Distributionally robust inventory routing problem to maximize the service level under limited budget

Ming Liu, Xin Liu, Feng Chu, Feifeng Zheng, Chengbin Chu

Distributionally robust inventory routing problem to maximize the service level under limited budget

Ming Liu[a], Xin Liu[a], Feng Chu[b], Feifeng Zheng[c], Chengbin Chu[d]

[a] *School of Economics & Management, Tongji University, Shanghai, 200092, China*

[b] *IBISC, Univ Évry, University of Paris-Saclay, 91025 Évry, France*

[c] *Glorious Sun School of Business & Management, Donghua University, China*

[d] *Systems Engineering Department, Université Paris-Est, ESIEE Paris, France*

Abstract

This paper studies a stochastic inventory routing problem with alternative handling modules and limited capital budget, under partial distributional information (i.e., the mean and covariance matrix of customer demands). The objective is to maximize the service level, i.e., the probability of jointly ensuring no stockout and respecting the warehouse capacities for all customers at the end of each period. A novel distributionally robust chance constrained formulation is proposed. The sample average approximation method and a model-based hierarchical approach based on problem analysis are developed. Computational results show that the latter approach is more efficient. We also draw some managerial insights.

Keywords: Inventory routing problem; Stochastic optimization; Distributionally robust; Ambiguity set; Service level

Email address: `feng.chu@univ-evry.fr` (Feng Chu[b])

Distributionally robust inventory routing problem to maximize the service level under limited budget

Abstract

This paper studies a stochastic inventory routing problem with alternative handling modules and limited capital budget, under partial distributional information (i.e., the mean and covariance matrix of customer demands). The objective is to maximize the service level, i.e., the probability of jointly ensuring no stockout and respecting the warehouse capacities for all customers at the end of each period. A novel distributionally robust chance constrained formulation is proposed. The sample average approximation method and a model-based hierarchical approach based on problem analysis are developed. Computational results show that the latter approach is more efficient. We also draw some managerial insights.

Keywords: Inventory routing problem; Stochastic optimization; Distributionally robust; Ambiguity set; Service level

1. Introduction

Inventory routing problem (IRP) considers simultaneously the inventory control, the transportation management and their coordination. IRP plays an important role in saving system cost and improving performance of companies (Rahimi et al., 2017). IRP has a wide range of applications in traditional commerce and e-commerce environment. One of the applications comes from the supply chain system of P&G and Wal-Mart (Yu et al., 2009a), where P&G can be considered as the central depot delivering products to Wal-Mart stores (i.e., customers). Another application is from the e-commerce supply chain system, e.g., Amazon and IKEA, where the manufacturer is responsible to delivery products to the product centers. Although there have been various works investigating IRP since it has been introduced by Bell et al. (1983), most of them focus on constant and dynamic customer demands (Savelsbergh and Song, 2008; Raa and Aghezzaf, 2009; Solyalı and Süral, 2011; Michel and Vanderbeck, 2012; Adulyasak et al., 2014).

Practitioners and researchers have the same consensus: distribution strategies have to consider fully the uncertain customer demands, which can be impacted by quickly changing market, flexible price and product promotion (Jiang et al., 2015). Ignoring the uncertainties may have dramatic consequences, such as loss of potential sale, early exhaust of cash flow and loss of new investment opportunities. Most existing works address the stochastic IRP under discrete scenarios or known probability distributions (Kleywegt et al., 2002; Huang and Lin, 2010; Yu et al., 2012; Zhalechian et al., 2016). However, according to Wagner (2008) and Delage and Ye (2010), the complete distributional information may not be well accessed, due to the following factors: (i) it is hard to obtain the information for new products, (ii) customers' reaction cannot be quickly collected and analyzed in time, (iii) and consequence of unforeseeable events and business strategies of competitors cannot be precisely estimated.

Besides, high customer service level is one of goals of companies. Customers' decision making and satisfaction are largely influenced by the service levels of companies. Bijvank and Vis (2012) indicate that customer satisfaction is commonly used as a differentiation strategy among competitors. Most existing related works focus on measuring service level by the rate of delivery delay or stockout under deterministic problem settings or known probability distribution of uncertain parameters (Rahimi et al., 2017), and respecting the service level as a problem constraint (Bijvank and Vis, 2012; Yu et al., 2012; Singh et al., 2015). In this work, each customer has a warehouse, and stockout or overfilling the warehouse may cause great loss for companies, such as huge cost and low customer satisfaction and loyalty (Rahimi et al., 2017). The pursued objective is (customer) service level, which is measured by the probability of jointly ensuring no stockout and respecting the warehouse capacities for all customers at the end of each period.

Moreover, each customer's warehouse possesses multi-functional equipments and has workers with different qualifications. Appropriately combining and managing these resources can significantly reduce warehouse handling cost and improve warehouse performance. Such equipment and worker combinations can be considered as handling modules with different capacities, performance or specific competence (Li et al., 2011; Li et al., 2014). In practice, the handling module selection is a tactical-level decision (Li et al., 2011). Once the handling module is selected, it remains unchanged during the planning horizon (Li et al., 2014). Therefore, in this work, the handling module selection is also regarded as a tactical-level decision.

Motivated by the above observations, to remain close to real IRP business cases, this paper investigates a distributionally robust IRP with various alternative handling modules and limited budget, especially only partial distributional information of customer demands is known (i.e., the mean and covariance matrix). The problem is shown in Figure 1, which is to determine (i) customer warehouse handling module selection in the tactical level, and (ii) vehicle delivery routes and volumes in each period and (iii) customer inventory quantity in each period in the operational level. The objective is to maximize the service level. Especially, the probability or risk of total system cost exceeding the budget is controlled. The contribution of this paper mainly includes:

(1) A new stochastic IRP with partial distributional information on customer demands, limited budget and handling module selection is studied.

(2) For the problem, a novel distributionally robust chance constrained formulation is first proposed, in which the objective function is probabilistic and the risk of system cost exceeding the budget is controlled by a chance constraint.

(3) Based on two approximation methodologies for the partial known probability distribution, a sample average approximation (SAA) method and a hierarchical approach are proposed, respectively.

The remainder of this paper is organized as follows. Section 2 gives a brief literature review. Section 3 describes the problem and provides a distributionally robust chance constrained formulation. In Section 4, two approximation methods for the partial known probability distribution are applied. Based on the two approximation methods, two solution methods, i.e., a classic sample average approximation (SAA) and a hierarchical approach, are designed, respectively. Computational results on randomly generated instances are reported and analyzed in Section 5. Section 6 summarizes this paper and suggests future research directions.

2. Literature review

The deterministic IRPs, to minimize the total cost or maximize the profit, have been well investigated in the literature (Bertazzi et al., 2002; Campbell and Savelsbergh, 2004; Aghezzaf et al., 2006; Yu et al., 2007; Yu et al., 2008; Yu et al., 2009a; Yu et al., 2009b; Bard and Nananukul, 2010; Coelho et al., 2012). Since this work falls within the scope of IRP with uncertain customer demands and handling module selection to maximize the service level, we only review the most related researches. Besides, as we adopt the framework of distribution robust optimization, we also briefly review the distributionally robust approaches.

2.1. IRP with uncertain customer demands

Most existing works address the IRP with stochastic customer demands under given set of scenarios or known probability distributions.

Huang and Lin (2010) consider a multi-item IRP with normally distributed demands, to minimize the total cost. A mixed integer programming (MIP) formulation and an ant colony optimization algorithm are developed. Solyalı et al. (2012) investigate a robust IRP, where uncertain demands are restricted in an interval. Robust optimization is achieved via uncertainty budget control, which measures the protection against the uncertainty. Two MIP formulations and a branch-and-cut algorithm are developed. Bertazzi et al. (2013) study an IRP with known discrete probability distributions of demands, to minimize the total cost. A dynamic programming (DP) formulation and a rollout algorithm are proposed. Shukla et al. (2013) investigate an IRP with known probability distributions of demands, to minimize the total cost. An MIP formulation and algorithm portfolios, based on evolutionary algorithms (EAs), are proposed. Rahim et al. (2014) consider an IRP with normally distributed demands, to minimize the total system cost. An MIP with stochastic parameters and an approximated deterministic model are developed. Bertazzi et al. (2015) study an IRP with known discrete probability distributions of demands, to minimize the total cost. They show that the expected cost based on the deterministic formulation (i.e., using average demand) is worse than that with probability distributions. Soysal et al. (2015) investigate an IRP with normally distributed customer demands, considering CO_2 emission and fuel consumption, to minimize the total system cost. Agra et al. (2018) consider a stochastic IRP with given set of scenarios of demands, to minimize the total system cost. The SAA method and several heuristic algorithms are developed.

In sum, to our best knowledge, researches considering the stochastic IRP, where only partial distributional information of customer demands is known, are very rare.

2.2. IRP considering the service level

Most related existing works focus on considering the service level as a problem constraint (Bijvank and Vis, 2012; Yu et al., 2012; Singh et al., 2015), or measuring service level by the rate of delivery delay or stockout under deterministic problem settings or known probability distribution of uncertain parameters.

