	40 Shengyan Fine Chemical Co., Ltd	0.29	1.26
23 Xiangpin Chemical Co., Ltd	0.35	1.08	41 Yusong Fine Chemical Co., Ltd	0.24	1.31
petrochemical industry	**113+**	**22+**	42 Xulin Fine Chemical Co., Ltd	0.27	1.27
24 RunJC Petrochemical Co., Ltd	0.50	1.13	43 JinSM Fine Chemical Co., Ltd	0.28	1.25
25 Yongchao Petrochemical Co., Ltd	0.22	1.40	44 Nuolan Fine Chemical Co., Ltd	0.40	1.37
26 HuiCY Petrochemical Co., Ltd	0.52	1.09	45 Wangsha Fine Chemical Co., Ltd	0.40	1.37
27 Hualong Petrochemical Co., Ltd	0.59	0.64	46 Barui Fine Chemical Co., Ltd	0.37	1.31
28 Yirun Petrochemical Co., Ltd	0.75	1.14	47 Yonghui Fine Chemical Co., Ltd	0.35	1.30
29 Fuhao Petrochemical Co., Ltd	0.75	1.22			

Table 8 Information of customers (part 2)

Node number n	11	12	13	14	15	16	17	18	19	20
μ_n (ton/h)	0.12	0.16	0.16	0.13	0.17	0.18	0.18	0.15	0.18	0.17
h_n (RMB/h)	72.80	102.20	78.40	67.20	98.00	86.10	63.00	95.20	68.60	88.90
mw_n (RMB)	12.93	8.48	12.11	9.27	10.70	13.54	12.95	9.61	9.87	9.49
Pw_n (h)	24	12	24	12	36	24	12	36	24	12
tw_n (h)	9.77	8.77	15.29	9.20	28.25	20.70	5.35	32.68	21.99	2.19
Node number n	21	22	23	24	25	26	27	28	29	30
μ_n (ton/h)	0.14	0.08	0.19	0.17	0.19	0.13	0.19	0.13	0.15	0.08
h_n (RMB/h)	58.80	78.40	68.60	69.46	62.76	85.30	63.37	86.52	73.12	73.12
mw_n (RMB)	9.02	9.36	11.29	10.12	8.60	9.31	9.31	13.99	8.86	10.87
Pw_n (h)	24	24	36	36	36	12	24	36	36	24
tw_n (h)	17.88	11.02	2.83	28.31	23.05	3.14	20.98	0.89	5.63	9.29
Node number n	31	32	33	34	35	36	37	38	39	40
μ_n (ton/h)	0.17	0.09	0.16	0.15	0.13	0.13	0.13	0.18	0.16	0.15
h_n (RMB /h)	48.74	49.35	60.93	54.84	81.04	59.19	54.67	71.23	41.13	48.15
mw_n (RMB)	9.53	11.71	13.59	10.42	11.77	10.73	13.78	13.39	11.24	12.01
Pw_n (h)	12	24	36	12	12	36	36	24	12	36
tw_n (h)	8.62	3.68	7.47	10.23	3.35	0.62	3.13	22.93	0.97	4.74
Node number n	41	42	43	44	45	46	47			
μ_n (ton/h)	0.09	0.13	0.12	0.13	0.18	0.17	0.20			
h_n (RMB /h)	60.69	56.18	53.17	45.65	53.17	65.21	70.73			
mw_n (RMB)	8.95	13.75	12.02	12.49	10.62	8.52	8.10			
Pw_n (h)	24	12	24	12	12	36	24			
tw_n (h)	15.05	2.76	2.71	7.86	4.37	21.02	0.92			

Table 9 Supplementary information

Service level α	0.90	Confidence level β	0.90	Variation coefficient of D_n	0.1

Vehicle capacity S^l	5 ton	Vehicle velocity vel	30 km/h	Fixed transportation cost u	50 RMB
DC capacity S_m	50 ton	DC ordering cost F_m	500 RMB	Customer ordering cost F_n	300 RMB
Fuel price c_f	6.38RMB/Litre				

To ensure the continuity of highways connecting Guangzhou to the surrounding cities, several roads of other cities and the corresponding population density are added onto the network for modeling, including Foshan, Zhongshan, Dongguan, Qingyuan, Shaoguan, and Huizhou city. A number of nodes are used to describe the road network of the above-mentioned cities, which includes the intersections and turnings. As a result, a total of 303 nodes are extracted from the map to describe the intersections and turnings of the real road network. Fig. 26 shows the distribution of nodes, restricted area, road network, and the heat map of population density in the area of interest. A darker color indicates the higher population density in the area. The restricted area in Guangzhou is within S81 Guangzhou Ring Expressway, including Tianhe District, Yuexiu District, and Liwan District, which is also the downtown area of Guangzhou. Hazmat transportation is prohibited in the area from 7:00 am to 10:00 pm each day. In what follows, a set of computational studies are carried out to assess the performance of our model and provide managerial implications from the perspective of practitioners.

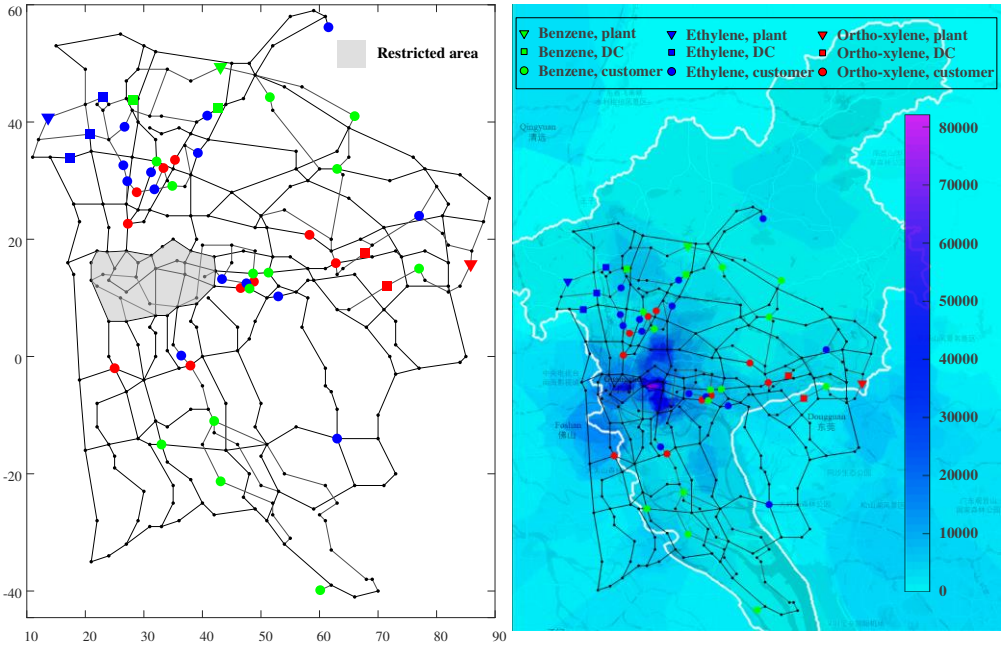


Fig. 26 Transportation network and population distribution representation for Guangzhou city

5.2 Results and discussion

5.2.1 Convergence analysis of cost and risk components

Prior to providing a detailed analysis of the convergence behavior of cost and risk components, we introduce some additional concepts. Fig. 27(a) illustrates the solution set in a typical iteration, where the light grey area and dark gray area indicate the distribution of the overall populations and the distribution of the top-rank set, respectively. The lower bound and upper bound of the area indicates the lowest value and highest value in this set, respectively.

