



# A hierarchical hub location model for the integrated design of urban and rural logistics networks under demand uncertainty

Zhi-Chun Li<sup>1</sup> · Xue Bing<sup>1</sup> · Xiaowen Fu<sup>2</sup> 

Accepted: 16 January 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

This paper contributes to the integrated design issue of urban and rural logistics networks under demand uncertainty. A hierarchical hub location model is proposed, which minimizes the expected total system cost by optimizing the locations, number and capacities of “urban–town–village” hierarchical logistics hubs. The interactions among the logistics hubs and among the hub–and–spoke connections, as well as the hub capacity constraints are explicitly considered in the presence of logistics demand uncertainty. A demand scenario–based branch–and–Benders–cut algorithm is developed to solve the proposed model. A case study of Jiangling urban–rural region in Hubei province of China is conducted for the illustration of the model and solution algorithm. The results generated by the proposed algorithm are benchmarked against those obtained by GUROBI solver and the practical scheme being currently implemented in the region. The results showed that the proposed methodology can greatly improve the efficiency of the urban–rural logistics system in terms of expected total system cost. It is important to explicitly model the demand uncertainty, otherwise a significant decision bias may emerge. The proposed algorithm outperforms the GUROBI solver in terms of problem size solved and computational time.

**Keywords** Integrated design · Urban and rural logistics networks · Hub–and–spoke network · Hierarchical hub location · Demand uncertainty · Stochastic optimization

## 1 Introduction

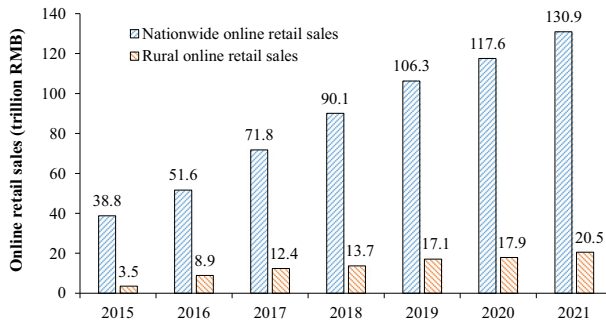
The urban and rural transportation systems, as the lifeline of urban and rural development, link all aspects of production, exchange, distribution, and consumption. With rapid developments of e–business and urbanization in developing countries such as China, logistics and

---

✉ Xiaowen Fu  
xiaowen.fu@polyu.edu.hk

<sup>1</sup> School of Management, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>2</sup> Department of Industrial and Systems Engineering, Hong Kong Polytechnic University, Hong Kong, China



**Fig. 1** Nationwide vs. rural online retail sales in China during 2015–2021

shipping demands between urban and rural regions have been growing fast, thanks to continuously improved transportation infrastructures. According to the annual report of agricultural products e-commerce development (ARAPED, 2022), the average annual growth rate of the Chinese rural online retail sales in the past seven years (2015–2021) reached 28.7% (from 3.5 to 20.5 trillion RMB), as shown in Fig. 1.<sup>1</sup> Both the nationwide and rural online retail sales had achieved rapid growth in the past years despite the COVID–19 pandemic. Online retail sales in rural regions account for a significant proportion of the nation’s total online retail sales. To accommodate the rapidly growing logistics demands between urban and rural regions, the supporting logistics network needs to be upgraded for efficient urban–rural cargo flow.

In this regard, the Chinese government is speeding up the integration of rural e-commerce and express delivery systems to better link production and consumption between urban and rural areas. According to the statistics recently issued by the State of Post Bureau of China, express delivery services now cover almost all towns in China, with over 37 billion express packages shipped to and from China’s rural regions in 2021. The associated exchanges of industrial and agricultural products between urban and rural areas reached 1.85 trillion RMB (or US\$291 billion). However, there are still some prominent problems in urban and rural deliveries, such as the “last-mile” problem and insufficient village logistics service points. Coordinated development between urban and rural logistics systems is a promising solution to these issues so as to realize rural revitalization and common prosperity.

In the past decades, the urban logistics industry received considerable investments and attention in many countries and areas around the world (Lagorio et al., 2016; Savelsbergh and Van Woensel, 2016; Buyukozkan & Ilicak, 2022). However, little attention was paid to the rural logistics industry. This may be attributed to the fact that the demand of rural logistics is more dispersed and often less-developed than that of urban logistics. Moreover, the design and operation of urban and rural networks are usually implemented separately. As a result, the logistics costs for shipments between urban and rural regions are usually high, and the associated logistics efficiency and inter-regional connectivity tend to be low. In order to address these issues, the Chinese government has recently implemented the program of “Integrated Urban–Rural Transportation System Development” (MOT, 2017). So far, a total of 113 counties have been selected as the first two batches of pilot areas in this program.<sup>2</sup> This

<sup>1</sup> RMB is the Chinese currency “Renminbi”. US\$1 approximates RMB6.35 as of January 1, 2022.

<sup>2</sup> [https://xxgk.mot.gov.cn/2020/jigou/ysfws/202006/t20200623\\_3315384.html](https://xxgk.mot.gov.cn/2020/jigou/ysfws/202006/t20200623_3315384.html); [https://xxgk.mot.gov.cn/2020/jigou/ysfws/202111/t20211102\\_3624076.html](https://xxgk.mot.gov.cn/2020/jigou/ysfws/202111/t20211102_3624076.html).

provides a favorable condition for the development of an integrated hierarchical urban–rural logistics network linking urban, town and village areas, on a basis of hub–and–spoke (H&S) network configuration. For example, with this program, Jiangling county’s government in Hubei province of central China has built 2 urban hubs, 4 town hubs, and 65 village hubs to provide the logistics services to a total of 8 towns and 107 villages in that county.<sup>3</sup> In general, there are two types of nodes in a typical H&S network: hub node and spoke node. The location of a hub largely determines the express delivery cargo distribution and the connections between the hub and spoke nodes (Shang et al., 2021; Zhong et al., 2018). In a typical urban–rural logistics network, there are three types of hubs and four kinds of hierarchical connections. The three hub types include urban logistics hubs, town logistics hubs and village logistics hubs. Accordingly, there are four connection layers, namely the top layer connecting urban logistics hubs, the second layer connecting town logistics hubs and urban logistics hubs, the third layer connecting village logistics hubs and town logistics hubs, and the bottom layer connecting spokes and hubs (urban, town or village hubs).

With the rapid development of rural e–commerce activities, the rural logistics demand has been growing rapidly in recent years. However, logistics demand usually fluctuates by the time of day, day of week, and seasons due to factors such as economic growth, population change, information technology improvement, weather, and various disruptions such as the COVID–19 pandemic, the effects of which cannot be accurately predicted. The uncertainty in the logistics demand could significantly affect the service price and cost, and the efficiency of the logistics network. Apparently, a frequent adjustment of the urban–rural logistics network is infeasible, partly due to the high cost that it can cause (e.g., it can be very costly to set up a new hub and the feeder networks even if it is possible to re–deploy equipment to another location). Therefore, it is important to optimize the design of the integrated urban–rural logistics network in the presence of uncertain demand.

In light of the above research needs, this paper addresses the challenge of integrated design of the urban–rural logistics networks under demand uncertainty. The main contributions of this study are two–fold. First, a hierarchical hub location model is proposed for the design of the urban–rural logistics networks, which minimizes the expected total system cost by determining the locations, number, and capacities of urban, town and village hubs. Different from the traditional urban logistics network design, the model proposed in this paper is an integrated optimization of the urban and rural logistics networks, in which the interactions among “urban–town–village” logistics hubs, and among hub–spoke connections in the urban–rural logistics system are considered. The effects of hub capacity constraints for each hub type are also incorporated, together with the effects of logistics demand uncertainty using a demand scenario–based approach. Second, a demand scenario–based branch–and–Benders–cut algorithm is developed to solve the proposed model. A case study of Jiangling urban–rural region in Hubei region of China is carried out for illustration purpose. The results generated by the proposed algorithm are compared with those obtained by GUROBI solver and the practical scheme being implemented in the region.

The remainder of this paper is organized as follows. In the next section, some related literature is reviewed. Section 3 formulates the model for the integrated design of the urban–rural logistics network. Section 4 presents the branch–and–Benders–cut algorithm. In Sect. 5, a case study is provided to illustrate the application of the proposed model and solution algorithm. Section 6 concludes this paper and provides suggestions for further studies.

<sup>3</sup> [http://nyj.jingzhou.gov.cn/ztzl\\_16/xcxz/202009/t20200903\\_518386.shtml](http://nyj.jingzhou.gov.cn/ztzl_16/xcxz/202009/t20200903_518386.shtml).