Bijvank and Vis (2012) investigate a deterministic IRP, where the service level is controlled. They measure the service level as the average fill rate, i.e., the fraction of demand satisfied. Singh et al. (2015) study a deterministic IRP, to maximize the service level and the efficiency of the operations, where the service level is measured as a function of the number of stockouts. Yu et al. (2012) investigate a stochastic IRP with split delivery to minimize the total system cost, where the uncertain demands respect a normal distribution. In their study, service level is measured as the probability of stockout and the probability of overfilling each customer's warehouse. They consider the service level as a problem constraint. Rahimi et al. (2017) propose a stochastic model for a multi-objective IRP, where the probability distributions of uncertain parameters, such as vehicle speed, customer demands and transportation costs, are known. They measure the service level as a function of delays and backorder frequency.

Therefore, to the best of our knowledge, there is no result on stochastic IRP with partial distributional information on the uncertain demands, to maximize the service level that is measured by the probability of jointly ensuring no stockout and respecting the warehouse capacities for all customers.

2.3. Handling modules

In practice, handling module selection can improve the efficiency and performance of a warehouse. However, most existing works considering handling modules mainly focus on facility location problems. Literature on IRP considering handling modules for inventories is very rare.

Li et al. (2011) investigate a two-stage capacitated facility location problem with handling cost and cross-docking tasks. A given alliance of workers and equipments, having both handling capacity and handling cost, is modeled as one handling module. A limited set of handling module is equipped at each depot location, where the handling cost of a module is incurred as soon as the module is used. Following that, Li et al. (2014) investigate a multi-product facility location problem, where different handling modules are considered. Irawan and Jones (2018) consider a facility location problem, where the locations of distribution centres should be determined and each center have several capacity levels, i.e., handling modules, to be selected.

The constructed warehouse possesses various equipments and workers. Appropriately selecting the equipment and worker combination can improve greatly the system performance. However, to the best of our knowledge, there is no result on IRP jointly considering handling modules.

2.4. Distributionally robust approaches

Under partial distributional information of uncertain demands, it is difficult and impractical to obtain sufficient and representative historical data to exactly estimate the probability distribution (Wagner, 2008; Delage and Ye, 2010; Ben-Tal et al., 2013). For such problems, distributionally robust optimization approaches have shown great power (Zhang et al., 2016; Zhang et al., 2017; Zhang et al., 2018), which include two popular approach classes: (i) one class of approaches is transforming the original problems into deterministic optimization problems based on random samples, among which the SAA method has been widely and successfully applied (Luedtke and Ahmed, 2008; Zhang et al., 2015); (ii) the other class focuses on employing convex approximations (Nemirovski and Shapiro, 2006), which is usually based on ambiguity sets that are assumed to include the true probability distribution.

Zhang et al. (2015) consider a stochastic chance constrained bin packing problem under partial known distributional information of item sizes. Based on the SAA, a two-stage stochastic MIP model is proposed to approximate the problem. Jebali and Diabat (2017) investigate a stochastic chance constrained operating room planning problem. An approximated two-stage stochastic MIP model, based on the SAA, is proposed. Zhang et al. (2016) also investigate a chance constrained bin packing problem, in which only the mean and covariance matrix of item weights are known. Two approximation models, based on two types of ambiguity sets, are proposed and compared with the SAA model. Zhang et al. (2017) investigate the chance constrained allocation of surgery blocks, where the surgery durations are uncertain with only the mean and covariance matrix known. Based on two ambiguity sets, a semidefinite programming (SDP) model and a second-order conic programming (SOCP) model, to approximate the original chance constrained formulation, are established. Zhang et al. (2018) study general distributionally robust chance constrained optimization problems with only mean and covariance matrix known. Two ambiguity sets are constructed for the partial known probability distribution, and two approximated SOCP models are proposed. Esfahani and Kuhn (2018) consider general distributionally robust optimization problems. They construct an ambiguity set named as Wasserstein ambiguity set, and propose an approximated model. The authors compare the approximated model with the SAA method. Liu et al. (2018) investigate a stochastic parallel machine problem, in which the job processing times are uncertain with only mean and covariance matrix known. A distributionally chance constrained formulation with a probabilistic objective is proposed for the problem. The authors propose the SAA method and a heuristic, which is grounded on an approximated model based on an ambiguity set.

Concluding, we broaden the realm of IRPs to maximize the service level under (i) partial demand information, (ii) handling module selection, and (iii) limited capital budget.

3. Problem description and formulation

In this section, we first give the problem statement and then propose a new distributionally robust chance constrained formulation with a probabilistic objective function for the problem.

3.1. Problem description

This paper studies a stochastic IRP with handling modules and limited capital budget. The considered problem includes a central depot 0, a set of customers $\{1, 2, ..., n\}$ and a fleet of vehicles, as shown in Figure 1. Note that on a complete graph, set $N = \{0, 1, 2, ..., n\}$ is used to denote the set of nodes including the central depot and customers. In the problem:

(1) The central depot can be considered as a central vendor or manufacturer, whose capacity is assumed to be unlimited (Yu et al., 2008; Yu et al., 2012). The central depot 0 is responsible for distributing products to customers, via a fleet of homogeneous vehicles with specified capacity. It is assumed that the numbers of vehicles and vehicle tours in each period are not constrained (Yu et al., 2012).

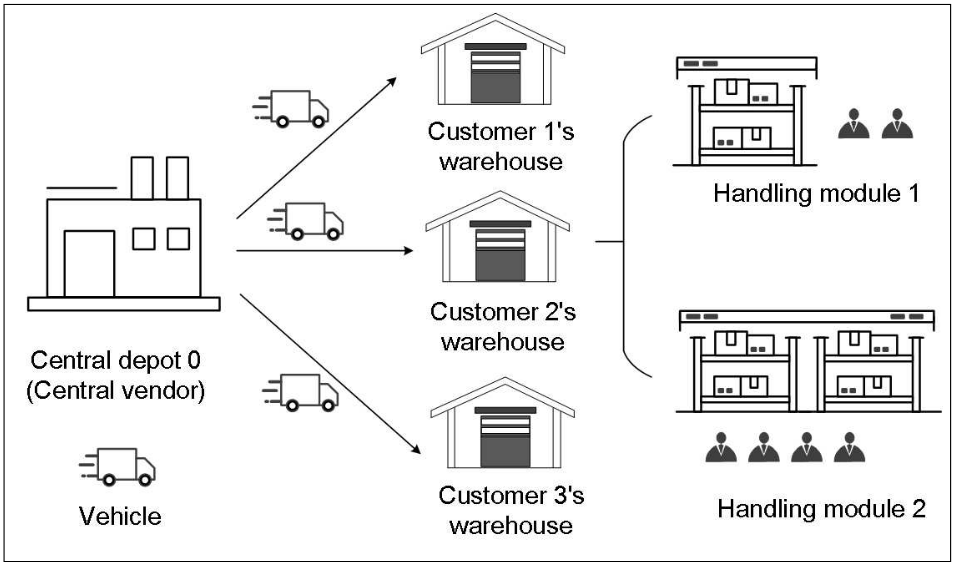

Figure 1: An illustrative example

(2) Given a planning time horizon $T = \{1, 2, ..., t, ..., |T|\}$, the demand of customer $i \in N\backslash\{0\}$ in period $t \in T$ is stochastic and denoted as ξ_{it}, and $\boldsymbol{\xi}_i = [\xi_{i1}, \xi_{i2}, ..., \xi_{i,|T|}]^\mathsf{T}$. Only partial distributional knowledge of $\boldsymbol{\xi}_i$, i.e., the mean and covariance matrix, is known. Besides, a customer's demand can be split and satisfied by two or more vehicles.

(3) Each customer $i \in N\backslash\{0\}$ has a warehouse. In other words, each customer is equipped with a dedicated warehouse, which is located at the corresponding customer's location. Each customer's warehouse possesses various equipments and workers. Different combinations of these resources are considered as handling modules with corresponding warehouse capacities and handling costs.

(4) The system capital budget B is predetermined, and the probability or risk for the total system cost exceeding the capital budget has to be controlled.

As stated above, the accurate probability distribution of customer demands, denoted by $\mathbb{P}$ in the following, is unknown. A limited set of samples, including historical data of customer demands $\{\boldsymbol{\xi}_i^r\}_{r=1}^{|R|}$, is given, where R denotes the set of samples indexed by r and $\boldsymbol{\xi}_i^r = [\xi_{i1}^r, \xi_{i2}^r, ..., \xi_{i,|T|}^r]^\mathsf{T}$ under sample $r \in R$. Then the empirical mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$ of customer demand $\boldsymbol{\xi}_i = [\xi_{it}, \xi_{i2}, ..., \xi_{i,|T|}]^\intercal$ are calculated as:

$$\boldsymbol{\mu}_i = \frac{1}{|R|}\sum_{r\in R}\boldsymbol{\xi}_i^r, \quad \boldsymbol{\Sigma}_i = \frac{1}{|R|}\sum_{r\in R}(\boldsymbol{\xi}_i^r - \boldsymbol{\mu}_i)(\boldsymbol{\xi}_i^r - \boldsymbol{\mu}_i)^\intercal, \quad \forall i \in N\backslash\{0\},$$

where $(\cdot)^\intercal$ denotes the transposition of the vector in parentheses.