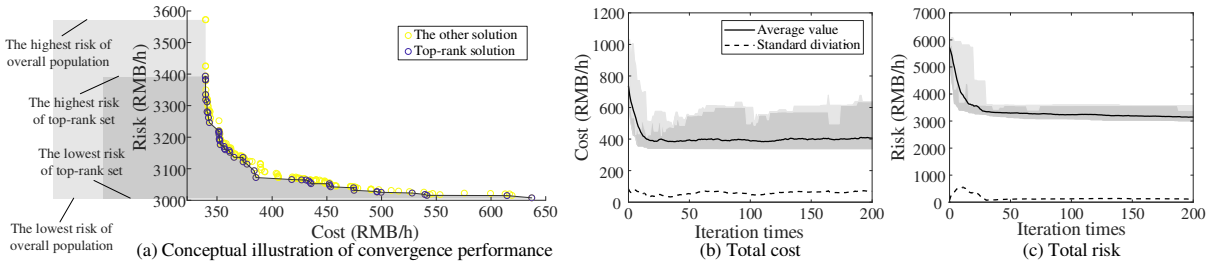


Fig. 27 Illustration and convergence process with respect to the total cost and risk

Figs. 27(b) and (c) present the iteration process with respect to the total cost and risk, respectively. It can be seen that the minimum cost and maximum total risk of the optimal solution set tend to stabilize after the 20th iteration. The minimum total cost is decreased by 44.24% as compared to that of the first iteration (607.89 RMB/h). The highest total risk drops to 3933.40 RMB/h, which is 39.73% lower than that of the first iteration (5630.25 RMB/h). After the 20th iteration, the minimum total risk continues to decrease, whereas the maximum total cost grows in turn, which suggests that the system risk can be further reduced at the expense of extra cost. Finally, the minimum total risk drops by 46.04%.

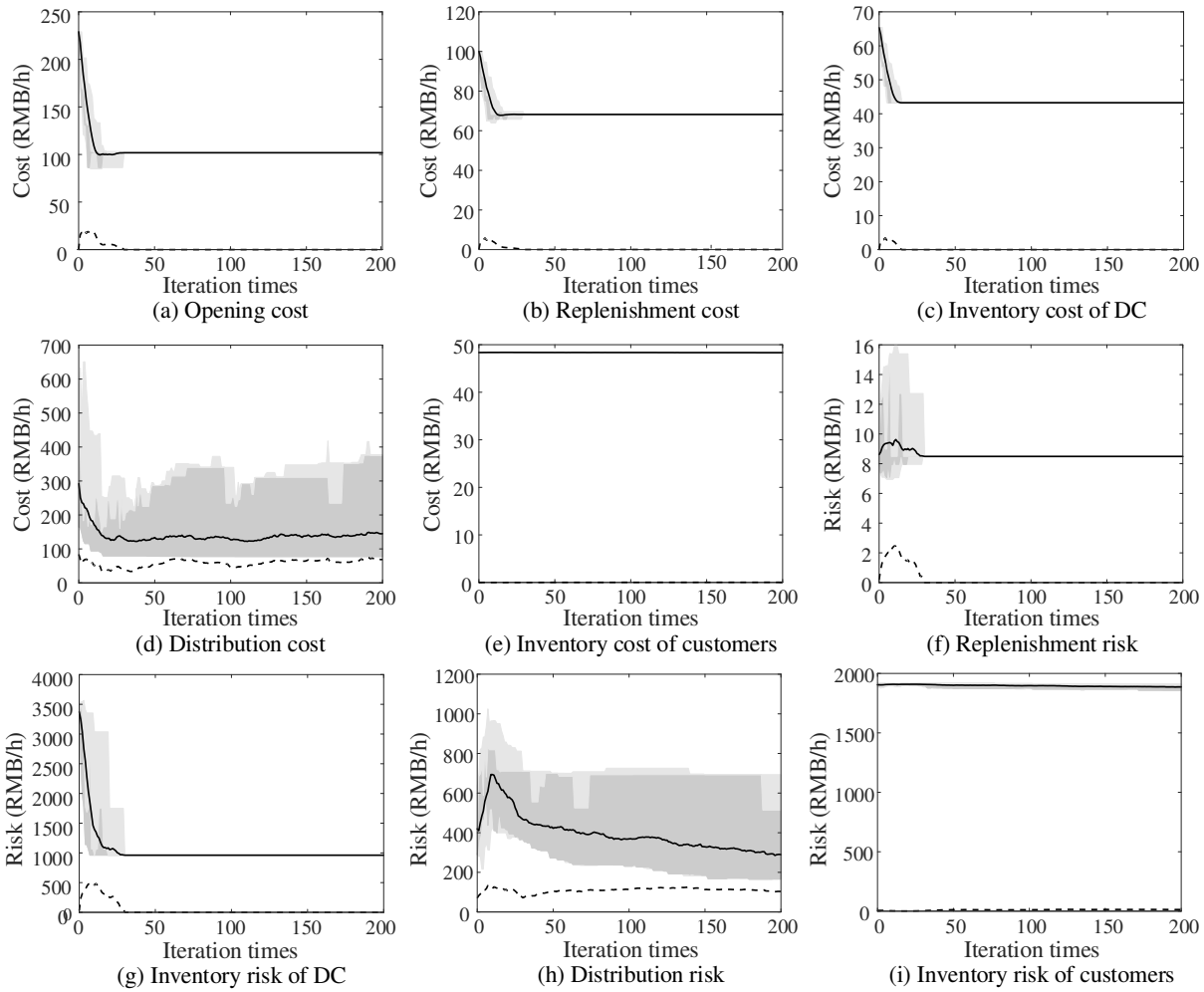


Fig. 28 Convergence process with respect to each component

We now analyze the convergence behavior of each system cost component. In our model, the total cost consists of 7 components: 1) the opening cost of DCs; 2) the replenishment cost from plants to DCs; 3) the

inventory cost of DCs; 4) the distribution cost from DCs to customers; and 5) the inventory cost of customers. The total risk consists of 5 components: 1) the replenishment risk from plants to DCs; 2) the inventory risk of DCs; 3) the distribution risk from DCs to customers; and 4) the inventory risk of customers.

Fig. 28 shows the convergence behavior of the aforementioned components. The decrease of the minimum total cost (i.e., the lower bound of total cost area) before the 30th iteration comes from the joint optimization of the location, replenishment, and distribution. After that, the opening cost, inventory cost and risk of distribution centers, replenishment cost, and risk converge to fixed values, and the change of total cost and total risk is mainly affected by the change of distribution plan. The changes in inventory cost and risk of customers are trivial. The reason is that the inventory level of customers varies linearly in each cycle with the minimum value 0 and the maximum value $D_n\tilde{P}_k$. Given the unit inventory cost per unit time h_n, then the average inventory cost per cycle is $h_n\,\mu_n\tilde{P}_k/2$, thus the average inventory cost per unit time is $h_n\,\mu_n\tilde{P}_k/2\tilde{P}_k$, that is, $h_n\,\mu_n/2$, which is a constant.