## 2 Literature review

The integration of urban–rural transportation networks plays a crucial role in the new urbanization and rural revitalization in China, and has attracted widespread attention in the last decade. For example, the studies of Yu et al. (2013), Zhong et al. (2018) and Lei et al. (2022) addressed the public transit hub location problems between urban and rural regions. Wang and Sun (2016) analyzed the relation of transportation infrastructure and rural development in China, and found that investment in transport infrastructure has a positive impact on China's rural development. Qu et al. (2022) pointed out that low-cost, accessible transportation systems connecting rural and urban areas are crucial for creating an efficient rural–urban partnership. However, these studies mainly focused on public passenger transport systems, whereas little attention has been paid to the design of urban–rural logistic networks.

The design of urban–rural logistics network can be considered as an extension of common hub location problems, also referred to as H&S network design problems in the literature. Among the hub location problems, the  $p$ -hub median problem and  $p$ -hub center problem are the most prevalent specifications (Campbell, 1994; Yaman, 2011). Since the pioneering work of O'Kelly (1986) in modeling of the  $p$ -hub median problem, the hub location problems have been well studied in the literature (Contreras et al., 2012; Contreras and O'Kelly, 2019; Alumur et al., 2021). For the convenience of readers, we have summarized in Table 1 some related studies about the hierarchical hub location problems, in terms of network structure, decision variables, uncertainty source and solution method.

It can be observed in Table 1 that the previous studies have mainly focused on hierarchical hub network (without direct spoke–spoke links) and hybrid hub network (allowing direct spoke–spoke links). Zhao et al. (2016) and Yao et al. (2019) proposed a hybrid H&S express network decision model and found that compared with the pure H&S structure, the hybrid H&S network can contribute to cost reduction, and at the same time can decrease detour and increase the timeliness and service level. The  $p$ -median problem in the hierarchical hub location models aims to determine the hub locations in a  $k$ -level network to minimize the total cost (Wang et al., 2020; Zhong et al., 2018). Shang et al. (2021) proposed a model for the design of a hierarchical urban–rural public passenger transport hub network. In addition to the difference in network structures, some studies have also considered other constraints. For example, Dukkanci and Kara (2017) considered travel time bounds in the design of a hierarchical H&S network. Alumur et al. (2018) presented a hub location model considering the service time limit and congestion at hubs. Wang et al. (2021) proposed a hub location model of agricultural product transportation network for exploring the impact of delivery time limit on the hub location. Kaveh et al. (2021) proposed a hub location model with an objective function of minimizing the total transportation and waiting time of the network. A limitation of these studies is the assumption of no direct link between spokes. Note that direct transportation may be more economical than connection via hubs when the logistics demand between the two spokes is sufficiently large. Therefore, in this paper a hierarchical H&S network allowing direct connections between spokes (i.e., a hierarchical hybrid H&S network) is analyzed.

Table 1 also shows that the existing studies on the hub location problem mainly considered the number and location of hubs as decision variables, whereas the capacity of hub has usually been ignored or pre-given. In general, hub capacity should meet the logistic demand for that hub, and thus it is necessary to consider hub capacity constraints in the model (Fotuhi & Huynh, 2018; Mišković et al., 2017). On the other hand, the capacities of hubs generally have a significant effect on the hub locations and flow allocation, and thus the locations

**Table 1** Summary of relevant studies on network design

References	Network structure	Decision variables	Uncertainty source	Solution method
Correia et al. (2010)	Hybrid H&S	Capacity of hub	Deterministic	CPLEX
Zhao et al. (2016)	Hybrid H&S	Location and allocation of hub	Deterministic	Heuristic
Dukkanci and Kara (2017)	Hierarchical H&S	Number, location and allocation of hub	Deterministic	Heuristic
Zokaee et al. (2017)	Hierarchical network	Location and allocation of hub	Demand, capacity, hub construction and transportation cost	Lingo
Mišković et al. (2017)	Hierarchical network	Capacity of hub	Transportation cost	Heuristic
Fotuhi et al. (2018)	Complete network	Number, capacity and location of hub	Demand	Heuristic
Zhong et al. (2018)	Hierarchical hybrid H&S	Number and location of hub	Deterministic	Heuristic
Alumur et al. (2018)	Hybrid H&S	Number and location of hub	Deterministic	CPLEX
Irawan and Jones (2019)	H&S	Capacity and location of hub	Deterministic	CPLEX
Shavarani et al. (2019)	Hierarchical H&S	Number and location of hub	Demand, hub construction cost, distance capacity of drones	Heuristic
Yao et al. (2019)	Hybrid H&S	Number and location of hub	Deterministic	Heuristic
Wang et al. (2020)	Hierarchical H&S	Number and location of hub	Demand, Transportation cost	GUROBI
Shang et al. (2021)	H&S	Number and location of hub	Travel time	Heuristic
Wang et al. (2021)	Hierarchical H&S	Location of hub	Deterministic	Heuristic
Kaveh et al. (2021)	Hierarchical H&S	Number and location of hub	Transportation cost, capacity, transportation and waiting time	Heuristic

and capacities of hubs should be jointly determined. In this regard, Correia and Captivo (2003), Correia et al. (2010) and Irawan and Jones (2019) addressed capacity–constrained hub location problems with multiple capacity levels, in which hub capacity is a decision variable. Yet, they did not propose an effective exact algorithm to solve the capacitated hub location problem, and the uncertainty is not incorporated.

In addition, there are some studies that considered uncertain hub location problems. In reality, hub location is generally a long–term decision, and the uncertainty in demand side and/or supply side may lead the current location decisions to be inefficient for future logistics demand. However, frequent changes of hub location are unrealistic due to substantial costs

involved. The uncertain facility location problems can be classified into the stochastic facility location problems and the robust facility location problems (Soyster, 1973; Shang et al., 2021). Both can capture the effects of uncertainty via various possible scenarios. The scenario probability is known in the stochastic optimization models. However, such a probability may be unknown in the robust optimization models. Robust location models are usually suitable for the situations in which it is difficult to collect data, or there is a lack of data about the probability distribution of the uncertain parameters (Shavarani et al., 2019; Zokaei et al., 2017).

With the literature review above, it is clear that the existing relevant studies mainly focused on the passenger public transport network and/or the urban transport network, with little attention paid to the modelling of comprehensive urban and rural logistics network. The integrated design of the urban and rural logistics networks are especially important for the mobility of human and goods between the two types of regions, the development of rural economy, and the reduction in the gap between urban and rural region development. Moreover, most studies only considered the locations of hubs, but not the hub capacity as decision variable, which can be an important issue in the design of urban–rural hierarchical network.

### 3 Model formulation

#### 3.1 Problem description

This paper aims to contribute to the integrated design of urban and rural logistics networks under demand uncertainty. The locations, number and capacities of the hubs in the urban and rural logistics networks will be optimized to minimize the expected total system cost. For illustration purpose, Fig. 2 shows a typical hierarchical urban–rural logistics network system with 19 hierarchical hubs (including 3 urban hubs, 6 town hubs and 10 village hubs) and 11 spoke nodes (i.e., small villages in rural region). Different from the traditional urban logistics network design, this paper focuses on an integrated optimization of the urban and rural logistics networks, in which the interactions among hierarchical “urban–town–village” logistics hubs are incorporated. The locations and capacities of multi–type hubs are thus endogenously determined by the model proposed in this paper. The hub capacity constraints

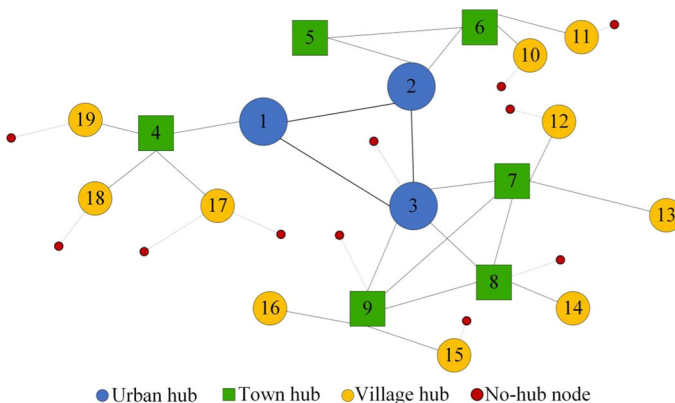


Fig. 2 A typical hierarchical urban–rural logistics network

for each hub type, which require that the logistics demand of each hub does not exceed its capacity, are also considered. In addition, the effects of logistics demand uncertainty are also modeled through using a demand scenario-based approach.

To facilitate the presentation of the essential ideas without losing generality, the following basic assumptions are made in this paper.

**A1** The urban hub network is a complete network, with each pair of urban hubs connected directly. This assumption has been widely adopted in the literature, e.g., see Zhong et al. (2018).

**A2** Town hubs connect to their associated urban hub, and the town hubs with the same urban hub connect with each other directly (e.g., see Alumur et al., 2012).