The problem is to determine (i) customer warehouse handling module selection in the tactical level; and (ii) vehicle delivery volumes and routes in each period and (iii) customer inventory quantity in each period in the operational level. The objective is to maximize the service level, i.e., the probability of jointly ensuring no stockout and respecting warehouse capacities for all customers at the end of each period.

3.2. The distributionally robust formulation

In this part, a distributionally robust chance constrained formulation for the problem is proposed. In the following, we give basic notations, define decision variables, then propose the formulation.

Parameters:

- T: Set of discrete periods indexed by s and t, and $T = \{1, 2, ..., |T|\}$.
- H: Set of handling modules indexed by h, and $H = \{1, 2, ..., |H|\}$.
- C: Vehicle capacity in volume.
- c_h^a: The fixed cost of handling module $h \in H$.
- c_{ij}: Variable delivering cost per unit of product on arc (i, j), and $c_{ij} = c_{ji}$, which satisfies the triangle inequality $c_{ij} + c_{jl} \geq c_{il},\ \forall i, j, l \in N\backslash\{0\}$.
- c_{i0}^b: Traveling cost of an empty vehicle from customer $i \in N\backslash\{0\}$ driving directly to the central depot 0.
- f_t: Fixed vehicle cost per tour in period $t \in T$.
- c_{it}^{inv}: Inventory cost for holding per unit product by customer $i \in N\backslash\{0\}$ in period $t \in T$.
- I_{i0}: Inventory level of customer $i \in N\backslash\{0\}$ at the beginning.

- W_h: The warehouse capacity of handling module $h \in H$.

- B: The system capital budget.

- α: The required maximum probability (or risk) for exceeding budget B.

- ξ_{it}: Stochastic demand of customer $i \in N \backslash \{0\}$ in period $t \in T$.

Decision variables:

- x_{ijt}: Integer variable, denoting the number of times that arc (i, j) is traversed in period $t \in T$.

- y_{ijt}: Continuous variable, denoting the volume of products transported on arc (i, j) in period $t \in T$.

- z_{ih}: Binary variable, equal to 1 if handling module $h \in H$ is selected for customer $i \in N \backslash \{0\}$, 0 otherwise.

- q_{it}: Continuous variable, denoting the delivery volume to customer $i \in N \backslash \{0\}$ in period $t \in T$.

Distributionally robust formulation [P1]:

$$\max \inf_{\mathbb{P}} \left\{ \mathrm{Prob}_{\mathbb{P}} \left(\begin{array}{r} 0 \leq I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \leq \sum_{h \in H} W_h \cdot z_{ih}, \\ \forall i \in N \backslash \{0\}, t \in T \end{array} \right) \right\} \tag{1}$$

$$\text{s.t.} \sum_{h \in H} z_{ih} = 1, \quad \forall i \in N \backslash \{0\} \tag{2}$$

$$\sum_{j \in N, j \neq i} x_{ijt} = \sum_{j \in N, j \neq i} x_{jit}, \quad \forall i \in N, t \in T \tag{3}$$

$$\sum_{j \in N, j \neq i} y_{jit} - \sum_{j \in N, j \neq i} y_{ijt} = q_{it}, \quad \forall i \in N, t \in T \tag{4}$$

$$\sum_{i \in N \backslash \{0\}} y_{0it} = \sum_{i \in N \backslash \{0\}} q_{it}, \quad \forall t \in T \tag{5}$$

$$y_{ijt} \leq C \cdot x_{ijt}, \quad \forall i \in N, j \in N \backslash \{0\}, j \neq i, t \in T \tag{6}$$

$$\inf_{\mathbb{P}} \left\{ \mathrm{Prob}_{\mathbb{P}} \left(\begin{array}{l} \sum_{i \in N\backslash\{0\}} \sum_{t \in T} \max \left\{ I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is},\ 0 \right\} \cdot c_{it}^{inv} \\ + \sum_{t \in T} \sum_{i \in N} \sum_{j \in N, j \neq i} c_{ij} \cdot y_{ijt} + \sum_{t \in T} \sum_{i \in N\backslash\{0\}} f_t \cdot x_{i0t} \\ + \sum_{i \in N\backslash\{0\}} \sum_{t \in T} c_{i0}^{b} \cdot x_{i0t} + \sum_{i \in N\backslash\{0\}} \sum_{h \in H} c_h^a \cdot z_{ih} \leq B \end{array} \right) \right\} \geq 1 - \alpha \tag{7}$$

$$q_{it} \geq 0, \quad \forall i \in N\backslash\{0\}, t \in T \tag{8}$$

$$y_{ijt} \geq 0, \quad \forall i, j \in N, j \neq i, t \in T \tag{9}$$

$$x_{ijt} \in \mathbb{Z}^{+}, \quad \forall i, j \in N, i \neq j, t \in T \tag{10}$$

$$z_{ih} \in \{0, 1\}, \quad \forall i \in N\backslash\{0\}, h \in H \tag{11}$$

In (1), we use $\inf_{\mathbb{P}} \{\cdot\}$ to denote the worst-case scenario, i.e., the robustness, and $\mathrm{Prob}_{\mathbb{P}}(\cdot)$ to denote the probability of the event in the parentheses under probability distribution $\mathbb{P}$. Note that only partial information on the probability distribution $\mathbb{P}$, i.e., the mean and covariance matrix, is known. The objective function (1) is to maximize the service level, which is measured by the probability of jointly ensuring (i) no stockout for all customers at the end of each period, i.e., $I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \geq 0,\ \forall i \in N\backslash\{0\}, t \in T$, and (ii) respecting the warehouse capacities for all customers at the end of each period, i.e., $I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \leq \sum_{h \in H} W_{ih} \cdot z_{ih},\ \forall i \in N\backslash\{0\}, t \in T$.

Constraint (2) ensures that there must be one handling module selected for customer i's warehouse, where $i \in N\backslash\{0\}$. Constraint (3) implies that the number of vehicles leaving node $i \in N$ is equal to the number of vehicles arriving at i. Constraint (4) denotes the flow conservation, which also serves as the subtour elimination. Constraint (5) implies that the total product volume transported from the central depot is equal to the total delivery volume to all customers in period $t \in T$. Constraint (6) respects the capacities of vehicles. Constraint (7) limits the probability of exceeding the budget less than α, where the total cost includes (i) the inventory holding cost for all customers, i.e., $\sum_{i \in N\backslash\{0\}} \sum_{t \in T} c_{it}^{inv} \cdot \left(\max \left\{ I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is}, 0 \right\} \right)$ and (ii) the variable transportation cost, i.e., $\sum_{t \in T} \sum_{i \in N} \sum_{j \in N, j \neq i} c_{ij} \cdot y_{ijt}$, and (iii) the fixed transportation cost, i.e., $\sum_{t \in T} \sum_{i \in N\backslash\{0\}} f_t \cdot x_{i0t}$, and (iv) the traveling cost for empty vehicles driving from customers to the central depot, i.e., $\sum_{i \in N\backslash\{0\}} \sum_{t \in T} c_{i0}^b \cdot x_{i0t}$, and (v) the total fixed cost for handling module selection, i.e., $\sum_{i \in N\backslash\{0\}} \sum_{h \in H} c_h^a \cdot z_{ih}$. Constraints (8)-(11) give the domains of decision variables.

4. Solution approaches

It is notoriously difficult to obtain exact solutions for general stochastic programs (Birge and Louveaux, 2011). According to Zymler et al. (2013), under partial distributional information, chance constrained problems can only be solved by conservative approximation. In this work, the objective function is probabilistic. Thus our problem is even more difficult, and obtaining the exact solution is quite intractable. To better solve the problem, we first propose an equivalent transformation of the original objective function, leading to an equivalent model [**P2**]. Then, two popular approximation methods for [**P2**] are proposed, i.e., the SAA method and the approximation method based on an ambiguity set for partial known $\mathbb{P}$. Based on the second approximation method, an MIP-based hierarchical approach is further developed.

4.1. Equivalent formulation

In this subsection, by introducing a risk level $\beta \in [0, 1]$, the original probabilistic objective function can be equivalently transformed as:

$$\begin{aligned} \min \quad & \beta \\ \text{s.t. } \inf_{\mathbb{P}} & \left\{ \text{Prob}_{\mathbb{P}} \left(\begin{array}{r} 0 \leq I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \leq \sum_{h \in H} W_h \cdot z_{ih}, \\ \forall i \in N\backslash\{0\}, t \in T \end{array} \right) \right\} \geq 1 - \beta, \end{aligned} \tag{12}$$

where risk level β implies the maximum probability of existing stockout or overfilling a customer's warehouse, and β is considered as a decision variable. For the problem, the maximization of probabilistic objective function (1) is in consistent with the minimization of β, with satisfying joint chance constraint (12) simultaneously.

Accordingly, the original formulation [**P1**] can be equivalently transformed into the following model:

$$
\begin{aligned}
[\mathbf{P2}]: \ \min \quad & \beta \\
\text{s.t.} \quad & (2)\text{-}(12)
\end{aligned}
$$

Under partial distributional information, it is difficult to solve [**P2**] by calling commercial solvers due to chance constraint (7) and joint chance constraint (12). Thus approximation methods are developed in the following.