The distribution cost consists of four components: transportation cost (including fixed cost and fuel consumption cost), ordering cost of customers, detour cost, and time windows penalty cost. The distribution risk consists of two components: transportation risk and detour risk. The convergence behavior of these six components is depicted in Fig. 29.

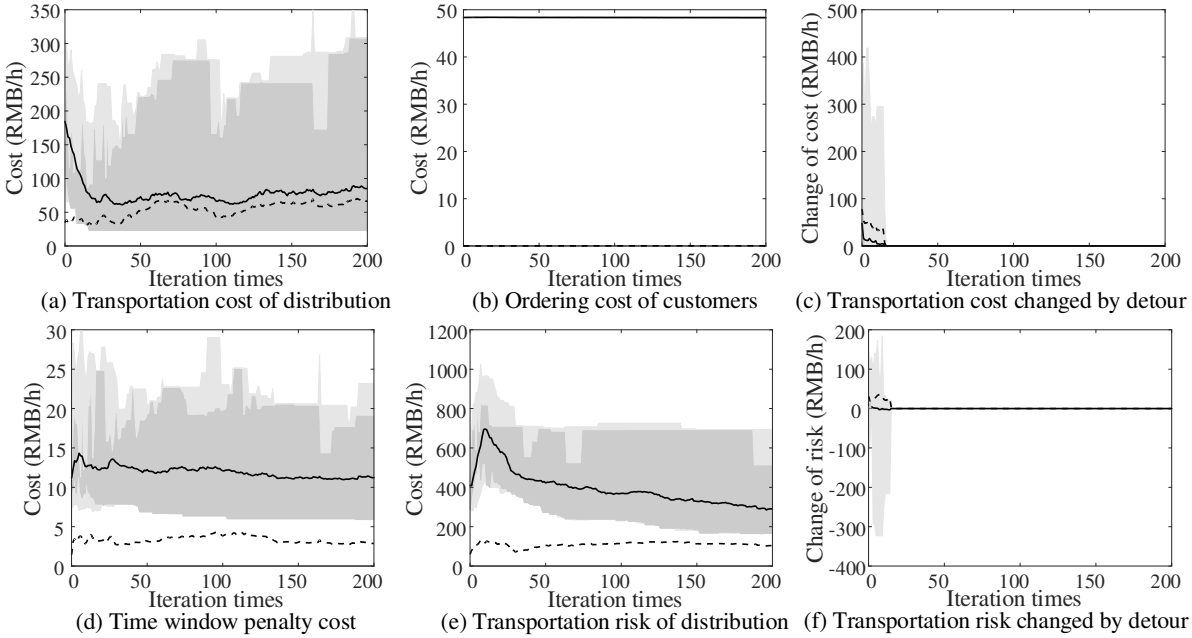


Fig. 29 Convergence process of each component with respect to the distribution cost

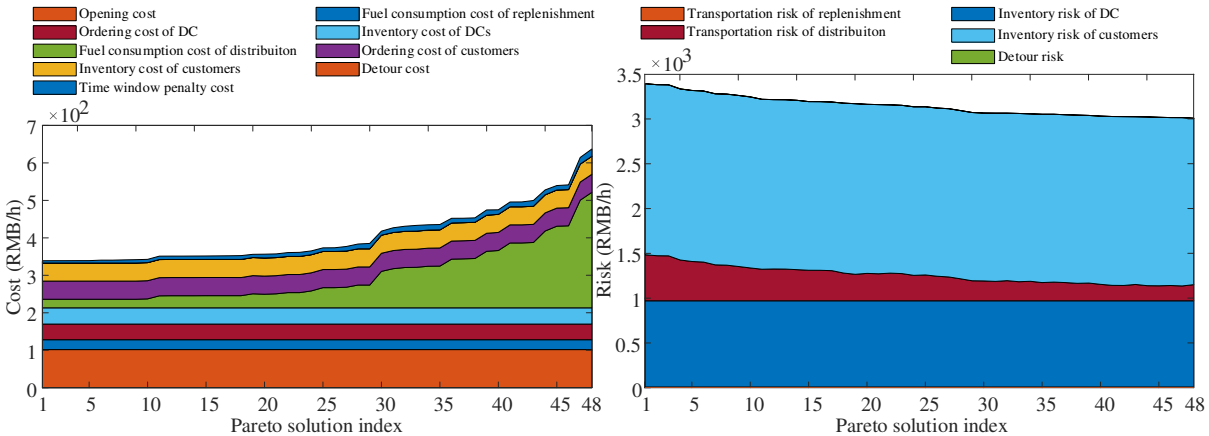
As we can see, the transportation cost of distribution and the transportation cost changed by detour account for a large proportion of the distribution cost, which may exert a great influence on the convergence behavior of total cost. As shown in Figs. 29(a) and (e), the minimum distribution cost decreases to its minimum value before the 20th iteration. Afterward, the maximum distribution cost generally increases, alongside the decrease of the minimum distribution risk. This suggests that some feasible solutions with higher total costs can reduce the system total risk at the expense of extra

transportation costs.

Figs. 29(c) and (f) show the change patterns of detour costs and risks, where the detour cost is caused by vehicles entering the restricted area during the forbidden period. Before the 20th iteration, the detour cost is relatively high in that a number of hazmat transportation routes go through the restricted area. In this range, the minimum detour risk is negative, which indicates the rationality of traffic restriction policy in Guangzhou city. Beyond this range, the detour cost and risk converge to 0, which indicates that the optimal solutions can avoid the detour cost caused by vehicles entering the restricted area during the forbidden period. In other words, the optimal distribution routes do not go through the restricted area. This suggests that, in practice, the periodic road closures policy in Guangzhou could be possibly upgraded to a full-time prohibition. In this way, urban safety can be ensured without significant impact on the existing hazmat logistics system.