**A3** Each village hub is only assigned to a town hub, and each town hub is only assigned to an urban hub (e.g., see Shang et al., 2021).

**A4** Each spoke (i.e., non-hub) node connects to one hub only. Direct connections between spoke nodes are not allowed (e.g., see Yaman, 2011).

### 3.2 Definitions of parameters and variables

#### (i) Sets

$N$ : Set of all nodes.

$P$ : Set of candidate urban hub locations,  $P \subseteq N$ .

$R$ : Set of candidate village hub locations,  $R \subseteq N$ .

$Q$ : Set of candidate town hub locations,  $Q \subseteq N$ ,  $Q \cap R = \emptyset$ .

$H$ : Set of all candidate hub locations,  $H = P \cup R \cup Q$ .

$Q_k$ : Set of alternative capacity levels for urban hub  $k \in P$ .

$Z_m$ : Set of alternative capacity levels for candidate town hub  $m \in Q$ .

$T_e$ : Set of alternative capacity levels for candidate village hub  $e \in R$ .

$S$ : Set of all stochastic demand scenarios,  $s \in S$ .

#### (ii) Parameters

$p_s$ : Probability of demand scenario  $s$ , with  $\sum_{s \in S} p_s = 1$ .

$w_{ij}^s$ : Logistics demand between origin-destination (O-D) pair  $ij$  under demand scenario  $s$ .

$D_{ij}$ : Distance between O-D pair  $ij$ .

$\Lambda$ : Transportation cost per unit of cargo per km.

$c_{ij}$ : Transportation cost between O-D pair  $ij$  per unit of cargo, with  $c_{ij} = \Lambda D_{ij}$ .

$\alpha_1$ : Discount coefficient of transportation cost per unit of cargo on connections between urban hubs.

$\alpha_2$ : Discount coefficient of transportation cost per unit of cargo on connections between town hubs and between urban hubs and town hubs.

$\alpha_3$ : Discount coefficient of transportation cost per unit of cargo on connections between village hubs and town hubs.

$C_{kq}$ : Construction cost of candidate village hub  $k$  with capacity level  $q \in Q_k$ .

$C_{mz}$ : Construction cost of candidate town hub  $m$  with capacity level  $z \in Z_m$ .

$C_{et}$ : Construction cost of candidate village hub  $e$  with capacity level  $t \in T_e$ .

$M_{kq}$ : Capacity of candidate village hub  $k$  with capacity level  $q \in Q_k$ .

$M_{mz}$ : Capacity of candidate town hub  $m$  with capacity level  $z \in Z_m$ .

$M_{et}$ : Capacity of candidate village hub  $e$  with capacity level  $t \in T_e$ .

$P'_k$ : Minimum number of urban hubs to be constructed.

$P_k$ : Maximum number of urban hubs to be constructed.

$P'_m$ : Minimum number of town hubs to be constructed.

$P_m$ : Maximum number of town hubs to be constructed.

$P'_e$ : Minimum number of village hubs to be constructed.

$P_e$ : Maximum number of village hubs to be constructed.

$\omega$ : Penalty coefficient.

(iii) *Decision variables*

$h_{kq}$ : 0–1 variable, equal to 1 if urban hub is located at node  $k$  with capacity level  $q$ , and 0 otherwise.

$d_{ez}$ : 0–1 variable, equal to 1 if town hub is located at node  $e$  with capacity level  $z$ , and 0 otherwise.

$b_{mt}$ : 0–1 variable, equal to 1 if village hub is located at node  $m$  with capacity level  $t$ , and 0 otherwise.

$x_{ie}$ : 0–1 variable, equal to 1 if node  $i$  is assigned to hub  $e$ , and 0 otherwise; if node  $e$  is a village hub, then  $x_{ee} = 1$ .

$u_{vm}$ : 0–1 variable, equal to 1 if village hub  $v$  is assigned to town hub  $m$ , and 0 otherwise; if node  $m$  is a village hub, then  $u_{mm} = 1$ .

$y_{mk}$ : 0–1 variable, equal to 1 if town hub  $m$  is assigned to urban hub  $k$ , and 0 otherwise; if node  $k$  is a urban hub, then  $y_{kk} = 1$ .

$z_{mn}$ : 0–1 variable, equal to 1 if there is a connection between town hub  $m$  and non–village hub  $n$ , and 0 otherwise.

$a_{vm}^{is}$ : Logistics volume from village hub  $v$  to town hub  $m$  which originates at node  $i$  under demand scenario  $s$ .

$g_{mn}^{is}$ : Logistics volume from town hub  $m$  to non–village hub  $n$  which originates at node  $i$  under demand scenario  $s$ .

$f_{kl}^{is}$ : Logistics volume from urban hub  $k$  to town hub  $l$  which originates at node  $i$  under demand scenario  $s$ .

$\eta_{se}$ : Logistics volume exceeding capacity of village hub  $e$  under demand scenario  $s$ .

$\eta_{sm}$ : Logistics volume exceeding capacity of town hub  $m$  under demand scenario  $s$ .

$\eta_{sk}$ : Logistics volume exceeding capacity of urban hub  $k$  under demand scenario  $s$ .

### 3.3 The model

Based on the assumptions and notations presented in the previous sections, the stochastic model for the urban–rural logistics hub network design is presented as follows:

$$\begin{aligned} \min F = & \sum_{z \in Z_m} \sum_{m \in Q} C_{mz} d_{mz} + \sum_{t \in T_e} \sum_{e \in R} C_{et} b_{et} + \sum_{q \in Q_k} \sum_{k \in P} C_{kq} h_{kq} \\ & + \sum_{s \in S} p_s \sum_{i \in N} \sum_{j \in N} (w_{ij}^s + w_{ji}^s) \sum_{e \in H} C_{ie} x_{ie} \\ & + \sum_{s \in S} p_s \sum_{i \in N} \sum_{k \in P} \sum_{l \in P \setminus \{k\}} \alpha_1 c_{kl} f_{kl}^{is} + \sum_{s \in S} p_s \sum_{i \in N} \sum_{m \in Q} \sum_{k \in P \cup Q \setminus \{m\}} \alpha_2 c_{mk} g_{mk}^{is} \\ & + \sum_{s \in S} p_s \sum_{i \in N} \sum_{v \in R} \sum_{m \in Q \setminus \{v\}} \alpha_3 c_{vm} a_{vm}^{is} \\ & + \omega_1 \sum_{s \in S} p_s \sum_{e \in R} \eta_{se} + \omega_2 \sum_{s \in S} p_s \sum_{m \in Q} \eta_{sm} + \omega_3 \sum_{s \in S} p_s \sum_{k \in P} \eta_{sk} \end{aligned} \quad (1)$$

$$\text{s.t. } \sum_{e \in H} x_{ie} = 1, \quad \forall i \in N \quad (2)$$



$$x_{ie} \leq x_{ee}, \forall i \in N, e \in H \tag{3}$$

$$\sum_{m \in Q} u_{vm} = x_{vv}, \forall v \in R, t \in T \tag{4}$$

$$u_{vm} \leq u_{mm}, \forall m \in Q, \forall v \in R, z \in Z \tag{5}$$

$$\sum_{k \in P} y_{mk} = u_{mm}, \forall m \in Q, z \in Z \tag{6}$$

$$y_{mk} \leq y_{kk}, \forall m \in Q, k \in P \tag{7}$$

$$\sum_{t \in T_e} b_{vt} \leq 1, \forall v \in R \tag{8}$$

$$\sum_{z \in Z_m} d_{mz} \leq 1, \forall m \in Q \tag{9}$$

$$\sum_{q \in Q_k} h_{kq} \leq 1, \forall k \in P \tag{10}$$

$$z_{mn} \geq y_{mk} + y_{nk} - 1, \forall m \in Q, n \in Q \setminus \{m\}, k \in P, k \neq m, k \neq n \tag{11}$$

$$2z_{mn} \leq y_{mk} + y_{nk}, \forall m \in Q, n \in Q \setminus \{m\}, k \in P, k \neq m \tag{12}$$

$$z_{mn} = e_{nm}, \forall m \in Q, n \in Q \setminus \{m\} \tag{13}$$

$$z_{ik} \geq y_{ik}, \forall i \in P \cup Q, k \in P \setminus \{i\}, s \in S \tag{14}$$

$$\sum_{m \in Q \setminus \{j\}} a_{jm}^{is} - \sum_{m \in Q \setminus \{j\}} a_{mj}^{is} = \sum_{r \in N} w_{ir}^s (x_{ij} - x_{rj}), \forall i \in N, j \in R, s \in S \tag{15}$$