4.2. The SAA

The basic idea of the SAA is to solve the stochastic optimization problems via Monte Carlo simulation and deterministic optimization techniques (Kleywegt et al., 2002; Hu et al., 2012). For the studied problem, only partial distributionally knowledge of customer demands, i.e., the mean and covariance matrix, is known. Based on the idea of SAA, we replace the partial known distribution $\mathbb{P}$ by an empirical one that satisfies the given conditions, corresponding to a finite set Ω of randomly generated scenarios. Following the expected-penalty-based SAA models (Zhang et al., 2015; Jebali and Diabat, 2017), we approximate the chance constraints (7) and (12) by a sample average estimation function, which is the weighted penalty of the stockout volume, the volume of stock beyond the warehouse capacity and the budget overruns. Note that β can be minimized through minimizing the stockout volume and the volume of stock beyond the warehouse capacity. Accordingly, an approximated SAA-based model [**P3**] is proposed. As the handling module selection is considered as a tactical decision, decision variable z_{il} denoting the handling module selection does not depend on scenario $\omega \in \Omega$ in [**P3**].

New parameters:

- Ω: Set of scenarios indexed by ω, and $\Omega = \{1, 2, ..., |\Omega|\}$.
- $\xi_{it}(\omega)$: Demand of customer $i \in N \backslash \{0\}$ in period $t \in T$ under scenario $\omega \in \Omega$.
- $\theta_1, \theta_2, \theta_3$: The weight coefficients of stockout volume, volume of the stock beyond the warehouse capacity and budget overruns in the objective function, respectively.

New decision variables:

- $I_{it}^{+}(\omega)$: Continuous variable, denoting the inventory level of customer $i \in N\backslash\{0\}$ at the end of period $t \in T$ under scenario $\omega \in \Omega$, and $I_{it}^{+}(\omega) = \max\left\{I_{i,0} + \sum_{s=1}^{t} q_{is}(\omega) - \sum_{s=1}^{t} \xi_{is}(\omega), 0\right\}$.

- $\upsilon_{it}(\omega)$: Continuous variable, denoting the stockout volume of customer $i \in N\backslash\{0\}$ in period $t \in T$ under scenario $\omega \in \Omega$.

- $\eta_{it}(\omega)$: Continuous variable, denoting the volume of stock beyond the warehouse capacity at customer $i \in N\backslash\{0\}$ in period $t \in T$ under scenario $\omega \in \Omega$.

- $\varphi(\omega)$: Continuous variable, denoting the budget overruns under scenario $\omega \in \Omega$.

[P3]:

$$\min\left\{\frac{1}{|\Omega|}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T}\sum_{\omega\in\Omega}\Big(\theta_1\cdot\upsilon_{it}(\omega)+\theta_2\cdot\eta_{it}(\omega)\Big)+\theta_3\cdot\sum_{\omega\in\Omega}\varphi(\omega)\right)\right\} \tag{13}$$

$$\text{s.t.} \quad \sum_{h\in H} z_{ih} = 1, \quad \forall i \in N\backslash\{0\} \tag{14}$$

$$\sum_{j\in N, j\neq i} x_{ijt}(\omega) = \sum_{j\in N, j\neq i} x_{jit}(\omega), \quad \forall i \in N, t \in T, \omega \in \Omega \tag{15}$$

$$\sum_{j\in N, j\neq i} y_{jit}(\omega) - \sum_{j\in N, j\neq i} y_{ijt}(\omega) = q_{it}(\omega), \quad \forall i \in N, t \in T, \omega \in \Omega \tag{16}$$

$$\sum_{i\in N\backslash\{0\}} y_{0it}(\omega) = \sum_{i\in N\backslash\{0\}} q_{it}(\omega), \quad \forall t \in T, \omega \in \Omega \tag{17}$$

$$y_{ijt}(\omega) \leq C \cdot x_{ijt}(\omega), \quad \forall i \in N, j \in N\backslash\{0\}, j \neq i, t \in T, \omega \in \Omega \tag{18}$$

$$I_{i,0} + \sum_{s=1}^{t} q_{is}(\omega) - \sum_{s=1}^{t} \xi_{is}(\omega) + \upsilon_{it}(\omega) \geq 0, \quad \forall i \in N\backslash\{0\}, t \in T, \omega \in \Omega \tag{19}$$

$$I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is}(\omega) - \eta_{it}(\omega) \leq \sum_{h \in H} W_h \cdot z_{ih},$$
$$\forall i \in N \backslash \{0\}, t \in T, \omega \in \Omega \tag{20}$$

$$\sum_{i \in N \backslash \{0\}} \sum_{t \in T} c_{it}^{inv} \cdot I_{it}^{+}(\omega) + \sum_{t \in T} \sum_{i \in N} \sum_{j \in N, j \neq i} c_{ij} \cdot y_{ijt}(\omega) + \sum_{t \in T} \sum_{i \in N \backslash \{0\}} f_t \cdot x_{i0t}(\omega)$$
$$+ \sum_{i \in N \backslash \{0\}} \sum_{t \in T} c_{i0}^{b} x_{i0t}(\omega) + \sum_{i \in N \backslash \{0\}} \sum_{h \in H} c_h^a \cdot z_{ih} - \varphi(\omega) \leq B, \quad \forall \omega \in \Omega \tag{21}$$

$$I_{it}^{+}(\omega) \geq I_{i,0} + \sum_{s=1}^{t} q_{is}(\omega) - \sum_{s=1}^{t} \xi_{is}(\omega), \quad \forall i \in N \backslash \{0\}, t \in T, \omega \in \Omega \tag{22}$$

$$I_{it}^{+}(\omega), \upsilon_{it}(\omega), \eta_{it}(\omega), \varphi(\omega), q_{it}(\omega) \geq 0, \quad \forall i \in N \backslash \{0\}, t \in T, \omega \in \Omega \tag{23}$$

$$y_{ijt}(\omega) \geq 0, \quad \forall i, j \in N, j \neq i, t \in T, \omega \in \Omega \tag{24}$$
$$x_{ijt}(\omega) \in \mathbb{Z}^{+}, \quad \forall i, j \in N, i \neq j, t \in T, \omega \in \Omega \tag{25}$$
$$z_{ih} \in \{0, 1\}, \quad \forall i \in N \backslash \{0\}, h \in H \tag{26}$$

The objective minimizes the weighted sum of the expected stockout volume, the expected volume of stock beyond the warehouse capacities and the expected budget overruns, to approximate chance constraints (7) and (12) in [**P2**].

Constraint (14) is the same as Constraint (2). Constraint (15) denotes that the numbers of vehicles arriving at node $i \in N$ and vehicles leaving $i \in N$ are the same, under scenario $\omega \in \Omega$. Flow conservation under scenario $\omega \in \Omega$ is ensured by Constraint (16). Constraint (17) denotes that the central depot is responsible for delivering products to all customers under scenario $\omega \in \Omega$. Constraint (18) respects the capacities of vehicles under scenario $\omega \in \Omega$. Constraint (19) defines the stockout volume $\upsilon_{it}(\omega)$ of customer $i \in N \backslash \{0\}$ in period $t \in T$ under scenario $\omega \in \Omega$. Constraint (20) calculates the volume of stock $\eta_{it}(\omega)$ beyond the warehouse capacity of customer $i \in N \backslash \{0\}$ in period $t \in T$ under scenario $\omega \in \Omega$. Constraint (21) defines the budget overruns $\varphi(\omega)$ under scenario $\omega \in \Omega$. Constraints (22)-(26) give the domains of decision variables.

MIP model [**P3**] can be optimally solved by calling commercial solvers, such as CPLEX. Due to the NP-hard nature of the problem, we observe that the computational time of the SAA increases dramatically with the problem size. Therefore, to efficiently solve large-scale problems, a two-stage hierarchical method is further designed.

4.3. The MIP-based hierarchical approach

In this part, based on the given mean and covariance of customer demands, we first construct an ambiguity set to approximately characterize the partial known probability distribution $\mathbb{P}$. Joint chance constraint (12) is then conservatively approximated by a set of individual chance constraints via Bonferroni's approximation. Besides, an approximation technique is developed for the chance constraint that restricts the system cost. Following the uncertainty approximation method, i.e., assuming an ambiguity set as in Delage and Ye (2010), and chance constraint approximation approach as in Zhang et al. (2017), an approximated MIP formulation is proposed. Based on that, a hierarchical approach is further developed.