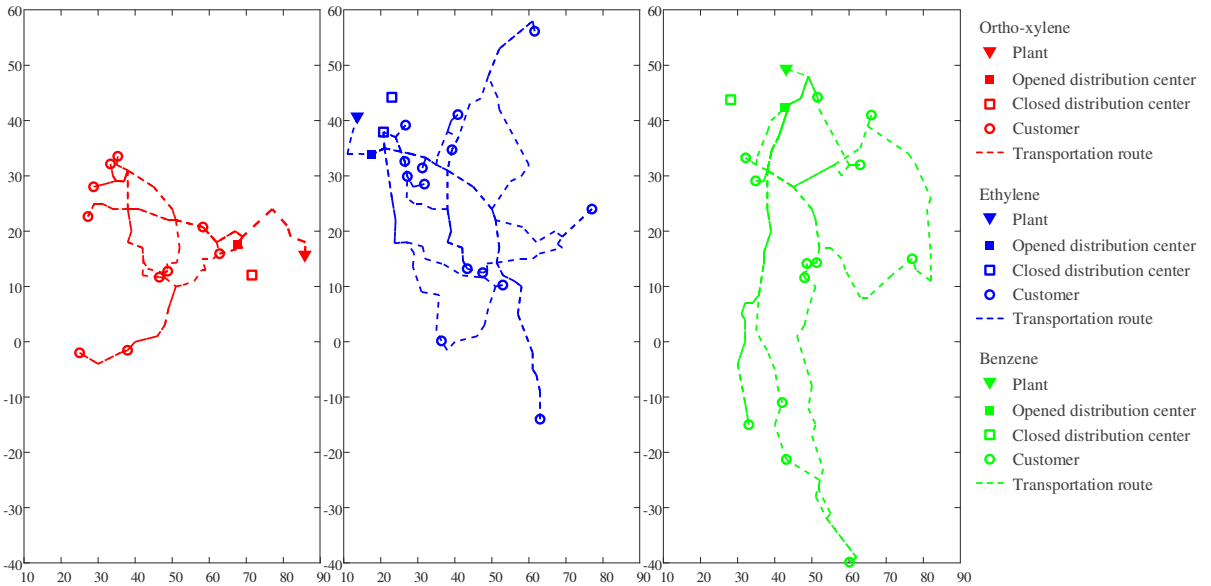
5.2.2 Analysis of Pareto solutions

Fig. 30 shows the cost and risk composition over all Pareto solutions. As the index of Pareto solutions increases, the cost increases while the risk decreases. This suggests that higher risk leads to smaller costs. The changes of opening cost, replenishment cost, inventory cost, detour cost, and time windows penalty cost are trivial over different solutions, whereas the fuel consumption cost of distribution increases dramatically. This is because to reduce the risk, a longer tour is required to pass the densely populated area, such that more distribution routes and fuel consumption costs are needed to avoid the excessively long distance for an individual vehicle. Correspondingly, the transportation risk reduces considerably as the index of Pareto solutions increases. Fig. 31 displays the distribution networks of different types of hazmat for solution 48, including the opened distribution centers and transportation routes. Table 10 provides the detailed replenishment and distribution schemes of solutions 1, 24, and 48, which includes the optimal order cycle for both each distribution center and distribution route (unit: hour) and the corresponding vehicle routing.



(a) Cost composition of Pareto solutions (b) Risk composition of Pareto solutions

Fig. 30 Cost and risk composition over Pareto solutions



(a) The supply network of ortho-xylene (b) The supply network of ethylene (c) The supply network of Benzene

Fig. 31 Diagram of the hazmat supply networks of solution 48

As different Pareto optimal solution represents different risk levels, how the index of the Pareto solution affects the fleet size and route lengths deserve some discussion. Fig. 32 presents the number of distribution routes, the average route length, and the resulting total route length (the number of distribution routes multiplied by the average route length) for different Pareto solutions, where a larger index indicates lower risk and higher cost. As the total risk decreases, the number of routes generally increases, and the average length of distribution routes decreases. Given this fact, the total route length slowly decreases when the number of routes remains unchanged (the grey area), but increases rapidly when the number of routes grows (the white area). Therefore, the total route length is mainly dependent on the number of routes. As can be seen from Fig. 30(a), the total risk reduces slowly over the Pareto solutions, whereas the number of distribution routes and the total length of distribution increases considerably, which also explains the increase in distribution costs in a nearly exponential way (Fig. 30(b)). These results suggest that, in practice, the total risk can be possibly reduced by partitioning the overall customers into several groups and frequently small shipments with multiple routes. However, the authorities need to trade-off the cost and risk, in that when the risk management reaches a certain level, a small reduction in risk will result in a significant increase in the marginal transportation cost (Fig. 30(b)).

Table 10 Replenishment and distribution schemes of solutions 1, 24 and 48

Solution No.	Type	Opened DC No.	Cycle of replenishment	Distribution routes	Cycle of distribution
1	Xylene	5	71.35	5-34-12-35-47-38-33-45-15-42-18-5	38.69
	Ethylene	8	59.70	8-24-23-39-22-32-26-14-29-43-20-44-40-11-36- 8	40.02
	Benzene	10	54.13	10-41-46-13-16-27-17-31-21-19-28-30-25-37-10	36.16
24	Xylene	5	71.35	5-38-12-35-47-34-33-45-15-42-18-5	38.69

	Ethylene	8	59.70	8-43-44-11-22-32-26-29-14-24- 8	42.07
				8-23-20-40-39-36-8	36.98
	Xylene	10	54.13	10-37-46-41-31-25-19-21-27-16-17-13-2 8-30-10	36.16
48	Xylene	5	71.35	5-38-33-47-34-35-12-5	38.33
				5-15-5	37.30
				5-42-45-18-5	40.00
	Ethylene	8	59.70	8-43-44-22-14-32-29- 8	43.40
				8-26-11- 8	42.32
				8-24-39-8	36.52
				8-23-20-40-36-8	36.87
	Benzene	10	54.13	10-30-41-37-10	44.72
				10-25-19-21-27-16-17-10	33.70
				10-13-10	35.74
				10-46-31-28-10	35.51

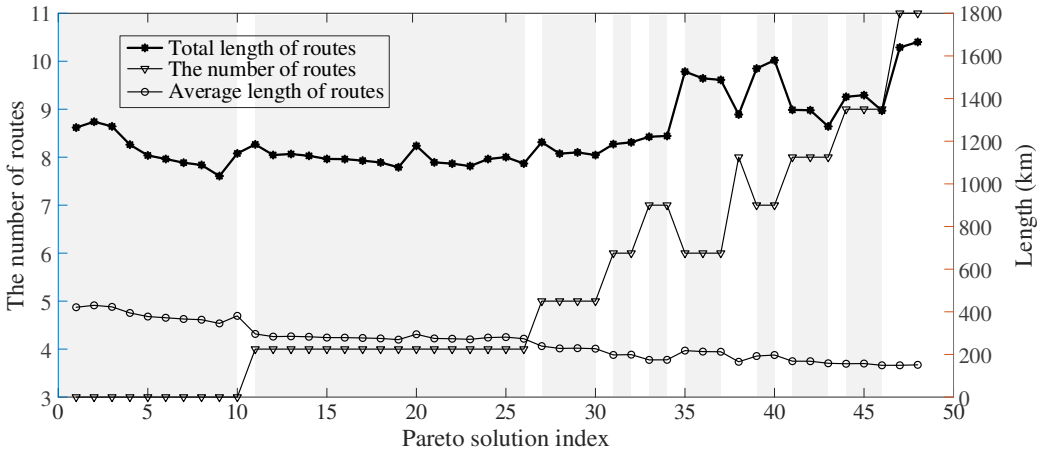


Fig. 32 The distribution schemes of Pareto optimal solutions

5.2.3 Sensitivity analysis

In this section, sensitivity analysis is conducted to investigate how the system performance changes with critical parameters, including peak risk constraint, and the level of service. In what follows, the highest/average/lowest total cost/risk indicates the highest/average/lowest value over all Pareto solutions. The total cost/risk of a specific Pareto solution is calculated by summing up the site-specific cost/risk. The cost composition in each chart represents the average values of respective components over Pareto solutions. Note that the risk of the detour is negative, which reduces the total risk.

5.2.3.1 Impact of risk superposition coefficient

The multi-class hazmat, which is characterized by the risk superposition coefficient, is a new feature in our model and the design of the distribution network. In this section, we analyze the impact of the risk superposition coefficient on the distribution network design and compare that without this effect. To this end, we introduce the concept of route overlapping ratio, which is defined as the route length overlapping

by different types of hazmat divided by the total route length.