$$\begin{aligned} & \sum_{n \in P \cup Q \setminus \{j\}} g_{mn}^{is} + \sum_{j \in R} a_{mj}^{is} - \sum_{j \in R} a_{jm}^{is} - \sum_{n \in P \cup Q \setminus \{j\}} g_{nm}^{is} \\ &= \sum_{r \in N} w_{ir}^s (x_{im} - x_{rm}), \forall i \in N, m \in Q, s \in S \end{aligned} \tag{16}$$

$$\begin{aligned} & \sum_{l \in Q \setminus \{j\}} g_{jl}^{is} + \sum_{k \in P \setminus \{j\}} f_{jk}^{is} - \sum_{k \in P \setminus \{j\}} f_{kj}^{is} - \sum_{l \in Q \setminus \{j\}} g_{lj}^{is} \\ &= \sum_{r \in N} w_{ir}^s (x_{ij} - x_{rj}), \forall i \in N, j \in P, s \in S \end{aligned} \tag{17}$$

$$\sum_{i \in N} \sum_{n \in Q} a_{ne}^{is} + \sum_{i \in N} \sum_{j \in N} w_{ij}^s x_{ie} \leq \sum_{t \in T_e} M_{et} b_{et} + \eta_{se}, \forall e \in R, s \in S_e \tag{18}$$

$$\sum_{i \in N} \sum_{v \in R} a_{vn}^{is} + \sum_{i \in N} \sum_{m \in P \cup Q} g_{mn}^{is} + \sum_{i \in N} \sum_{j \in N} w_{ij}^s x_{in} \leq \sum_{z \in Z_m} M_{nz} d_{nz} + \eta_{sn}, \forall n \in Q, s \in S \tag{19}$$

$$\sum_{i \in N} \sum_{m \in Q} g_{mk}^{is} + \sum_{i \in N} \sum_{l \in P} f_{lk}^{is} + \sum_{i \in N} \sum_{j \in N} w_{ij}^s x_{ik} \leq \sum_{q \in Q_k} M_{kq} h_{kq} + \eta_{sk}, \forall k \in P, s \in S \tag{20}$$

$$P'_m \leq \sum_{z \in Z_m} \sum_{m \in Q} d_{mz} \leq P_m \tag{21}$$

$$P'_e \leq \sum_{t \in T_e} \sum_{e \in R} b_{et} \leq P_e \tag{22}$$

$$P'_k \leq \sum_{q \in Q_k} \sum_{k \in P} h_{kq} \leq P_k \tag{23}$$

$$g_{jl}^{is} \leq z_{jl} \sum_{r \in N} w_{jr}^s, \forall i \in N, j \in Q, l \in P \cup Q \setminus \{j\}, s \in S \quad (24)$$

$$b_{et}, d_{nz}, h_{kq}, x_{ie}, u_{vm}, y_{mk}, z_{mn} \in \{0, 1\}, \forall i \in N, e \in H, v \in R, m \in H, k \in P, n \in H \setminus \{m\} \quad (25)$$

$$a_{vm}^{is}, g_{mn}^{is}, f_{kl}^{is}, \eta_{se}, \eta_{sn}, \eta_{sk} \geq 0, \forall i \in N, m \in H, k \in P, n \in H \setminus \{m\}, l \in P \setminus \{k\}, s \in S \quad (26)$$

The objective function Eq. (1) represents the total cost of the urban–rural logistics system under all the demand scenarios. The cost includes hub construction costs, transportation costs and penalty costs. The first three terms are the construction costs of village hubs, town hubs and urban hubs, respectively. The fourth term is the transportation costs between the no–hub nodes and hubs. The fifth term is the transportation costs between the urban hubs. The sixth term is the transportation costs between the town hubs and between the urban hubs and the town hubs. The seventh term is the transportation costs between the town hubs and the village hubs. If the demand of a hub exceeds its capacity, then a penalty is incurred. The final three terms are the penalty costs due to violations of the capacity constraints of urban, town, and village hubs, respectively.

Constraints (2) and (3) indicate that each non–hub node is assigned to one hub. Constraints (4) and (5) indicate that each village hub is assigned to one town hub. Constraints (6) and (7) indicate that each town hub is assigned to one urban hub. Constraints (8), (9) and (10) ensure that at most one capacity level is attached to a hub. Constraints (11), (12) and (13) indicate that if town hubs  $i$  and  $j$  are assigned to the same urban hub, then a connection is formed between  $i$  and  $j$ . Otherwise,  $i$  and  $j$  are not connected. Constraint (14) represents the connection between  $i$  and  $k$  if town hub  $i$  is assigned to urban hub  $k$ . Constraint (15) indicates the flow conservations at village hubs. Constraint (16) indicates the flow conservations at town hubs. Constraint (17) indicates the flow conservations at urban hubs. Constraints (18), (19) and (20) are the capacity constraints of the village hubs, town hubs and urban hubs, respectively. Constraints (21), (22) and (23) are the boundary constraints of total number of selected hubs. Constraint (24) indicates that if there is no connection between two hubs, then the logistics quantity between them is 0. Constraint (25) is a 0–1 integer constraint. Constraint (26) is nonnegative constraint.

## 4 Solution algorithm

Benders decomposition algorithm is a typical method for solving the complex mixed integer programming problem with integer variables and coupling constraints (Tang et al., 2013; Taherkhani et al., 2020; Jenabi et al., 2022). As suggested by Contreras et al. (2012), Benders decomposition algorithm has a strong ability to solve large–scale instances of the hub location problem. In each iteration of Benders decomposition, an integer master problem must be solved and the computational time of this master problem generally increases with the number of added cuts (Rei, et al., 2009). Since Benders cuts can be generated from any master problem solution and not just from an optimal integer solution, Naoum-sawaya and Elhedhli (2013) explored the use of Benders cuts in a branch–and–cut framework, called the Branch–and–Benders–Cut (BBC) algorithm. In this paper, we use the BBC algorithm to solve the model proposed in the previous section. A multi-cut strategy is employed to improve the convergence and efficiency of the basic BBC algorithm by considering the property of multiple scenarios of the urban–rural logistics hierarchical hub location problem.

Within the standard branch-and-cut framework, the BBC algorithm proceeds by solving the Linear Relaxation of Master Problem (denoted as LRMP) at each node of the branch-and-bound tree, and valid Benders cuts are added to the model of LRMP in each iteration. The process iterates until the gap between the lower bound and the upper bound is sufficiently small, thus yielding the global optimal solution. The framework for the algorithm is formulated as follows.

Fixing the values of the binary variables  $h_{kq}$ ,  $d_{ez}$ ,  $b_{et}$ ,  $x_{ie}$ ,  $u_{vm}$ ,  $y_{mk}$ , and  $z_{mn}$  (e.g., denoted as  $\bar{h}_{kq}$ ,  $\bar{d}_{ez}$ ,  $\bar{b}_{et}$ ,  $\bar{x}_{ie}$ ,  $\bar{u}_{vm}$ ,  $\bar{y}_{mk}$ ,  $\bar{z}_{mn}$ , each of which is 0 or 1), we can obtain the Benders sub-problems, represented by SP, below:

$$\begin{aligned} \min \quad & \sum_{i \in N} \sum_{k \in P} \sum_{l \in P \setminus \{k\}} \alpha_1 c_{kl} f_{kl}^{is} + \sum_{i \in N} \sum_{m \in Q} \sum_{k \in P \cup Q \setminus \{m\}} \alpha_2 c_{mk} g_{mk}^{is} + \sum_{i \in N} \sum_{v \in R} \sum_{m \in Q \setminus \{v\}} \alpha_3 c_{vm} a_{vm}^{is} \\ & + \omega_1 \sum_{e \in R} \eta_{se} + \omega_2 \sum_{n \in Q} \eta_{sn} + \omega_3 \sum_{k \in P} \eta_{sk} \end{aligned} \tag{27}$$

$$\text{s.t.} \quad \sum_{m \in Q \setminus \{j\}} a_{jm}^{is} - \sum_{m \in Q \setminus \{j\}} a_{mj}^{is} = \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}), \forall i \in N, j \in R, s \in S \tag{28}$$

$$\begin{aligned} & \sum_{n \in P \cup Q \setminus \{j\}} g_{mn}^{is} + \sum_{j \in R} a_{mj}^{is} - \sum_{j \in R} a_{jm}^{is} - \sum_{n \in P \cup Q \setminus \{j\}} g_{nm}^{is} \\ & = \sum_{r \in N} w_{ir}^s (\bar{x}_{im} - \bar{x}_{rm}), \forall i \in N, m \in Q, s \in S \end{aligned} \tag{29}$$

$$\begin{aligned} & \sum_{l \in Q \setminus \{j\}} g_{jl}^{is} + \sum_{k \in P \setminus \{j\}} f_{jk}^{is} - \sum_{k \in P \setminus \{j\}} f_{kj}^{is} - \sum_{l \in Q \setminus \{j\}} g_{lj}^{is} \\ & = \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}), \forall i \in N, j \in P, s \in S \end{aligned} \tag{30}$$