4.3.1. Ambiguity set

Ambiguity set includes a family of distributions characterized via known properties of the partial known $\mathbb{P}$. Traditional ambiguity sets focus on exactly matching the given mean and covariance matrix of uncertain parameters (El Ghaoui et al., 2003; Calafiore and Ghaoui, 2006). Delage and Ye (2010) construct a new ambiguity set considering estimation errors, and the ambiguity set has been successfully applied (Cheng et al., 2013; Zhang et al., 2017). Therefore, following the idea in Delage and Ye (2010), an ambiguity set $\mathcal{P}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \gamma_1, \gamma_2)$, where $\boldsymbol{\mu} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, ..., \boldsymbol{\mu}_n]$ and $\boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, ..., \boldsymbol{\Sigma}_n]$, is applied:

$$\mathcal{P}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \gamma_1, \gamma_2) = \left\{ \mathbb{P} : \begin{array}{l} (\mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}_i] - \boldsymbol{\mu}_i)^{\mathsf{T}} (\boldsymbol{\Sigma}_i)^{-1} (\mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}_i] - \boldsymbol{\mu}_i) \le \gamma_1, \\ \mathbb{E}_{\mathbb{P}}[(\boldsymbol{\xi}_i - \boldsymbol{\mu}_i)(\boldsymbol{\xi}_i - \boldsymbol{\mu}_i)^{\mathsf{T}}] \preceq \gamma_2 \boldsymbol{\Sigma}_i, \quad \forall i \in N \backslash \{0\}. \end{array} \right\},$$

where $\mathbb{E}[\cdot]$ denotes the expected value. Besides, $\gamma_1 \ge 0$ and $\gamma_2 \ge \gamma_1$ are two parameters of ambiguity set $\mathcal{P}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \gamma_1, \gamma_2)$, which restricts that the true mean vector for $\boldsymbol{\xi}_i$ lies in an ellipsoid centered at $\boldsymbol{\mu}_i$ with radius γ_1 and the true covariance matrix for $\boldsymbol{\xi}_i$ is in a positive semi-definite cone which is bounded by $\gamma_2 \boldsymbol{\Sigma}_i$, where $i \in N \backslash \{0\}$. Thus the partial known $\mathbb{P}$ is considered to be included in $\mathcal{P}$, i.e., $\mathbb{P} \in \mathcal{P}$, in the following.

4.3.2. The Bonferroni approximation

Following Bonferroni's inequality, if the individual risk levels ϵ_{it}^1 and ϵ_{it}^2, for violating inequalities $I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \geq 0$ and $I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \leq \sum_{h \in H} W_h \cdot z_{ih}$, are known for each $i \in N \backslash \{0\}$ and $t \in T$, Constraint (12) can be approximated by the following ones:

$$\inf_{\mathbb{P} \in \mathcal{P}} \text{Prob}_{\mathbb{P}} \left(I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \geq 0 \right) \geq 1 - \epsilon_{it}^1, \quad \forall i \in N \backslash \{0\}, t \in T \tag{27}$$

$$\inf_{\mathbb{P} \in \mathcal{P}} \text{Prob}_{\mathbb{P}} \left(I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is} \leq \sum_{h \in H} W_h \cdot z_{ih} \right) \geq 1 - \epsilon_{it}^2, \quad \forall i \in N \backslash \{0\}, t \in T \tag{28}$$

where ϵ_{it}^1 and ϵ_{it}^2 are restricted by $\sum_{i \in N \backslash \{0\}} \sum_{t \in T} (\epsilon_{it}^1 + \epsilon_{it}^2) \leq \beta$ (Calafiore and Ghaoui, 2006). Constraint (27) implies that the probability of no stockout for customer $i \in N \backslash \{0\}$ in period $t \in T$ is no less than $1 - \epsilon_{it}^1$. Constraint (28) requires that the probability of no overstock for customer $i \in N \backslash \{0\}$ in period $t \in T$ is no less than $1 - \epsilon_{it}^2$.

The approximation quality depends largely on the selection of individual risk levels, i.e., ϵ_{it}^1 and ϵ_{it}^2. However, according to Sun et al. (2014), the problem of finding the optimal individual risk levels is very difficult and intractable. In this paper, we apply a popular approach to set ϵ_{it}^1 and ϵ_{it}^2, by evenly dividing β, i.e., $\epsilon_{it}^1 = \epsilon_{it}^2 = \frac{1}{2} \cdot \frac{\beta}{n \cdot |T|}$ (Chung et al., 2012). Therefore, an approximated model [**P4**] with distributionally robust individual chance constraints, i.e., Constraints (27) and (28), is formulated as:

[P4]: $\min \ \beta$

s.t. (2)-(11), (27), (28)

$$\sum_{i \in N \backslash \{0\}} \sum_{t \in T} \left(\epsilon_{it}^1 + \epsilon_{it}^2 \right) \leq \beta \tag{29}$$

$$0 \leq \epsilon_{it}^1, \epsilon_{it}^2, \beta \leq 1, \quad \forall i \in N \backslash \{0\}, t \in T \tag{30}$$

4.3.3. Approximation of the chance constraint for the system cost

Formulation [**P4**] is difficult to be solved, due to the non-linear expression in Constraint (7), i.e., $\max\left\{I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is}, 0\right\}$. Thus we develop a conservative approximation for Constraint (7) (please see Appendix A). To do this, a new continuous and nonnegative variable, $\psi_{it} \geq 0$, is first introduced. Constraint (7) can be approximated by the following three inequalities:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathrm{Prob}_{\mathbb{P}}\left(\psi_{it} \geq I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is},\ \forall i \in N\backslash\{0\}, t \in T\right) \geq 1 - \alpha \tag{31}$$

$$\sum_{i\in N}\sum_{t\in T} c_{it}^{inv} \cdot \psi_{it} + \sum_{t\in T}\sum_{i\in N}\sum_{j\in N, j\neq i} c_{ij} \cdot y_{ijt} + \sum_{t\in T}\sum_{i\in N\backslash\{0\}} f_t \cdot x_{i0t} +$$
$$\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{i0}^{b} x_{i0t} + \sum_{i\in N\backslash\{0\}}\sum_{h\in H} c_h^{a} \cdot z_{ih} \leq B \tag{32}$$

$$\psi_{it} \geq 0,\ \forall i \in N\backslash\{0\}, t \in T \tag{33}$$

According to Bonferroni's inequality, if the individual risk level α_{it} for $\psi_{it} < I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is}$ is known, joint chance constraint (31) can be approximated by the following individual chance constraints:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathrm{Prob}_{\mathbb{P}}\left(\psi_{it} \geq I_{i,0} + \sum_{s=1}^{t} q_{is} - \sum_{s=1}^{t} \xi_{is}\right) \geq 1 - \alpha_{it},\ \forall i \in N\backslash\{0\}, t \in T \tag{34}$$

where $\sum_{i\in N\backslash\{0\}}\sum_{t\in T} \alpha_{it} \leq \alpha$ and α_{it} is obtained in this paper via evenly dividing α, i.e., $\alpha_{it} = \frac{\alpha}{n\cdot|T|}$. Thus, formulation [**P2**] can be approximated by the following chance constrained model [**P5**]:

$$\begin{aligned} \textbf{[P5]:} \quad & \min\ \beta \\ \text{s.t.} \quad & (2)\text{-}(6),\ (8)\text{-}(11),\ (27)\text{-}(30),\ (32)\text{-}(34) \end{aligned}$$

4.3.4. An approximated MIP formulation

Owing to the distributionally robust chance constraints (27), (28) and (34), [**P5**] still cannot be directly solved by calling commercial solvers. By applying the approximation method in Zhang et al. (2017), the distributionally robust chance constraints can be approximated by (35), (36) and (37), respectively (please see Appendix B). Accordingly, an MIP model [**P6**] is proposed to approximate formulation [**P2**]. In the following, related new parameters and new decision variables are first specified, and then [**P6**] is further presented.

New parameters:

- γ_1, γ_2, a, b: Parameters of ambiguity set $\mathcal{P}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \gamma_1, \gamma_2)$, where $\gamma_1 \geq 0$ and $\gamma_2 \geq \gamma_1$ and $\gamma_1 = \frac{b}{1-a-b}$, $\gamma_2 = \frac{1+b}{1-a-b}$.

- α_{it}: $\alpha_{it} \geq \text{Prob}\left(\psi_{it} < I_{i,0} + \sum\limits_{s=1}^{t} q_{is} - \sum\limits_{s=1}^{t} \xi_{is}\right)$, i.e., the maximal probability that variable ψ_{it} is less than the inventory level of customer $i \in N\backslash\{0\}$ in period $t \in T$, and $\alpha_{it} = \frac{\alpha}{n \cdot |T|}$.

- $\boldsymbol{\delta}_{it}$: Vector $[1, ..., 1, 0, ..., 0]^{\mathsf{T}}$, $\forall i \in N\backslash\{0\}, t \in T$, where the first t elements in this vector are 1.

- $\boldsymbol{\pi}_{it}$: Vector $[-1, ..., -1, 0, ..., 0]^{\mathsf{T}}$, $\forall i \in N\backslash\{0\}, t \in T$, where the first t elements are -1.

New decision variables:

- β: Continuous variable, denoting the maximum probability of existing stockout or overfilling a customer's warehouse.

- $\epsilon^1_{it}, \epsilon^2_{it}$: Continuous variable, denoting individual risk probability for existing stockout and overfilling the warehouse of customer $i \in N\backslash\{0\}$ in period $t \in T$, respectively. It is calculated as $\epsilon^1_{it} = \epsilon^2_{it} = \frac{1}{2} \cdot \frac{\beta}{n \cdot |T|}$.

- ψ_{it}: Continuous variable to linearize $\max\left\{I_{i,0} + \sum\limits_{s=1}^{t} q_{is} - \sum\limits_{s=1}^{t} \xi_{is}, 0\right\}$, which is nonnegative.