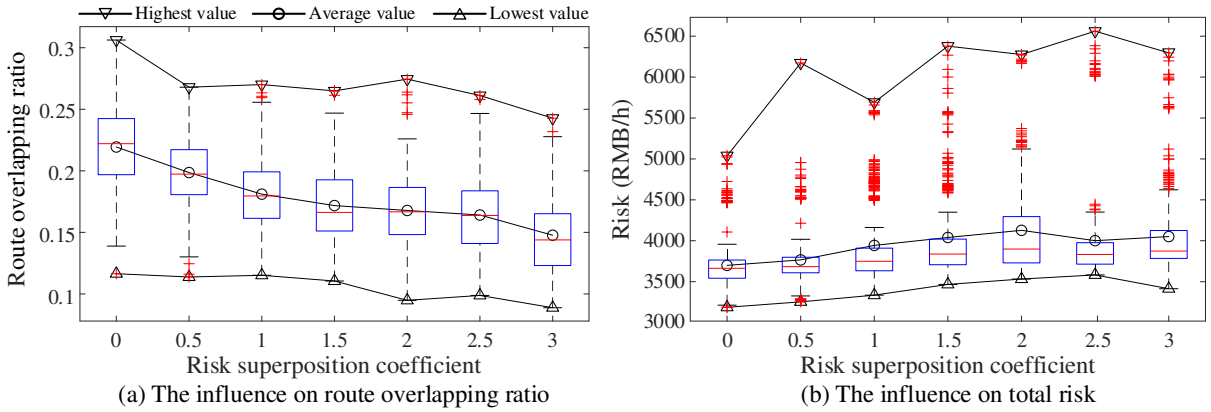


Fig. 33 Influence of risk superposition coefficient on the system

To facilitate the analysis, we assume that the risk superposition coefficients of the three types of hazmat $\delta_{ll'}$ are identical. Each experiment is repeated by 5 times. The results are shown in Fig. 33, where the error bars indicate the distribution of optimal solutions. As we can see, as the value of the risk superposition coefficient increases, the route overlapping ratio generally decreases, and the total risk increases. This is because, with the enhancement of interactions between different types of hazmat, the routes for different types of hazmat should be distributed as separately as possible to reduce the additional total superimposed risk. This suggests that the distribution network should be carefully designed in the provision of multi-class hazmat.

5.2.3.2 Impact of peak risk constraint

In this study, the peak risk is not allowed to exceed a specific threshold, as shown in constraints (71). We now investigate the impact of peak risk constraints on the cost and risk composition, and the results are shown in Fig. 34. As shown in Fig. 34(a), with the increase of peak risk constraint, the highest cost of Pareto optimal solution drops, while the average cost and the lowest cost decrease slightly, which mainly results from the reduction in the fuel consumption cost of distribution. This provides a practical insight that to achieve a small expected cost, we should endure a relatively large risk.

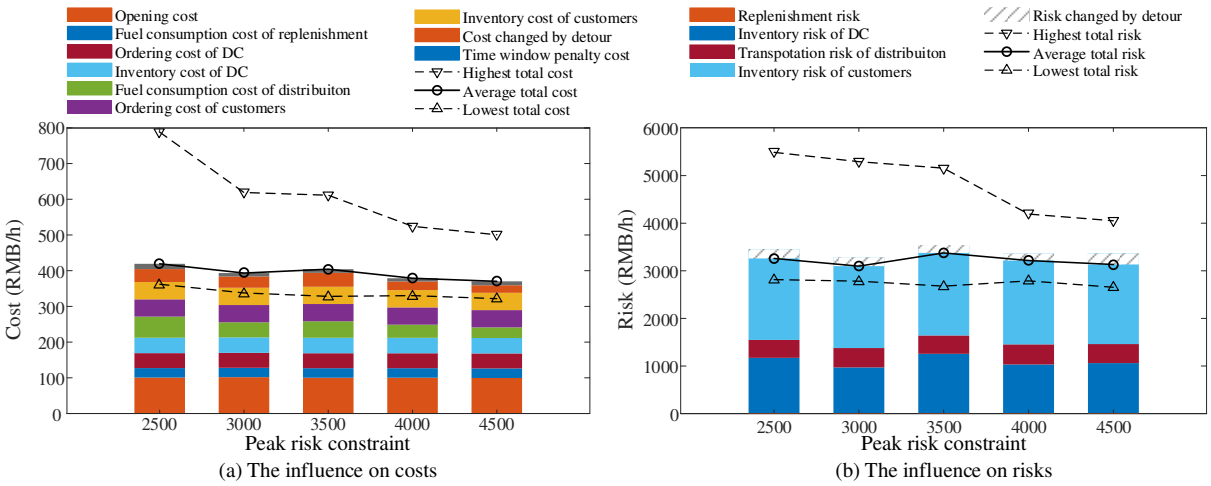


Fig. 34 Impact of the peak risk constraint on the costs and risks

Interestingly, the relationship between the highest total risk and peak risk constraint in Fig. 33(b) constitutes a Pareto frontier, which indicates the trade-off between total risk minimization and risk equilibration. Specifically, the highest total risk decreases considerably with the increase of peak risk constraint. This is because a strict peak risk constraint on each site would redistribute the transportation and inventory schemes more evenly over the whole network, which may, in turn, penalize zones with low risk, ending up with an increase of total risk. While the peak risk constraint indicates the risk equity, this result provides a practical insight that, in the hazmat supply chain management, the authorities should not try to find a perfect distribution of risk, and they should make a trade-off the risk equity and total exposed risk.

5.2.3.3 Impact of the level of service

The level of service α refers to the probability that customers will not be out of stock, which is essential to planning the distribution redundancy and safety stock. Fig. 35 shows the impact of the level of service on the system costs and risk. As the level of service increases, the average cost and the highest cost in Pareto optimal solutions increase substantially. Notably, the average cost and highest cost increase dramatically when the level of service grows from 0.95 to 0.99, which mainly results from the increase of detour cost.

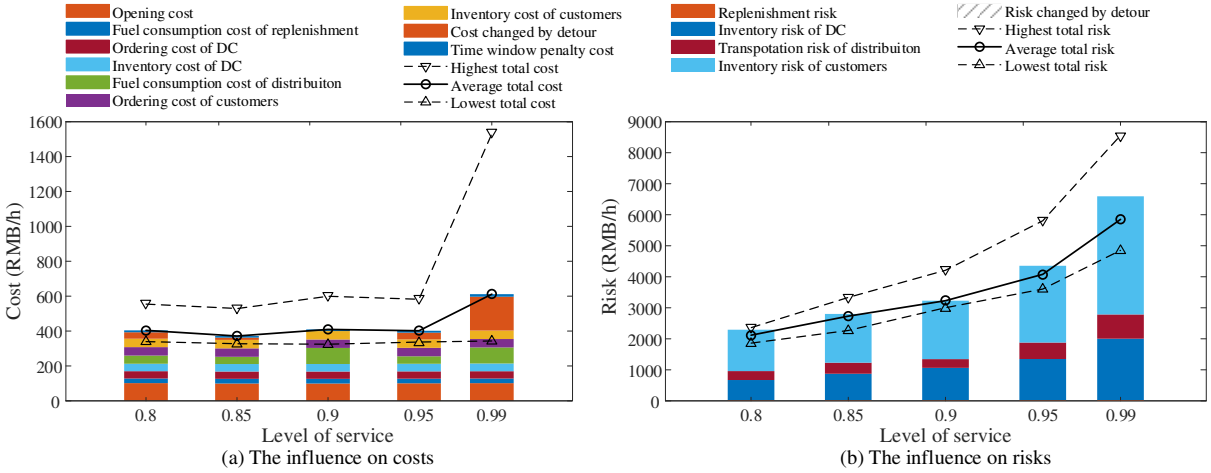


Fig. 35 Impact of the level of service on the costs and risks

As shown in Fig. 35(b), there is a positive correlation between the level of service and total risk. As the level of service increases, the average cost, the highest cost, and the lowest cost in Pareto optimal solutions increase considerably. This is because a larger amount of distribution redundancy and safety stock is required when the level of service is higher, which results in a higher level of risk. This suggests that, in the hazmat supply chain management, the operator should make a trade-off between the risk and the level of service.