$$\eta_{se} - \sum_{i \in N} \sum_{n \in Q} a_{ne}^{is} \geq \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ie} - \sum_{t \in T_e} M_{et} \bar{b}_{et}, \forall e \in R, e \notin Q, s \in S \tag{31}$$

$$\eta_{sn} - \sum_{i \in N} \sum_{v \in R} a_{vn}^{is} - \sum_{i \in N} \sum_{m \in P \cup Q} g_{mn}^{is} \geq \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{in} - \sum_{z \in Z_m} M_{en} \bar{d}_{ez}, \forall n \in Q, s \in S \tag{32}$$

$$\eta_{sk} - \sum_{i \in N} \sum_{m \in Q} g_{mk}^{is} - \sum_{i \in N} \sum_{l \in P} f_{lk}^{is} \geq \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ik} - \sum_{q \in Q_k} M_{kq} \bar{h}_{kq}, \forall k \in P, s \in S, q \in Q_k \tag{33}$$

$$g_{jl}^{is} \leq \bar{z}_{jl} \sum_{r \in N} w_{jr}^s, \forall i \in N, j \in Q, l \in P \cup Q \setminus \{j\}, s \in S \tag{34}$$

$$a_{vm}^{is}, g_{mn}^{is}, f_{kl}^{is}, \eta_s \geq 0, \forall i \in N, m \in H, k \in P, n \in H \setminus \{m\}, l \in P \setminus \{k\}, s \in S \tag{35}$$

The dual form of the above SP model, represented by DSP, can be formulated as

$$\begin{aligned} \max & \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \varepsilon_m^s + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{im} - \bar{x}_{rm}) \right] \phi_m^s + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \gamma_k^s \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ie} - \sum_{t \in T_e} M_{et} \bar{b}_{et} \right) \iota_e^s + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{in} - \sum_{z \in Z_m} M_{enz} \bar{d}_{ez} \right) \kappa_n^s \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ik} - \sum_{q \in Q_k} M_{kq} \bar{h}_{kq} \right) \lambda_k^s + \bar{z}_{jl} \sum_{r \in N} w_{jr}^s v_{jl}^s \end{aligned} \quad (36)$$

$$\text{s.t. } \gamma_l^{is} - \gamma_k^{is} - 1 \leq \alpha_1 c_{kl}, \forall i \in N, k \in P, l \in P, s \in S \quad (37)$$

$$\phi_k^{is} - \phi_m^{is} + \gamma_j^{is} - \gamma_m^{is} + \kappa_m^{is} + v_{mk}^{is} \leq \alpha_2 c_{mk}, \forall i \in N, m \in Q, k \in P \cup Q, j \in P, s \in S \quad (38)$$

$$\varepsilon_v^{is} - \varepsilon_m^{is} + \phi_m^{is} - \phi_v^{is} - \iota_v^{is} - \kappa_m^{is} \leq \alpha_3 c_{vm}, \forall i \in N, v \in R, m \in Q, s \in S \quad (39)$$

$$\iota_e^{is} \leq \omega_1, \forall i \in N, e \in R, s \in S \quad (40)$$

$$\kappa_n^{is} \leq \omega_2, \forall i \in N, n \in Q, s \in S \quad (41)$$

$$\lambda_k^{is} \leq \omega_3, \forall i \in N, k \in P, s \in S \quad (42)$$

$$\iota, \kappa, \lambda, v \leq 0 \quad (43)$$

where  $\varepsilon_m^{is}$ ,  $\phi_m^{is}$ ,  $\gamma_k^{is}$ ,  $\iota_e^{is}$ ,  $\kappa_n^{is}$ ,  $\lambda_k^{is}$  and  $v_{jl}^{is}$  are the dual variables of constraints (27)–(35), respectively.

With the introduction of the auxiliary variable  $\zeta$ , the Benders master problem, represented by MP, can then be formulated as

$$\begin{aligned} \min & \sum_{z \in Z_m} \sum_{m \in Q} C_{mz} d_{mz} + \sum_{t \in T_e} \sum_{e \in R} C_{et} b_{et} + \sum_{q \in Q_k} \sum_{k \in P} C_{kq} h_{kq} \\ & + \sum_{s \in S} p_s \sum_{i \in N} \sum_{j \in N} (w_{ij}^s + w_{ji}^s) \sum_{e \in H} c_{ie} x_{ie} + \zeta \end{aligned} \quad (44)$$

$$\text{s.t. Constraints (2) – (14), and (21) – (26)} \quad (45)$$

$$0 \geq \sum_{s \in S} p_s \left\{ \begin{aligned} & \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \varepsilon_m^{is} + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{im} - \bar{x}_{rm}) \right] \phi_m^{is} \\ & + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \gamma_k^{is} + \bar{z}_{jl} \sum_{r \in N} w_{jr}^s v_{jl}^{is} \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ie} - M_{el} \bar{b}_{el} \right) \iota_e^{is} + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{in} - M_{nz} \bar{d}_{ez} \right) \kappa_n^{is} \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ik} - M_{kq} \bar{h}_{kq} \right) \lambda_k^{is} \end{aligned} \right\} \tag{46}$$

$$\zeta \geq \sum_{s \in S} p_s \left\{ \begin{aligned} & \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \varepsilon_m^{is} + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{im} - \bar{x}_{rm}) \right] \phi_m^{is} \\ & + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \gamma_k^{is} + \bar{z}_{jl} \sum_{r \in N} w_{jr}^s v_{jl}^{is} \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ie} - M_{el} \bar{b}_{el} \right) \iota_e^{is} + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{in} - M_{nz} \bar{d}_{ez} \right) \kappa_n^{is} \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ik} - M_{kq} \bar{h}_{kq} \right) \lambda_k^{is} \end{aligned} \right\} \tag{47}$$

Once the first-stage variables are fixed, the sub-problems can be solved independently (Birge & Louveaux, 2011). To obtain more information from each sub-problem and improve the lower bound of MP more effectively, the multi-cut strategy is applied to accelerate Benders decomposition (Rahmaniani et al., 2018). The MP can be stated as

$$\begin{aligned} & \min \sum_{z \in Z_m} \sum_{m \in Q} C_{mz} d_{mz} + \sum_{t \in T_e} \sum_{e \in R} C_{et} b_{et} + \sum_{q \in Q_k} \sum_{k \in P} C_{kq} h_{kq} \\ & + \sum_{s \in S} p_s \sum_{i \in N} \sum_{j \in N} (w_{ij}^s + w_{ji}^s) \sum_{e \in H} c_{ie} x_{ie} + \sum_{s \in S} p_s \zeta_s \end{aligned} \tag{48}$$

$$\text{s.t. Constraints (45) - (46)} \tag{49}$$

$$\begin{aligned} \zeta_s \geq & \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \varepsilon_m^{is} + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{im} - \bar{x}_{rm}) \right] \phi_m^{is} \\ & + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \gamma_k^{is} + \bar{z}_{jl} \sum_{r \in N} w_{jr}^s v_{jl}^{is} + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ie} - M_{el} \bar{b}_{el} \right) \iota_e^{is} \\ & + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{in} - M_{nz} \bar{d}_{ez} \right) \kappa_n^{is} + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ik} - M_{kq} \bar{h}_{kq} \right) \lambda_k^{is} \end{aligned} \tag{50}$$

The step-by-step procedure of the BBC algorithm is as follows.

**Initialization:** Set  $Z_U \leftarrow +\infty, Z_u \leftarrow +\infty, Z_l \leftarrow -\infty$ . Set the cut pool  $P = \emptyset$ . Initialize the root node with the linear relaxation of the Benders master problem (48)–(50).

**repeat**

Initialization:  $Z_u \leftarrow +\infty, Z_l \leftarrow -\infty$

Solve the LRMP at the selected node and get the optimal solution  $(h, d, b, x, u, y, z, \zeta_s)$ .