[P6]: min β

s.t. (2)-(6), (8)-(11), (29), (30), (32), (33)

$$
\sqrt{\frac{1}{1-a-b}}\left(1+\sqrt{\frac{\epsilon_{it}^{1}\cdot b}{1-\epsilon_{it}^{1}}}\right)\sqrt{\boldsymbol{\delta}_{it}{}^{\mathsf{T}}\boldsymbol{\Sigma}_{i}\boldsymbol{\delta}_{it}}
$$

$$
\leq\sqrt{\frac{\epsilon_{it}^{1}}{1-\epsilon_{it}^{1}}}\left(I_{i,0}+\sum_{s=1}^{t}q_{is}-\boldsymbol{\mu}_{i}{}^{\mathsf{T}}\boldsymbol{\delta}_{it}\right),\quad \forall i\in N\backslash\{0\},t\in T \tag{35}
$$

$$
\sqrt{\frac{1}{1-a-b}}\left(1+\sqrt{\frac{\epsilon_{it}^{1}\cdot b}{1-\epsilon_{it}^{1}}}\right)\sqrt{\boldsymbol{\delta}_{it}{}^{\mathsf{T}}\boldsymbol{\Sigma}_{i}\boldsymbol{\delta}_{it}}
$$

$$
\leq\sqrt{\frac{\epsilon_{it}^{1}}{1-\epsilon_{it}^{1}}}\left(I_{i,0}+\sum_{s=1}^{t}q_{is}-\boldsymbol{\mu}_{i}{}^{\mathsf{T}}\boldsymbol{\delta}_{it}\right),\quad \forall i\in N\backslash\{0\},t\in T \tag{36}
$$

$$
\sqrt{\frac{1}{1-a-b}}\left(1+\sqrt{\frac{\alpha_{it}\cdot b}{1-\alpha_{it}}}\right)\sqrt{\boldsymbol{\pi}_{it}{}^{\mathsf{T}}\boldsymbol{\Sigma}_{i}\boldsymbol{\pi}_{it}}
$$

$$
\leq\sqrt{\frac{\alpha_{it}}{1-\alpha_{it}}}\left(\psi_{it}-I_{i,0}-\sum_{s=1}^{t}q_{is}-\boldsymbol{\mu}_{i}{}^{\mathsf{T}}\boldsymbol{\pi}_{it}\right),\quad \forall i\in N\backslash\{0\},t\in T \tag{37}
$$

$$
\psi_{it}\geq 0,\quad \forall i\in N\backslash\{0\},t\in T \tag{38}
$$

Based on [**P6**], a two-stage hierarchical approach named as the MIP-based hierarchical approach is further developed and shown in Algorithm 1. The MIP-based hierarchical approach is based on the trial-and-error rule: finding a minimal β such that [**P6**] is feasible during the iterations and obtaining the corresponding feasible solution. Since [**P6**] is a conservative approximation, budget with larger value is required to satisfy Constraint (32). The input capital budget B' is set to be the product of its original value and a multiplier $\theta_0 \geq 1$, i.e., $B' = B\cdot\theta_0$, to find feasible solutions. If no feasible solution can be found under current θ_0, the value of θ_0 will be updated as $\theta_0 = \theta_0 + \Delta$, where Δ is the step size of θ_0. In the beginning of the method, estimated risk probabilities ϵ_{it}^1 and ϵ_{it}^2 are given. $MAXITER$ implies the maximum number of iterations. U and O denote two sets that store values of ϵ_{it}^1 and ϵ_{it}^2 in each iteration, respectively. U^* and O^* denote two sets that store ϵ_{it}^1 and ϵ_{it}^2 values, under which [**P6**] is feasible.

As shown in Algorithm 1, during the k-th iteration: (i) [**P6**] is solved

Algorithm 1: MIP-based hierarchical approach

Input: Parameters for the problem:
$n,\ |T|,\ |H|,\ C,\ B,\ c_h^a,\ c_{ij},\ c_{i0}^b,\ f_t,\ c_{it}^{inv},\ \boldsymbol{\mu}_i,\ \boldsymbol{\Sigma}_i,\ \forall i,j \in N, t \in T, h \in H.$

$k = 1$; (Iteration counter)

$\epsilon_{it}^{1,k} = \frac{0.618}{n \cdot |T|},\ \epsilon_{it}^{2,k} = \frac{0.618}{n \cdot |T|},\ \forall i \in N, t \in T$;

$U = \emptyset,\ O = \emptyset$; (Set of the values of ϵ_{it}^1 and ϵ_{it}^2)

$U^* = \emptyset,\ O^* = \emptyset$; (Set of the values of ϵ_{it}^1 and ϵ_{it}^2 so that [**P6**] is feasible)

while θ_0 **do**

  for $k = 1 : MAXITER$ **do**

    Solve formulation [**P6**] by calling CPLEX;

    $U = U \cup \epsilon_{it}^1$ and $O = O \cup \epsilon_{it}^2$;

    if *The problem is feasible* **then**

      $U^* = U^* \cup \epsilon_{it}^1$ and $O^* = O^* \cup \epsilon_{it}^2$;

      $\epsilon_{it}^{1,k+1} = \epsilon_{it}^{1,k} - 0.618 \cdot \left(\epsilon_{it}^{1,k} - \max_{u \in \{1,\ldots,|U|\}} \left\{ U(u) | U(u) < \epsilon_{it}^{1,k} \right\} \right)$

      and

      $\epsilon_{it}^{2,k+1} = \epsilon_{it}^{2,k} - 0.618 \cdot \left(\epsilon_{it}^{2,k} - \max_{o \in \{1,\ldots,|O|\}} \left\{ O(o) | O(o) < \epsilon_{it}^{2,k} \right\} \right)$;

    else

      $\epsilon_{it}^{1,k+1} = \epsilon_{it}^{1,k} + 0.618 \cdot \left(\min_{u \in \{1,\ldots,|U|\}} \left\{ U(u) | U(u) > \epsilon_{it}^{1,k} \right\} - \epsilon_{it}^{1,k} \right)$,

      $\epsilon_{it}^{2,k+1} = \epsilon_{it}^{2,k} + 0.618 \cdot \left(\min_{o \in \{1,\ldots,|O|\}} \left\{ O(o) | O(o) > \epsilon_{it}^{2,k} \right\} - \epsilon_{it}^{2,k} \right)$;

    end

  end

  end

  $\theta = \theta + \Delta$;

end

Compare the values in Sets U^* and O^* and obtain the minimum objective β.

Output: A service level $(1 - \beta)$ and the corresponding solution.

and its feasibility is checked by calling CPLEX; (ii) Risk levels $\epsilon_{it}^{1,k}$ and $\epsilon_{it}^{2,k}$ are updated via the step search methods, such as the bisection search and the golden-section search. Since the golden-section search method has been widely used and performs well (Nazareth and Tseng, 2002; Mourad and Reilly, 2010; Chen et al., 2015), it is applied in this study. The process continues until the stop criterion is met. Then by comparing the values of risk levels in U^* and O^*, the minimum values of ϵ_{it}^1 and ϵ_{it}^2 and the corresponding objective value β and solution can be obtained.

5. Computational experiments

In this section, the proposed solution approaches are evaluated in randomly generated instances. The approaches are coded in MATLAB_2014b and combined by CPLEX 12.6 solver. All numerical experiments are conducted on a personal computer with Core I5 and 3.30GHz processor and 8GB RAM under Windows 7 Operation System. The computational times of the methods are limited to 3600 seconds. Each instance is tested 10 times by each approach, to obtain its average value.

5.1. Out-of-sample test

We test the solutions obtained by the two methods in a large set of scenarios, namely the out-of-sample test (Zhang et al., 2016; Zhang et al., 2017; Zhang et al., 2018). In the scenarios, customer demands are generated following a Log-Normal distribution and satisfying the specified information, representing the realization of customer demands (Xie and Ahmed, 2018). To determine the number of scenarios, we have tested 1000, 5000, 10,000 scenarios (Bertsimas et al., 2017; Zhang et al., 2017). Experiments show that three scenario numbers provide very similar results, but 5000 and 10,000 scenarios are extremely time-consuming. Besides, Xie and Ahmed (2018) use 1000 scenarios to evaluate their methods. Thus, we test the solutions obtained by the two methods in 1000 scenarios. As the handling module selection is considered as the tactical-level decision, it should be determined before the realization of customer demands. That is, for each method, based on its obtained handling module selection:

(1) Under each scenario out of the 1000 ones, the vehicle routing decision and the customer inventory quantities in each period are determined by solving [**P3**] via calling CPLEX.

(2) Then the out-of-sample performance of each method is evaluated by the following metrics:

  (i) the service level by the following formula: $\frac{s_1}{1000} \times 100\%$, where s_1 denotes the number of scenarios without stockout and overfilling customers' warehouses;

  (ii) the risk level calculated by $\frac{s_2}{1000} \times 100\%$, where s_2 is the number of scenarios in which the total system cost exceeds the budget;

  (iii) the coverage level obtained by $\frac{times}{1000 \times n \times |T|} \times 100\%$, where $1000 \times n \times |T|$ denotes the total frequency of customer demands and *times* is the number of times for no stockout and overfilling the customers' warehouses.