6 Concluding remarks

In this paper, we introduce a new multi-class hazmat distribution network design problem with inventory and superimposed risk (MHND) in a multi-echelon supply chain. Although their in-transit and storage are incompatible, the superimposed risks among different materials stemming from possible

chemical reactions once accidents (e.g., leakage, explosion) happen are explicitly considered. The long-term cyclic time windows penalty cost with periodic route closures policy are also formulated. A new population-based risk definition is proposed to evaluate the risk for the population at any location and any time with respect to its multi-class hazmat logistics system. In particular, we introduce risk superposition coefficients for multi-class hazmat, which allows for capturing possible superimposed risks (e.g., chemical reaction) among different hazmat and accommodates a general system with more than two hazmat types. This makes our model scalable for multiple types of hazmat, which has significant implications in the realm of hazmat logistics. Commendably, we develop a knowledge-based NSGA-II algorithm with cyclic dissimilarity-based elitist selection (NSGA-II-CD) to solve the problem to obtain an efficient Pareto frontier for minimizing the total cost and minimizing the total risk. A cyclic dissimilarity-based elitist selection (CD) operator is devised to address the issue of speeding proliferation, which can produce better-spread non-dominated solutions with a limited computation burden. Further test instances were conducted to prove the scalability and robustness of the proposed algorithm.

The proposed model was applied to a metropolitan-wide real-world case study in Guangzhou, China. The impact of population density is explicitly considered. The analysis of the cost and risk convergence process indicates that it is possible to achieve significant reductions in population exposure while reducing the total cost. We thus suggest that, from the perspective of the traffic management sector, the time-dependent road closures policy in Guangzhou could be possibly upgraded to a full-time prohibition. In addition, we also reveal a couple of new insights that have not been looked at before, which are summarized as follows: First and foremost, the transportation risk and inventory risk are highly correlated in the hazmat logistics systems. As such, in addition to the transportation risk, it is necessary to examine the occupational risk associated with the temporary storage of hazardous materials in distribution centers and customer points. Second, there is a positive convex relation between risk minimization and risk equilibration. The authorities should not try to find a perfect distribution of risk, and they should make a trade-off the risk equity and total exposed risk. Third, there is a positive correlation between the level of service and total risk. Therefore, in practice, the agencies should make a trade-off between economic viability of the system, exposed risk, and maintaining good service for customers. Fourth, the interactions between different types of hazmat may exert great influence on the distribution network design. The route overlapping ratio for different types of hazmat will be reduced with greater interactions.

This paper opens up new research directions. For instance, while this paper primarily focuses on the city-wide case study, future research can investigate the multi-class hazmat location inventory routing problem over an intermodal transportation network. Another possible extension is to model the correlated demand to more accurately represent the demand characteristics. In addition, designing a new meta-heuristic to further improve the algorithmic performance may be interesting work.

Appendix A. Derivation of Eqs. (1)-(3)

As mentioned previously in (A8), in the hazmat supply chain, the customers are assumed to determine their optimal order quantity based on the principle of economic order quantity. According to the cyclic

inventory theory, the order cycle of each customer n on route k is its ordering quantity $\bar{O}_n$ divided by consumption rate D_n:

$$\bar{P}_n = \frac{\bar{O}_n}{D_n} = \frac{\sum_{n\in N_k} \bar{O}_n}{\sum_{n\in N_k} D_n} \quad ,\forall n \in N_k$$

where N_k represents the set of customers on route k, so that we can abbreviate $\sum_{n\in N} X_{n,k} \cdot$ to $\sum_{n\in N_k} \cdot$.

The ordering cost of route k per unit time is equal to the total ordering cost divided by the order cycle.

$$\frac{\sum_{n\in N_k} F_n}{\bar{P}_n} \quad ,\forall n \in N_k$$

The inventory holding cost of route k is the summation of customers on this route.

$$\sum_{n\in N_k} \frac{h_n \bar{O}_n}{2} = \frac{\sum_{n\in N_k} h_n D_n}{2} \bar{P}_n \quad ,\forall n \in N_k$$

Therefore, the total inventory cost of this route is the summation of these two parts. Applying the first-order condition on the total inventory cost yields the optimal order cycle time that minimizes the total inventory cost as

$$\bar{P}_n = \sqrt{\frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n D_n}} \quad ,\forall n \in N_k$$

As a result, the optimal order quantity for customer n should be

$$\bar{O}_n = \bar{P}_n D_n = D_n \frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n D_n} \quad ,\forall n \in N_k$$

Due to the demand uncertainty, a reliable replenishment plan from the distribution centers to the customers should be made to avoid the shortage. To this end, replenishment redundancy is required to meet a certain confidence level.

Given the independent and identically normal distributed demand, i.e., $D_n \sim N(\mu_n, \sigma_n^2)$, then we have

$$\sum_{n\in N_k} h_n D_n \sim N\left(\sum_{n\in N_k} h_n \mu_n \,, \sum_{n\in N_k} h_n^2 \sigma_n^2\right) \quad ,\forall n \in N_k$$

Since the demand is independently and identically distributed, according to the above formula, the replenishment times (the reciprocal of the order cycle) also follow the normal distribution.

$$\frac{1}{\bar{P}_n^2} = \frac{\sum_{n\in N_k} h_n D_n}{2\sum_{n\in N_k} F_n} \sim N\left[\frac{\sum_{n\in N_k} h_n \mu_n}{2\sum_{n\in N_k} F_n}, \left(\frac{\sum_{n\in N_k} h_n \sigma_n}{2\sum_{n\in N_k} F_n}\right)^2\right] \quad ,\forall n \in N_k$$

Taking the expected value as the actual replenishment times, we have that:

$$\frac{1}{\tilde{P}_k} = \sqrt{\frac{\sum_{n\in N_k} h_n \mu_n}{2\sum_{n\in N_k} F_n}} \Rightarrow \tilde{P}_k = \sqrt{\frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n \mu_n}}$$

We now calculate the corresponding delivery quantity. Given the actual cycle time, the optimal order quantity of customer *n* and route *k* can be calculated as follows:

$$\bar{Q}_n = D_n \tilde{P}_k = D_n \sqrt{\frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n \mu_n}} \quad ,\forall n \in N_k$$

Since the optimal order quantity $\bar{Q}_k = \sum_{n\in N_k} \bar{Q}_n$ is a linear combination of multiple independent and identically distributed random variables, it is a random variable following a normal distribution.

$$\bar{Q}_k \sim N(\mu_k, \sigma_k^2) \quad ,\mu_k = \sqrt{\frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n \mu_n}} \sum_{n\in N_k} \mu_n \quad ,\sigma_k^2 = \sqrt{\frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n \mu_n}} \sum_{n\in N_k} \sigma_n^2$$

Assume that the level of service should be higher than a threshold, we have that

$$P\{\bar{Q}_k \le \tilde{Q}_k\} \ge \alpha$$

where P is the probability and α is the confidence level.