**While** ( $|Z_l - Z_u| > \varepsilon$ ,  $Z_l < Z_u$  and the problem is feasible) **do**

**For each**  $s \in S$  **do**

Solve  $DSP(36)$ –(43) and get the extreme point  $(i^*, \kappa^*, \lambda^*, \nu^*, \varepsilon^*, \phi^*, \gamma^*)$ ;

**If**

$$\zeta_s \geq \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \varepsilon_m^{is} + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{im} - \bar{x}_{rm}) \right] \phi_m^{is} + \left[ \sum_{r \in N} w_{ir}^s (\bar{x}_{ij} - \bar{x}_{rj}) \right] \gamma_k^{is} + \bar{z}_{jl} \sum_{r \in N} w_{jr}^s \nu_{jr}^{is} \\ + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ij} - M_{ct} \bar{b}_{ct} \right) t_v^{is} + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ij} - M_{nc} \bar{d}_{nc} \right) \kappa_n^{is} + \left( \sum_{i \in N} \sum_{j \in N} w_{ij}^s \bar{x}_{ij} - M_{kq} \bar{h}_{kq} \right) \lambda_k^{is}$$

**then**

Generate optimal cut;

Add the generated cut to the master problem at the selected node and to the cut pool  $P$ ;

Update the upper bound

$$Z_u \leftarrow \min \left\{ Z_u, \sum_{s \in S} p_s (DSP_s^*) + f(h, d, b, x, u, y, z) \right\};$$

$$\text{Update the lower bound } Z_l \leftarrow \max \left\{ Z_l, f(h, d, b, x, u, y, z) + \sum_{s \in S} p_s \zeta_s \right\};$$

**If**  $(h, d, b, x, u, y, z, \zeta_s)$  doesn't violate the generated cuts **then**

$$\text{Update the upper bound } Z_u \leftarrow \min \left\{ Z_u, f(h, d, b, x, u, y, z) + \sum_{s \in S} p_s \zeta_s \right\};$$

**End**

**End**

**if**  $Z_l > Z_u$  or if the LRMP at the selected node is infeasible **then**

Fathom the node;

Go back to repeat;

**End**

**If** the solution  $(h, d, b, x, u, y, z, \zeta_s)$  satisfies the integer constraints **then**

Remove the current sub–problem from the active sets;

$$\text{Update } Z_U \leftarrow \min \left\{ Z_U, f(h, d, b, x, u, y, z) + \sum_{s \in S} p_s \zeta_s \right\};$$

Generate a new sub–problem with integer cuts;

**else**

Create a new node of the current sub–problem by branching and append the cuts to set  $P$  and the newly generated sub–problem;

**End**

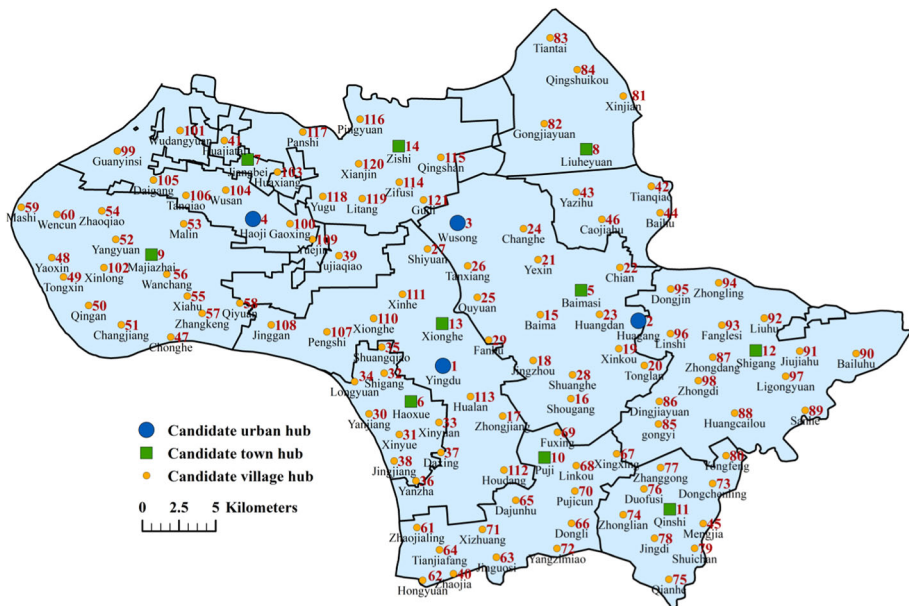
**Until** there is no non–fathomed node from the tree exist;

## 5 Case study

### 5.1 Data input

In this section, we apply the proposed model and solution algorithm to the urban–rural express logistics system of the Jiangling county in Hubei province of Central China. The reasons for choosing such a county for the case study are as follows. First, the Jiangling county has been selected as a delivery logistics pilot county of Hubei Province, and the Jiangling government has recently launched the construction of a three–tiered rural logistics network of county, town and village levels. Second, we have detailed data of this county, such as socio–economic, geographical, demographic and logistics data, thus that our case study can be constructed for realistic scenarios.

The total revenue of the courier industry of the Jiangling region in 2021 hit 29.064 million RMB, with a total of 2.5956 million parcels delivered. This region has an area of 1048.74 square kilometers and a population size of 278,200 as of 2021. It consists of 8 towns (i.e., Baimasi town, Haoxue town, Zishi town, Xionghe town, Shagang town, Puji town, Qishi town and Majiazhai town) and 107 villages, as shown in Fig. 3. According to the data provided by the Jiangling county government, there are currently two urban hubs (No. 1 and 2), 4 town hubs (No. 8, 9, 10, 12), and 65 village hubs (No. 15, 18, 19, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 38, 39, 41, 42, 45, 46, 48, 50, 52, 56, 57, 60, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 82, 85, 86, 87, 88, 90, 91, 92, 93, 94, 96, 97, 98, 99, 103, 106, 108, 110, 111, 112, 114, 120). The aim of the numerical example is to illustrate the optimal solutions of the locations, number, and capacities of the county, town and village hubs for this region, and to judge the gap between the present hub scheme and the optimal scheme. To do so, we consider all of the 107 villages as the candidate village hubs, 10 nodes (No. 5, 6, 7, 8, 9, 10,



**Fig. 3** Geographical distributions of urban, towns and villages in Jiangling region, China

**Table 2** Input data of capacities and construction costs for candidate hubs

	Capacity (allowable maximum number of express parcels)	Construction cost per hub (RMB/hub)
Urban hub	10,000	150,000
	15,000	200,000
	20,000	250,000
Town hub	2000	50,000
	4000	80,000
	6000	110,000
Village hub	150	10,000
	200	20,000
	300	30,000

11, 12) as the candidate town hubs and 4 nodes (No. 1, 2, 3, 4) as the candidate urban hubs. The stochastic logistics demand  $q_{ij}$  between any O–D pair  $ij$  is assumed to follow a truncated normal distribution with  $TN(\bar{q}_{ij}, \sigma_{ij})$ , where  $\bar{q}_{ij}$  and  $\sigma_{ij}$  are the mean and standard variance of the stochastic demand  $q_{ij}$ . The data for the mean  $\bar{q}_{ij}$  and variance  $\sigma_{ij}$  of the O–D demand are not shown here for saving space, together with the geographical locations of towns and villages. However, these data are available from the authors on request. In this example, the average transportation cost,  $\Lambda$ , per express package per km is RMB0.2/km (Lian, 2019). The discount factors  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  for the average transportation cost per package of inter–hub connections are 0.7, 0.8 and 0.9, respectively (Shang et al., 2021; Zhong et al., 2018). The values of the penalty coefficients  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  are large enough (e.g., 1000) such that the capacity constraints can be satisfied when the algorithm iterations terminate.

According to the suggestions of Jiangling county government, only three possible capacity levels are considered for the urban, town and village hubs, as shown in Table 2. The construction cost for each hub capacity level is also shown in Table 2. In order to consider the effects of stochastic demand, Monte Carlo simulations are implemented to generate a set of demand scenarios, leading to a scenario probability distribution  $\{p_s\}$  with  $\sum_{s \in S} p_s = 1$ . The proposed solution algorithm is carried out on a personal computer (Dell, AMD Ryzen 7 5800H 3.20 GHz CPU, RAM 16.0 GB). The solutions obtained are compared with those by the software of GUROBI solver.