5.2. An illustrative example

Preliminary analyses are conducted to adjust the parameters for the proposed two solution methods (please see Appendix C), such that each method can obtain solutions with high quality under the selected parameters, to make them comparable. Parameters for the proposed two solution methods are shown in Table 1. In the following, θ_1, θ_2, θ_3 and $|\Omega|$ for the SAA are set to be 6, 6, 1 and 20, and γ_1, γ_2, Δ for the MIP-based hierarchical approach are set to be 0.2 and 0.4.

Table 1: Input parameters for the solution methods

	SAA				MIP-based hierarchical approach			
Parameters	θ_1	θ_2	θ_3	$\|\Omega\|$	γ_1	γ_2	Δ	$MAXITER$
Values	6	6	1	20	0.2	0.4	0.2	10

An illustrative example based on a small network is studied to compare the proposed two solution methods with an exact method by calling CPLEX, which will be detailed later. The parameters of the example are presented in Table 2, where there are 3 customers and 2 time periods in total. It is assumed that the customer demands are independent and its probability distribution is discrete (Table 3). The total number of possible scenarios is $4^6 = 4096$, where 4 denotes the number of probabilities in Table 3 and 6 is the number of demands of 3 customers during 2 periods. According to Table

Table 2: Input parameters for the small network

Parameters	Value
Number of customers (n)	3
Number of time periods ($\|T\|$)	2
Vehicle capacity (C)	6
Fixed vehicle cost (f_t)	[5,10]
Inventory holding cost (c_{it}^{inv})	5
Inventory level at the beginning (I_{i0})	0
Coordinates	$(1,5),(1,4),(5,5)$
Warehouse capacities (W_h)	$[4,6,8]$
Handling cost (c_h^a)	$[100,200,300]$
Capital budget (B)	1000

Table 3: The probability distribution of customer demands

Demands	2	4	6	8
Probability	0.4	0.4	0.1	0.1

Table 4: Experimental results of the small network instance

	Exact method			SAA			MIP-based hierarchical approach		
Customers\ Handling modules	1	2	3	1	2	3	1	2	3
1	1			1			1		
2	1			1			1		
3	1			1			1		
Time (seconds)	312.6			15.2			13.8		
Service level (%)	100			100			100		
Risk level (%)	0			0			0		
Coverage level (%)	100			100			100		

3, the given mean and standard deviation of each customer's demand can be calculated as 3.8 and 3.77.

Table 4 reports the handling module selection obtained by the exact method, the SAA and the MIP-based hierarchical approach. Note that the exact method is based on enumerating all 4096 scenarios, in which CPLEX is called to directly solve [**P3**]. In each row $i \in \{1, 2, 3\}$, if there is a number "1", it means that the handling module of customer i's warehouse is $h \in H$, i.e., $z_{ih} = 1$. The computational times of the exact method, the SAA and the MIP-based hierarchical approach are 312.6, 15.2 and 13.8 seconds, respectively. We can observe from Table 4 that the service level, risk level and coverage level are the same by the three methods, but the computational time of the SAA and the MIP-based hierarchical approach are only 4.86% and 4.41% of the exact method by calling CPLEX, respectively. The results for the tested instance show that the SAA and the MIP-based hierarchical approach perform well for the small-size instance.

However, in practice, as the probability distribution of demands is more complex and the possible scenarios cannot be enumerated, the exact method is not appropriate for large-size instances. Therefore, in the following, we apply the SAA and the MIP-based hierarchical approach to test large-size instances.

5.3. Numerical experiments on randomly generated instances

Numerical experiments on randomly generated instances with different scales are conducted. The tested data is detailed in Table 5. Parameters C, f_t, c_{it}^{inv}, I_{i0} are randomly and uniformly generated as in line with Yu et al. 2012 from the intervals $[100, 400]$, $[400, 700]$, $[0.5, 2]$, $[50, 400]$, and $|T| = 5$. The SAA is based on a finite-sample approximation, via a set of scenarios Ω with randomly generated demand. The way (i.e., selected empirical distribution) to randomly generate demands may impact the performance of the SAA. Therefore, two common distributions are employed, i.e., Uniform distribution and Normal distribution. Since the correlation of customer demands depends on the specific application background, it is necessary to preprocess the limited historical data. Besides, the correlation may be positive or negative. In order to illustrate the proposed approaches, we consider a neutral correlation, i.e., the customer demands in different time periods are assumed as independently distributed (Chen et al., 2010). The computational results are reported in Table 6, where "-" denotes that no feasible

Table 5: Input data of the tested instances

Parameters	Related to previous work	Randomly generated	Value
Mean demand (μ_{it})	✓		$[100, 400]$
Standard deviation of demand (σ_{it})	✓		$0.2\mu_{it}$
Vehicle capacity (C)	✓		$[100, 400]$
Fixed vehicle cost (f_t)	✓		$[400, 700]$
Inventory holding cost (c_{it}^{inv})	✓		$[0.5, 2]$
Inventory level at the beginning (I_{i0})	✓		$[50, 400]$
Vehicle delivering cost (c_{ij})	✓		Geometrical distance from a 10×10 square
Traveling cost of an empty vehicle (c_{i0}^{b})	✓		$10 \times c_{i0}$
Warehouse capacities (W_h)		✓	$[200, 300, 400, 500]$
Handling cost (c_h^a)		✓	$[400, 600, 800, 1000]$
Capital budget (B)		✓	$11.5 \times 10^3 \times n$

solution can be found within 3600 seconds.

From columns 6, 10 and 14 in Table 6, it can be obtained that the computational times of the proposed two solution methods increase with the problem scale. The computational times in columns 6 and 10 of the SAA are much greater than those of the MIP-based hierarchical approach. From the 25-th row in Table 6, it can be observed that the SAA loses its power to solve the instances exceeding 51 customers. Besides, columns 4-10 in Table 6 show that solutions obtained by the SAA under Normal distribution and Uniform distribution demands are very similar, in terms of the service level, risk level, coverage level and computational time. Thus it can be concluded that the impact of different demand distribution for the SAA is quite small. In addition, the solutions obtained by the SAA and the MIP-based hierarchical approach are similar.

Moreover, it can be also obtained from Table 6 that with the expansion of the problem scale, the service level has not become worse. The reasons might be as follows. As in line with Yu et al. (2012), the fleet of vehicles and the number of tours a vehicle performs in each period are not restricted. That is, the increase of customers' demand can be satisfied by employing more vehicles or performing more vehicle tours. Thus, the expansion of the problem scale, i.e., the increase of demand, mainly yields more cost. Besides, the capital budget in the computational experiments is set to be $11.5 \times 10^3 \times n$, where n denotes the number of customers. That is, the

Table 6: Impact of the instance size, and the way to generate customer demands on the SAA

Set	n	SAA with demands generated from Normal Distribution				SAA with demands generated from Uniform Distribution				MIP-based hierarchical approach			
		Service level (%)	Risk level (%)	Coverage level (%)	Time (seconds)	Service level (%)	Risk level (%)	Risk level (%)	Time (seconds)	Service level (%)	Risk level (%)	Coverage level (%)	Time (seconds)
1	5	50.12	0.00	96.83	22.3	50.31	0.00	96.86	34.6	50.33	0.00	96.98	4.5
2	7	100.00	0.00	100.00	41.2	100.00	0.00	100.00	52.3	100.00	0.00	100.00	7.8
3	9	90.01	0.01	99.69	52.5	90.12	0.01	99.64	56.8	90.01	0.00	99.75	13.7
4	11	99.99	0.02	99.99	85.7	99.99	0.01	99.99	87.3	99.99	0.01	99.99	14.6
5	13	99.99	0.00	99.99	119.9	99.99	0.00	99.99	120.5	99.99	0.00	99.99	14.9
6	15	39.17	0.00	98.21	161.6	40.21	0.00	98.33	174.2	40.16	0.00	98.35	21.4
7	17	12.02	0.00	95.02	223.5	12.45	0.00	95.11	230.8	11.95	0.00	95.09	33.6
8	19	99.99	0.00	99.99	269.8	99.99	0.00	99.99	273.4	99.99	0.00	99.99	42.9
9	21	99.99	0.00	99.99	370.4	99.99	0.00	99.99	382.5	99.99	0.00	99.99	49.8
10	23	76.10	0.00	99.53	493.3	76.28	0.00	99.64	504.8	76.25	0.00	99.49	59.8
11	25	99.99	0.00	99.99	649.7	99.99	0.00	99.99	633.1	99.99	0.00	99.99	72.1
12	27	99.99	0.00	99.99	899.5	99.99	0.00	99.99	889.7	99.99	0.00	99.99	85.5
13	29	78.21	0.00	98.75	1112.9	78.33	0.00	98.98	1100.5	78.27	0.00	98.86	104.2
14	31	99.99	0.00	99.99	1424.6	99.99	0.00	99.99	1429.7	99.99	0.00	99.99	122.3
15	33	99.99	0.00	99.99	1826.1	99.99	0.00	99.99	1997.9	99.99	0.00	99.99	150.4
16	35	75.96	0.00	94.44	2530.2	75.94	0.00	94.35	2598.9	76.25	0.00	94.46	193.1
17	37	99.99	0.00	99.99	3106.5	99.99	0.00	99.99	3567.2	99.99	0.00	99.99	213.3
18	39	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	251.1
19	41	53.21	0.00	86.49	3600.0	52.16	0.00	86.97	3600.0	53.67	0.00	87.59	270.5
20	43	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	281.2
21	45	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	300.6
22	47	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	338.0
23	49	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	386.4
24	51	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	3600.0	99.99	0.00	99.99	466.7
25	53	-	-	-	-	-	-	-	-	65.49	0.00	94.98	512.3

capital budget is proportional to the number of customers. Therefore, the capital budget increases with the expansion of the problem scale.