Let $q_k = \bar{Q}_k - \tilde{Q}_k$, then we have

$$\mathrm{E}(q_k) = \mathrm{E}(\bar{Q}_k) - \tilde{Q}_k = \tilde{P}_k \mu_k - \tilde{Q}_k$$
$$\mathrm{D}(q_k) = \mathrm{D}(\bar{Q}_k) = \tilde{P}_k^2 \sigma_k^2$$

Let $\eta = \frac{q_k - \mathrm{E}(q_k)}{\sqrt{\mathrm{D}(q_k)}}$, then $\eta \sim N(0,1)$, and $q_k = \bar{Q}_k - \tilde{Q}_k \le 0$, η is equivalent to

$$\eta = \frac{q_k - \mathrm{E}(q_k)}{\sqrt{\mathrm{D}(q_k)}} \le \frac{\mathrm{E}(q_k)}{\sqrt{\mathrm{D}(q_k)}}$$

Therefore, the delivery quantity constraint $\bar{Q}_k \le \tilde{Q}_k$ is equivalent to

$$P\left\{\eta \le \frac{-\mathrm{E}(q_k)}{\sqrt{\mathrm{D}(q_k)}}\right\} \ge \alpha$$

The above inequity can be transformed into the following formula

$$\sqrt{\mathrm{D}(q_k)} z_\alpha + \ \mathrm{E}(q_k) = \tilde{P}_k(\mu_k + \sigma_k z_\alpha) - \tilde{Q}_k \le 0$$

where z_α is the quantile statistics with a confidence level of α.

$$\tilde{Q}_k \ge \tilde{P}_k(\mu_k + \sigma_k z_\alpha)$$

As a result, the reliable delivery quantity of route k takes the following form:

$$\tilde{Q}_k \ge \tilde{P}_k(\mu_k + \sigma_k z_\alpha) = \frac{2\sum_{n\in N_k} F_n}{\sum_{n\in N_k} h_n \mu_n} \sum_{n\in N_k} (\mu_n + \sigma_n z_\alpha) = \frac{2\sum_{n\in N} X_{n,k} F_n}{\sum_{n\in N} X_{n,k} h_n \mu_n} \sum_{n\in N} X_{n,k}(\mu_n + \sigma_n z_\alpha)$$

where $\mu_k = \tilde{P}_k \sum_{n\in N} X_{n,k}\, \mu_n$, $\tilde{P}_k = \sqrt{\frac{2\sum_{n\in N} X_{n,k} F_n}{\sum_{n\in N} X_{n,k} h_n \mu_n}}$

This completes the derivation of Eqs. (1)-(3).

Appendix B. Proof of Proposition 1

Proof: Let $\frac{2\pi}{T_2} T_1 = \omega$, $\frac{2\pi}{T_2}\varphi_1 + \varphi_2 = \varphi$, we have:

$$\lim_{T\to+\infty} \frac{1}{T} \sum_{t=1}^{T} \cos\left[\frac{2\pi}{T_2}(T_1 t + \varphi_1) + \varphi_2\right] = \lim_{T\to+\infty} \frac{1}{T} \sum_{t=1}^{T} \cos(\omega t + \varphi)$$

When the order cycle is not an integer multiple of the time window cycle, i.e., $\omega \ne 2\kappa\pi, \kappa \in N^+$, using the integration and difference formula yields:

$$\cos\omega t \sin\frac{\omega}{2} = \frac{1}{2}\left(\sin\frac{2t+1}{2}\omega - \sin\frac{2t-1}{2}\omega\right)$$

Let $t = 1,2,3,\ldots,T$, summing up the above equations yields:

$$\sin\frac{\omega}{2}\sum_{t=1}^{T}\cos\omega t = \frac{1}{2}\left(\sin\frac{2T+1}{2}\omega - \sin\frac{\omega}{2}\right) = \cos\frac{T+1}{2}\omega\sin\frac{T\omega}{2}$$

Dividing both sides by $\sin\frac{\omega}{2}$, then we have:

$$\sum_{t=1}^{T}\cos\omega t = \frac{\cos\frac{T+1}{2}\omega\sin\frac{T\omega}{2}}{\sin\frac{\omega}{2}} \quad ,\sum_{t=1}^{T}\sin\omega t = \frac{\sin\frac{T+1}{2}\omega\sin\frac{T\omega}{2}}{\sin\frac{\omega}{2}}$$

Based on the trigonometric function induction formula, the following simplified formula can be obtained:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}\cos(\omega t+\varphi) = \lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}(\cos\omega t\cos\varphi - \sin\omega t\sin\varphi)$$

$$= \lim_{T\to+\infty}\frac{1}{T}\frac{\sin\frac{T\omega}{2}}{\sin\frac{\omega}{2}}\left(\cos\frac{T+1}{2}\omega\cos\varphi - \sin\frac{T+1}{2}\omega\sin\varphi\right) = 0 \quad ,\omega \neq 2\kappa\pi$$

On the other hand, when the tolerance T_1 of the independent variable is equal to the integral times of the trigonometric function's period T_2, we have:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}\cos\left[\frac{2\pi}{T_2}(T_1 t+\varphi_1)+\varphi_2\right] = \lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}\cos\left[2\kappa t\pi + (\frac{2\pi}{T_2}\varphi_1+\varphi_2)\right]$$

$$= \lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}\cos(\frac{2\pi}{T_2}\varphi_1+\varphi_2) = \cos(\frac{2\pi}{T_2}\varphi_1+\varphi_2) \quad ,\exists\kappa\in N^+, T_1 = \kappa T_2$$

This completes the proof of Proposition 1.

Appendix C. Proof of Corollary 1

Proof: According to the Fourier theory, any periodic function can be formulated as the sum of several trigonometric functions (totally E) and a constant A_0, and the constant is the mean of the periodic function in one period:

$$f(t) = A_0 + \sum_{\varepsilon=1}^{E}A_\varepsilon\cos(\frac{2\pi}{T_2}\varepsilon t+\varphi_\varepsilon) = A_0 + \sum_{\varepsilon=1}^{E}A_\varepsilon\cos(\frac{2\pi}{T_2/\varepsilon}t+\varphi_\varepsilon) \quad ,A_0 = \frac{1}{T_2}\int_0^{T_2}f(t)\mathrm{d}t$$

When T_1 is not equal to the integral times of T_2, it is also not equal to the integral times of T_2/ε that is the period of the trigonometric function.

$$T_1 \neq \kappa T_2 \quad ,\forall\kappa\in N^+ \quad \Rightarrow \quad T_1 \neq \kappa T_2/\varepsilon \quad ,\forall\kappa,\varepsilon\in N^+$$

Therefore, we have that:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}A_\varepsilon\cos\left[\frac{2\pi}{T_2}\varepsilon(T_1t+\varphi_1)+\varphi_\varepsilon\right]=\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}A_\varepsilon\cos\left[\frac{2\pi}{T_2/\varepsilon}(T_1t+\varphi_1)+\varphi_\varepsilon\right]=0$$

And

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}f(T_1t+\varphi_1)=A_0+\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{\varepsilon=1}^{E}A_\varepsilon\cos(\frac{2\pi}{T_2}\varepsilon t+\varphi_\varepsilon)=A_0+0=A_0$$

On the other hand, when T_1 is equal to the integral times of T_2, we have:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}f(T_1t+\varphi_1)=\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}f(\kappa T_2t+\varphi_1)=\lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}f(\varphi_1)=f(\varphi_1)$$

This completes the proof of Corollary 1.