## 5.2 Discussion of results

### 5.2.1 Comparison of solution performances of different solvers

We first look at the performance of the solution of the BBC algorithm proposed in this paper and the GUROBI solver under different demand scenario sizes. For all demand scenario sizes, the algorithmic iterations terminate when the relative gap of the upper and lower bounds of the objective function values reaches 0.1%. In order to control the computational time, we set the upper bound of the computational time as 36,000 s. Once this upper bound is reached, the iterations terminate automatically. Table 3 shows the solution performances of the BBC algorithm and the GUROBI solver when the demand scenario size changes from 5 to 80. It can be seen that the GUROBI solver can find the global optimal solution for small–sized

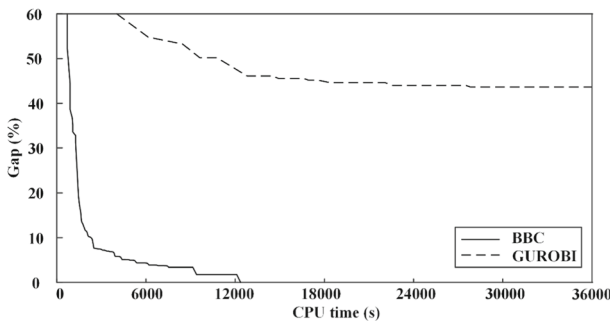


**Table 3** Solution performance of BBC algorithm and GUROBI solver

Scenario size	BBC			GUROBI		
	Expected total cost (RMB)	Gap (%)	CPU time (s)	Expected total cost (RMB)	Gap (%)	CPU time (s)
5	$1.7265 \times 10^7$	0.00	133	$1.7265 \times 10^7$	0	1490
10	$1.6580 \times 10^7$	0.00	726	$1.6580 \times 10^7$	0	8163
20	$1.6025 \times 10^7$	0.00	2394	$1.6025 \times 10^7$	0	28,284
30	$1.5554 \times 10^7$	0.00	5683	$1.5816 \times 10^7$	12.51	36,000
40	$1.5117 \times 10^7$	0.00	7267	$1.5309 \times 10^7$	28.77	36,000
50	$1.4803 \times 10^7$	0.00	8897	$1.5006 \times 10^7$	35.61	36,000
60	$1.4373 \times 10^7$	0.00	9786	$1.4657 \times 10^7$	39.24	36,000
70	$1.4259 \times 10^7$	0.00	12,388	$1.4790 \times 10^7$	48.37	36,000
80	$1.4237 \times 10^7$	0.03	16,963	$1.4585 \times 10^7$	46.85	36,000

scenarios only (i.e., 5, 10, 20). As the scenario number increases to 30, the computational time exceeds 36,000 s, and the iterations stop. As a result, the gap of the upper and lower bounds of the iterative solution is 12.51%. It can also be seen that the BBC algorithm proposed in this paper can obtain the global optimal solution under all the scenarios, and the solution speed outperforms the GUROBI solver, in terms of the computational time. In addition, as the scenario exceeds 70, the expected system cost stabilizes at a level of about  $\text{RMB}1.4259 \times 10^7$ . In the following analysis, a demand scenario size of 70 is adopted.

Figure 4 shows the change of the relative gap of the upper and lower bounds of the objective function values during the course of iterations with CPU time for the BBC algorithm and GUROBI solver when the demand scenario size is 70. It can be observed that the GUROBI solver cannot converge within a specified time threshold of 36,000 s, whereas the BBC algorithm can rapidly converge within about 12,000 s. Therefore, the BBC algorithm proposed in this paper is promising in solving the design models of large-scale urban-rural logistics networks.



**Fig. 4** The convergences of the proposed BBC algorithm and the GUROBI solver

**Table 4** Locations and the number of urban, town and village hubs under different schemes

Scheme	Number and locations of hubs	Expected total cost (RMB)
The present urban–town–village hub location scheme	2 urban hubs: 1, 2 4 town hubs: 8, 9, 10, 12 65 village hubs: 15, 18, 19, 21, 23, 24, 25, 26, 27, 28, 30, 31, 33, 34, 35, 38, 39, 41, 42, 45, 46, 48, 50, 52, 56, 57, 60, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 82, 85, 86, 87, 88, 90, 91, 92, 93, 94, 96, 97, 98, 99, 103, 106, 108, 110, 111, 112, 114, 120	$1.6471 \times 10^7$
The optimal solution generated by the proposed BBC algorithm	3 urban hubs: 1, 2, 4 6 town hubs: 5, 8, 9, 10, 12, 14 82 village hubs: 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 38, 39, 41, 40, 42, 43, 45, 46, 49, 50, 52, 53, 56, 57, 60, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 76, 77, 78, 79, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 96, 97, 98, 100, 102, 103, 105, 108, 109, 110, 111, 112, 114, 115, 117, 118, 119, 120	$1.4259 \times 10^7$

### 5.2.2 Comparison of optimal solution and existing scheme

As previously stated, the Jiangling county government is currently adopting 2 nodes (No. 1, 2) as the urban hubs, 4 nodes (No. 8, 9, 10, 12) as the town hubs and 65 villages as the village hubs. We now identify the difference of the present scheme and the optimal solution generated by the proposed BBC algorithm. Table 4 shows the locations and number of the urban, town and village hubs with the optimal solution and the present scheme. It can be seen that the BBC algorithm leads to the optimal solution of 91 hubs, with an expected total cost of  $\text{RMB}1.4259 \times 10^7$ . The 91 hubs include 3 urban hubs (No. 1, 2, 4), 6 town hubs (No. 5, 8, 9, 10, 12, 14) and 82 village hubs (No. 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 38, 39, 41, 40, 42, 43, 45, 46, 49, 50, 52, 53, 56, 57, 60, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 76, 77, 78, 79, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 96, 97, 98, 100, 102, 103, 105, 108, 109, 110, 111, 112, 114, 115, 117, 118, 119, 120). However, the current implemented scheme generates an expected total cost of  $\text{RMB}1.6471 \times 10^7$ , which is 15.69% higher than that of the optimal scheme.

### 5.2.3 Comparison of solutions of stochastic model and deterministic model

Finally, we look at the difference in the solutions of stochastic model and deterministic model. The deterministic model refers to the situation of  $\sigma_{ij} = 0$  in the truncated normal distribution  $TN(\bar{q}_{ij}, \sigma_{ij})$ . Table 5 shows the solutions of the stochastic model and the deterministic model. It can be seen that the optimal number and locations of the hubs significantly change across the two models. Specifically, compared with the stochastic model, the deterministic model adds nodes 95 and 101 as the village hubs, but removes the village hubs of No. 5, 30, 39, 40,

**Table 5** Comparison of the results of stochastic model and deterministic model

Model	Number and locations of hubs	Expected total cost (RMB)
The stochastic model	3 urban hubs: 1, 2, 4 6 town hubs: 5, 8, 9, 10, 12, 14 82 village hubs: 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 38, 39, 41, 40, 42, 43, 45, 46, 49, 50, 52, 53, 56, 57, 60, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 76, 77, 78, 79, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 96, 97, 98, 100, 102, 103, 105, 108, 109, 110, 111, 112, 114, 115, 117, 118, 119, 120	$1.4259 \times 10^7$
The deterministic model	3 urban hubs: 1, 2, 4 5 town hubs: 8, 9, 10, 12, 14 77 village hubs: 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 33, 34, 35, 38, 41, 42, 43, 45, 46, 49, 50, 52, 53, 56, 57, 60, 61, 62, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 76, 77, 78, 79, 82, 84, 86, 87, 88, 91, 92, 93, 94, 95, 96, 97, 98, 100, 101, 102, 103, 105, 108, 109, 110, 111, 112, 114, 115, 117, 118, 119, 120	$1.3486 \times 10^7$

72, 85, 89 and 90. As a result, the expected total cost of the deterministic model decreases by 5.28%. This means that the deterministic model will cause an underestimation of the total system cost, compared with the stochastic model.

## 6 Conclusion and further studies

In this paper, a hierarchical hub location model is proposed to minimize the expected total system cost for the integrated design of urban and rural logistics networks under demand uncertainty. The locations, number, and capacities of urban–town–village hubs are simultaneously optimized and the effects of logistics demand uncertainty are incorporated. A demand scenario–based branch–and–Benders–cut algorithm is presented to solve the proposed model. A case study of Jiangling urban–rural region in Hubei province of China is implemented, together with a comparison of the results generated by the proposed algorithm and those obtained with GUROBI solver and the present scheme in that region.

The results showed that ignoring the effects of demand uncertainty could lead to a significant decision bias. Therefore, there is indeed a need for the planner to take into account the effects of future demand fluctuations in the urban–rural logistics network design. The proposed methodology in this paper outperforms the GUROBI solver in terms of the size of problem solved and the computational time. It can also significantly improve the efficiency of the urban–rural logistics system in terms of the expected total system cost. The model and solution algorithm proposed in this paper can thus serve as a useful tool for future design of urban–rural logistics networks and for the evaluation of urban–rural transportation and logistics policies.

Although the numerical results presented in this paper are consistent with reality, some real–world features are not fully captured, which may be explicitly considered in further studies. First, only logistics service is considered, and the passenger transportation service

has been ignored. In reality, both human and goods movements are served, sometimes simultaneously through the urban–rural transportation network. It is thus necessary to design an integrated urban–rural transportation network for promoting the human and goods mobility. Second, only demand uncertainty is considered. In reality, uncertainty may also come from the supply side. There is thus a need to jointly consider the uncertainty effects from the demand and supply sides. Third, only travel distance is considered as an indicator of transport cost. It will be more reasonable to consider both travel time and travel distance in the transport cost computation, which is left for a future study.

**Acknowledgements** The work described in this paper was supported by grants from the National Natural Science Foundation of China (72131008, 71890974/71890970) and the Fundamental Research Funds for the Central Universities (2021GCRC014).

## Declarations

**Conflict of interest** The authors declare that there is no financial or non–financial interests that are directly or indirectly related to the work submitted for publication.

## References

- Aalaei, A., & Davoudpour, H. (2017). A robust optimization model for cellular manufacturing system into supply chain management. *International Journal of Production Economics*, 183, 667–679.
- Alumur, S. A., Campbell, J. F., Contreras, I., Kara, B. Y., Marianov, V., & O’Kelly, M. E. (2021). Perspectives on modeling hub location problems. *European Journal of Operational Research*, 291(1), 1–17.
- Alumur, S. A., Nickel, S., Rohrbeck, B., & Saldanha-da-Gama, F. (2018). Modeling congestion and service time in hub location problems. *Applied Mathematical Modelling*, 55, 13–32.
- Alumur, S. A., Yaman, H., & Kara, B. Y. (2012). Hierarchical multimodal hub location problem with time-definite deliveries. *Transportation Research Part E*, 48(6), 1107–1120.
- ARAPED (Annual report for agricultural products e-commerce development), 2022. <https://www.chyxx.com/industry/1102307.html>.
- Birge, J.R., Louveaux, F., 2011. Introduction to Stochastic Programming. Springer Series in Operations Research and Financial Engineering, 1–469.
- Buyukozkan, G., & Ilicak, O. (2022). Smart urban logistics: Literature review and future directions. *Socio-Economic Planning Sciences*, 81, 101197.
- Campbell, J.F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research*, 72(2), 387–405.
- Contreras, I., Cordeau, J., & Laporte, G. (2012). Exact solution of large-scale hub location problems with multiple capacity levels. *Transportation Science*, 46(4), 439–459.
- Contreras, I., & O’Kelly, M. (2019). Hub location problems. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location Science* (pp. 327–363). Springer.
- Correia, I., & Captivo, M. E. (2003). A Lagrangean heuristic for a modular capacitated location problem. *Annals of Operations Research*, 122(1), 141–161.
- Correia, I., Nickel, S., & Saldanha-da-Gama, F. (2010). Single-assignment hub location problems with multiple capacity levels. *Transportation Research Part B*, 44(8–9), 1047–1066.
- Dukkanci, O., & Kara, B. Y. (2017). Routing and scheduling decisions in the hierarchical hub location problem. *Computers & Operations Research*, 85, 45–57.
- Fotuhi, F., & Huynh, N. (2018). A reliable multi-period intermodal freight network expansion problem. *Computers & Industrial Engineering*, 115, 138–150.
- Hu, D. P. (2002). Trade, rural-urban migration, and regional income disparity in developing countries: A spatial general equilibrium model inspired by the case of China. *Regional Science and Urban Economics*, 32(3), 311–338.
- Irawan, C. A., & Jones, D. (2019). Formulation and solution of a two-stage capacitated facility location problem with multilevel capacities. *Annals of Operations Research*, 272(1), 41–67.
- Jabbarzadeh, A., Fahimnia, B., & Seuring, S. (2014). Dynamic supply chain network design for the supply of blood in disasters: A robust model with real world application. *Transportation Research Part E*, 70(1), 225–244.

- Jain, M., Korzhenevych, A., & Hecht, R. (2018). Determinants of commuting patterns in a rural-urban megaregion of India. *Transport Policy*, 68, 98–106.
- Jenabi, M., Fatemi, G., Torabi, S. A., & Sammak, J. (2022). An accelerated Benders decomposition algorithm for stochastic power system expansion planning using sample average approximation. *Opsearch*, 59, 1304–1336.
- Kaveh, F., Tavakkoli-Moghaddam, R., Triki, C., Rahimi, Y., & Jamili, A. (2021). A new bi-objective model of the urban public transportation hub network design under uncertainty. *Annals of Operations Research*, 296(1), 131–162.
- Lagorio, A., Pinto, R., & Golini, R. (2016). Research in urban logistics: A systematic literature review. *International Journal of Physical Distribution & Logistics Management*, 46(10), 908–931.
- Lei, X., Chen, J., Zhu, J., Guo, X., Liu, P., & Jiang, X. (2022). How to locate urban-rural transit hubs from the viewpoint of county integration? *Physica a: Statistical Mechanics and Its Applications*, 606, 1–12.
- Lian, Y. X. (2019). Two-stage method for location selection of rural logistics distribution Center Based on Location Potential Theory. Masters Thesis, Lanzhou Jiaotong University.
- Mišković, S., Stanimirović, Z., & Grujičić, I. (2017). Solving the robust two-stage capacitated facility location problem with uncertain transportation costs. *Optimization Letters*, 11(6), 1169–1184.
- MOT (Ministry of Transportation), 2017. The pilot cities for integrated urban-rural transportation system development. [https://xxgk.mot.gov.cn/2020/jigou/ysfws/202006/t20200623\\_3315384.html](https://xxgk.mot.gov.cn/2020/jigou/ysfws/202006/t20200623_3315384.html).
- Naoum-Sawaya, J., & Elhedhli, S. (2013). An interior-point Benders based branch-and-cut algorithm for mixed integer programs. *Annals of Operations Research*, 210(1), 33–55.
- O'Kelly, M. E. (1986). The location of interacting hub facilities. *Transportation Science*, 20(2), 92–106.
- Qu, X., Wang, S., & Niemeier, D. (2022). On the urban-rural bus transit system with passenger-freight mixed flow. *Communications in Transportation Research*, 2, 100054.
- Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2018). Accelerating the Benders decomposition method: Application to stochastic network design problems. *SIAM Journal on Optimization*, 28(1), 875–903.
- Rei, W., Cordeau, J., Gendreau, M., & Soriano, P. (2009). Accelerating benders decomposition by local branching. *INFORMS Journal on Computing*, 21(2), 333–345.
- Savelsbergh, M., & Van Woensel, T. (2016). City logistics: Challenges and opportunities. *Transportation Science*, 50(2), 579–590.
- Scott, A. J. (1971). Operational analysis of nodal hierarchies in network systems. *Journal of the Operational Research Society*, 22(1), 25–37.
- Shang, X., Yang, K., Jia, B., & Gao, Z. (2021). The stochastic multi-modal hub location problem with direct link strategy and multiple capacity levels for cargo delivery systems. *Transportmetrica A*, 17(4), 380–410.
- Shavarani, S. M., Mosallaeipour, S., Golabi, M., & İzbirak, G. (2019). A congested capacitated multi-level fuzzy facility location problem: An efficient drone delivery system. *Computers & Operations Research*, 108, 57–68.
- Taherkhani, G., Alumur, S. A., & Hosseini, M. (2020). Benders decomposition for the profit maximizing capacitated hub location problem with multiple demand classes. *Transportation Science*, 54(6), 1446–1470.
- Tang, L., Jiang, W., & Saharidis, G. K. (2013). An improved Benders decomposition algorithm for the logistics facility location problem with capacity expansions. *Annals of Operations Research*, 210(1), 165–190.
- Wang, M., Cheng, Q., Huang, J., & Cheng, G. (2021). Research on optimal hub location of agricultural product transportation network based on hierarchical hub-and-spoke network model. *Physica a: Statistical Mechanics and Its Applications*, 566, 125412.
- Wang, S., Chen, Z., & Liu, T. (2020). Distributionally robust hub location. *Transportation Science*, 54(5), 1189–1210.
- Wang, Z., & Sun, S. (2016). Transportation infrastructure and rural development in China. *China Agricultural Economic Review*, 8(3), 516–525.
- Yaman, H. (2011). Allocation strategies in hub networks. *European Journal of Operational Research*, 211(3), 442–451.
- Yang, H., & Cherry, C. (2012). Statewide rural-urban bus travel demand and network evaluation: An application in Tennessee. *Journal of Public Transportation*, 15, 97–111.
- Yao, G. X., Zhu, C. J., & Dai, P. Q. (2019). A research on the construction of hybrid hub-and-spoke logistics network of urban-rural integration. *Industrial Engineering Journal*, 22(6), 1–7.
- Yu, B., Zhu, H., Cai, W., Ma, N., Kuang, Q., & Yao, B. (2013). Two-phase optimization approach to transit hub location: The case of Dalian. *Journal of Transport Geography*, 33, 62–71.
- Zhao, J., Zhang, J. J., & Yan, C. H. (2016). Research on hybrid hub-spoke express network decision with point-point direct shipment. *Chinese Journal of Management Science*, 24, 58–65.

- Zhong, W., Juan, Z., Zong, F., & Su, H. (2018). Hierarchical hub location model and hybrid algorithm for integration of urban and rural public transport. *International Journal of Distributed Sensor Networks*, 14(4), 1–14.
- Zokaei, S., Jabbarzadeh, A., Fahimnia, B., & Sadjadi, S. J. (2017). Robust supply chain network design: An optimization model with real world application. *Annals of Operations Research*, 257(1), 15–44.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.