Acknowledgements

This work is jointly supported by the National Science Foundation of China (Project No. 72071079, 71890972, 71890970).

References

Assadipour G., Ke G., Verma M., 2015. Planning and managing intermodal transportation of hazardous materials with capacity selection and congestion. Transportation Research Part E, 76, 45-57.

Batta R, Chiu S., 1988. Optimal obnoxious paths on a network: Transportation of hazardous materials. Operations Research, 36(1), 84-92.

Beneventti G D., Bronfman A., Germán P., Vladimir M., 2019. A multi-product maximin hazmat routing-location problem with multiple origin-destination pairs. Journal of Cleaner Production, 240: 118193.

Barth M., Younglove T., Scora G., 2005. Development of a heavy-duty diesel modal emissions and fuel consumption model. California Partners for Dominant Transit and Highways (PATH), UC Berkeley: Technical Report.

Bertzzi L., Paletta G., Speranza M.G., 2002. Deterministic order-up-to level policies in an inventory routing problem. Transportation Science, 36(1), 119-132.

Bronfman A., Marianov V., Paredes-Belmar G., Lüer-Villagra A., 2016. The maxisum and maximin-maxisum HAZMAT routing problems. Transportation Research Part E, 93, 316-333.

Bronfman, A., Paredes-Belmar G., Marianov V., 2019. A multi-product maximin hazmat routing-location problem with multiple origin-destination pairs. Journal of Cleaner Production, 240, 118193.

Carotenuto P., Giordani S., Ricciardelli S., 2007. Finding minimum and equitable risk routes for hazmat shipments. Computers & Operations Research, 34, 1304-1327.

Chen C., Tian Z., Yao B., 2017. Optimization of two-stage location-routing-inventory problem with time-windows in food distribution network. Annuals of Operation Research, 273(1-2), 111-134.

Dadkar Y., Nozick L., Jones D., 2010. Optimizing facility use restrictions for the movement of hazardous

materials. Transportation Research Part B, 44, 267-281.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE transactions on evolutionary computation, 6(2), 182-197.

Erkut E, Tjandra SA, Verter V. Hazardous materials transportation. Handbook in OR&MS, 2007, 14, 539-621.

Fan T., Chiang W., Russell R., 2015. Modeling urban hazmat transportation with road closure consideration. Transportation Research Part D, 35, 104-115.

Fontaine P., Minner S., 2018. Benders decomposition for the hazmat transport network design problem. European Journal of Operational Research, 267, 996-1002.

Fontaine, P., Crainic, T. G., Gendreau, M., Minner, S., 2020. Population-based risk equilibration for the multimode hazmat transport network design problem. European Journal of Operational Research, 284(1), 188-200.

Gopalan R., Kolluri K.S., Batta R., Karwan M.H., 1990. Modeling equity of risk in the transportation of hazardous materials. Operations Research, 38(6), 961-973.

Ghaderi A., Burdett R.L., 2019. An integrated location and routing approach for transporting hazardous materials in a bi-modal transportation network. Transportation Research Part E, 127, 49-65.

Huang Y., Zhao L., Woensel T V., Gross J P., 2017. Time-dependent vehicle routing problem with path flexibility. Transportation Research Part B, 95, 169-195.

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. Computers&Industrial Engineering, 128, 861-876.

Hosseini S.D., Verma M., 2017. A Value-at-Risk (VAR) approach to routing rail hazmat shipments. Transportation Research Part D, 54, 191-211.

Hosseini S.D., Verma M., 2018. Conditional value-at-risk (CVaR) methodology to optimal train configuration and routing of rail hazmat shipment. Transportation Research Part B, 110, 79-103.

Javid, A.A., Azad, N., 2010. Incorporating location, routing and inventory decisions in supply chain network design. Transportation Research Part E, 46(5), 582-597.

Kara B. Y., Verter V. Designing a road network for hazardous materials transportation. Transportation Science, 2004, 38(2), 188-196.

Kheirkhah A., Navidi H., Bidgoli M., 2015. A bi-level network interdiction model for solving the hazmat routing problem, International Journal of Production Research, 54(2), 459-471.

Kourniotis S.P., Kiranoudis C.T., Markatos N.C., 2000. Statistical analysis of domino chemical accidents. Journal of Hazardous Materials, 71(1-3), 239-252.

Kumar A., Roy D., Verter V., Sharma D., 2018. Integrated fleet mix and routing decision for hazmat transportation: A developing country perspective. European Journal of Operational Research, 264, 225-238.

Mahmoudsoltani F., Shahbandarzadeh H., Moghdani R., 2018. Using Pareto-based multi-objective Evolution algorithms in decision structure to transfer the hazardous materials to safety storage centre. Journal of Cleaner Production, 184, 893-911.

Mohri S.S., Asgari N., Farahani R.Z., Bourlakis M., Laker B., 2020. Fairness in hazmat routing-scheduling: A bi-objective Stackelberg game. Transportation Research Part E, 140, 102006.

Nicolet-Monnier M., Gherghe A.V., 1995. Quantitative risk assessment of hazardous materials transport systems. Kluwer, 5(8): 1915-1928.

Paredes-Belmar G., Bronfman A., Marianov V., Latorre-Núnez, G., 2017. Hazardous materials collection with multiple-product loading, Journal of Cleaner Production, 141, 909-919.

Reilly A., Nozick L., Xu N., Jones D., 2012. Game theory-based identification of facility use restrictions for the movement of hazardous materials under terrorist threat. Transportation Research Part E, 48, 115-131.

Romero N., Nozick L.K., Xu N., 2016. Hazmat facility location and routing analysis with explicit consideration of equity using the Gini coefficient. Transportation Research Part E, 89, 165-181.

Rabbani M., Heidari R., Yazdanparast R., 2019. A stochastic multi-period industrial hazardous waste location-routing problem: Integrating NSGA-II and Monte Carlo simulation. European Journal of Operational Research, 272, 945-961.

Samanlioglu F., 2013. A multi-objective mathematical model for the industrial hazardous waste location-routing problem. European Journal of Operational Research, 226, 332-340.

Szeto W.Y., Farahani R.Z., Sumalee A., 2017. Link-based multi-class hazmat routing-scheduling problem: A multiple demon approach. European Journal of Operational Research, 261, 337-354.

Verma M., Verter V., Gendreau M., 2011. A tactical planning model for railroad transportation of dangerous goods. Transportation Science, 45(2), 163-174.

Wu W., Zhou W., Lin Y., Xie Y., Jin W. A hybrid metaheuristic algorithm for location inventory routing problem with time windows and fuel consumption. Expert Systems With Applications, 2021, 166, 114034.

Xie Y., Lu W., Wang W., Quadrifoglio L., 2012. A multimodal location and routing model for hazardous materials transportation. Journal of Hazardous Materials, 227-228, 135-141.

Zhao J., Huang L., Lee D.H., Peng Q., 2016. Improved approaches to the network design problem in regional hazardous waste management systems. Transportation Research Part E, 88, 52-75.

Zitzler E, Deb K, Thiele L., 2000. Comparison of multi-objective evolutionary algorithms: Empirical study. Evolutionary Computation, 8(2), 173-195.