5.4. Sensitivity analyses

Table 7: Input data of the case

Customers (i)	I_{i0}	μ_{it}					c_{it}^{inv}				
		1	2	3	4	5	1	2	3	4	5
1	59	345	171	366	382	155	1.37	1.61	1.32	1.15	0.85
2	52	334	139	171	249	283	0.50	0.76	0.84	1.58	1.60
3	64	297	319	145	238	215	0.55	0.58	0.76	1.45	0.96
4	85	316	184	191	116	345	1.32	1.16	1.59	0.75	1.34
5	57	387	259	194	376	162	1.78	0.84	0.59	1.94	1.06
6	51	116	142	380	236	133	1.19	1.07	1.44	0.79	1.65
7	90	258	216	293	247	131	1.91	0.69	1.05	0.84	1.19
8	77	179	230	318	400	325	1.45	1.11	1.70	1.92	1.51
9	55	378	288	128	323	343	1.60	1.95	0.61	1.64	1.72
10	75	294	331	301	109	280	1.44	0.76	0.87	0.78	0.79
11	70	284	188	214	204	289	1.70	0.72	0.67	0.64	1.25
12	88	365	317	229	259	218	0.75	1.69	1.93	1.00	1.26
13	89	241	164	206	136	332	0.76	1.45	1.99	0.76	1.41
14	91	174	384	104	171	112	1.42	1.41	1.09	1.81	1.32
15	96	288	143	183	307	172	0.51	0.69	1.44	0.86	1.46

In this part, sensitivity analyses on a instance with 15 customers are conducted. The input data is generated in the way detailed in Table 5, such that: (1) four handling modules are set as $W_1 = 200$, $W_2 = 400$, $W_3 = 600$, $W_4 = 800$, and $c_1^a = 400$, $c_2^a = 600$, $c_3^a = 800$, $c_4^a = 1000$; (2) the coordinates of customers and central depot are first generated from a 10×10 square, and c_{ij} is then calculated as the geometrical distance between customers i and j and $c_{i0}^b = 10 \times c_{i0}$; and (3) I_{i0}, μ_{it} and c_{it}^{inv} are shown in Table 7. In the following, unless otherwise specified, (i) the fixed vehicle costs in $|T| = 5$ periods: $f_1 = 615, f_2 = 456, f_3 = 650, f_4 = 424, f_5 = 480$, (ii) the vehicle capacity is set to be 300, (iii) the capital budget is set as $11.5 \times 10^3 \times n$, where $n = 15$, and (iv) the standard deviation is set as $\sigma_{it} = 0.2 \times \mu_{it}$. For the MIP-based hierarchical approach, the limited risk level for exceeding the capital budget is set as $\alpha = 0.5$, and the maximum number of iteration is 10.

The impact of the budget value is first examined, and the budget is set from $9 \times 10^3 \times n, 9.5 \times 10^3 \times n, ..., 12 \times 10^3 \times n$. The numerical results are reported in Table 8 and Figure 2. It can be obtained that the service level, risk level and coverage level obtained by the SAA and the MIP-based hierarchical approach are very similar. However, from columns 5 and 9 in Table 8, we can observe that the computational time of the SAA is about 7 times larger than that of the MIP-based hierarchical approach. As the

Table 8: The impact of capital budget B

B ($\times 10^3 \times n$)	SAA Service level (%)	SAA Risk level (%)	SAA Coverage level (%)	SAA Time (seconds)	MIP-based hierarchical approach Service level (%)	MIP-based hierarchical approach Risk level (%)	MIP-based hierarchical approach Coverage level (%)	MIP-based hierarchical approach Time (seconds)
9.0	0.00	100.00	89.58	160.1	0.00	100.00	89.84	21.4
9.5	0.00	100.00	89.54	159.8	0.00	100.00	89.54	21.2
10.0	0.00	100.00	89.59	153.2	0.00	100.00	89.58	20.9
10.5	0.00	65.00	91.13	155.2	0.00	65.10	91.14	20.8
11.0	1.00	25.00	94.72	157.8	1.02	25.01	94.73	22.4
11.5	40.00	2.00	98.21	152.3	40.21	2.01	98.21	21.7
12.0	92.00	0.00	99.83	156.4	92.31	0.00	99.84	22.7

MIP-based hierarchical approach has a relatively slight advance, in Figure 2, we only plot the performance of this method. The service level and the coverage level increase with the capital budget, as shown in Figure 2. That may be because that handling module with higher warehouse capacity and more vehicles can be selected with larger capital budget. Moreover, as the risk level is measured by the probability of exceeding the capital budget, thus it is understanding that the risk level decreases when the capital budget is getting larger, as illustrated in Figure 2.

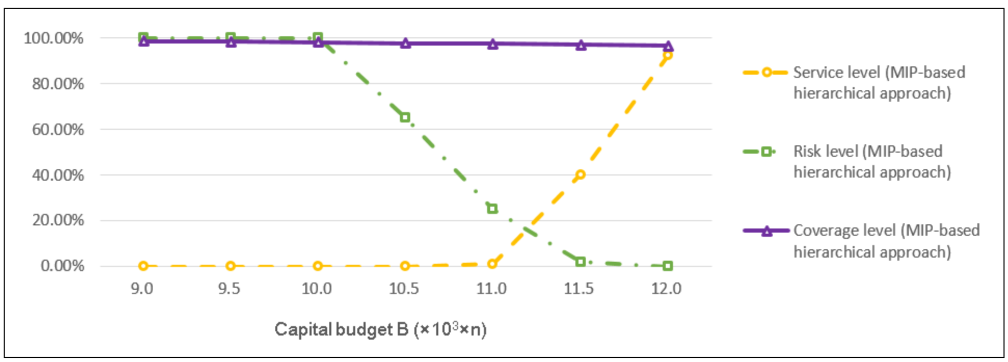

Figure 2: The impact of capital budget

We then examine the sensitivity of the solutions with the standard deviation, i.e., σ_{it}, and the standard deviation is set from 0, $0.1 \times \mu_{it}$, ..., $1 \times \mu_{it}$. Besides, as the standard deviation σ_{it} is set to be related to the mean customer demand μ_{it}, we also test whether the results are influenced by parameter settings, and each mean customer demand μ_{it} is randomly and uniformly generated from intervals [100,200], [100,300], [100, 400] and [100, 500]. Numerical results are reported in Figure 3, from which we can observe that the changing trends of the service level, the coverage level and the risk level are different under different μ_{it}. When μ_{it} is small, i.e., $\mu_{it} \in [100, 200]$, the solution quality is very high under all tested σ_{it}, such that the coverage level and the service level are very close to 1 and the risk level is close to 0. Under $\mu_{it} \in [100, 300]$, larger μ_{it} may exist. With the increase of the σ_{it}, the number of customers with large demands increases, thus the service level decreases. In addition, under $\mu_{it} \in [100, 400]$, we observe that with the increase of σ_{it}, the service level increases at first, then fluctuates slightly, and finally decreases. The reason may be that when σ_{it} starts to increase from 0, the customer demands are relatively fixed in the given interval [100,400] and there are more smaller demands, and thus more demands may be satisfied under given vehicle and warehouse capacities. However, as σ_{it} continues to increase, the number of small customer demands and large customer demands increase simultaneously, thus the service level fluctuates slightly. When σ_{it} is getting larger, there may exist the situation where the demand is too large to be satisfied, thus the service level decreases. Besides, the risk level increases with σ_{it}. The reason may be that when σ_{it} is getting larger, handling module with larger module capacity and higher handling cost will be selected to balance the demand fluctuation, leading to a larger cost. Moreover, when $\mu_{it} \in [100, 500]$, there are more larger μ_{it}, leading to more larger customer demands, and thus the solution quality is quite poor under all tested σ_{it}.

The impact of the vehicle capacity is further examined, and the capacity is set from 100, 150, ..., 400. Computational results of the MIP-based hierarchical approach are shown in Figure 4. It can be obtained that the service level and coverage level increase with the vehicle capacity. The reason may be that when the vehicle capacity is getting larger, more demands can be covered by one vehicle. Besides, the risk level decreases with the increase of the vehicle capacity. That may be because that larger vehicle capacity requires less vehicles to meet customer demands, leading to less system cost.

The impact of the inventory level of each customer at the beginning is tested, and the input or the initial value of the inventory level is set from $I_0 - 50$, $I_0 - 40$, ..., $I_0 + 50$. Numerical results of the MIP-based hierarchical approach are shown in Figure 5. It can be obtained that the service level and the coverage level increase with the input inventory level. The reason may be that when the input inventory level at the beginning increases, same number of vehicle tours can satisfy more demands. Besides, the impact of the input inventory level on the risk level is little.

In sum, it can be obtained that (i) solutions obtained by the SAA and the MIP-based hierarchical approach are similar, in terms of service level,

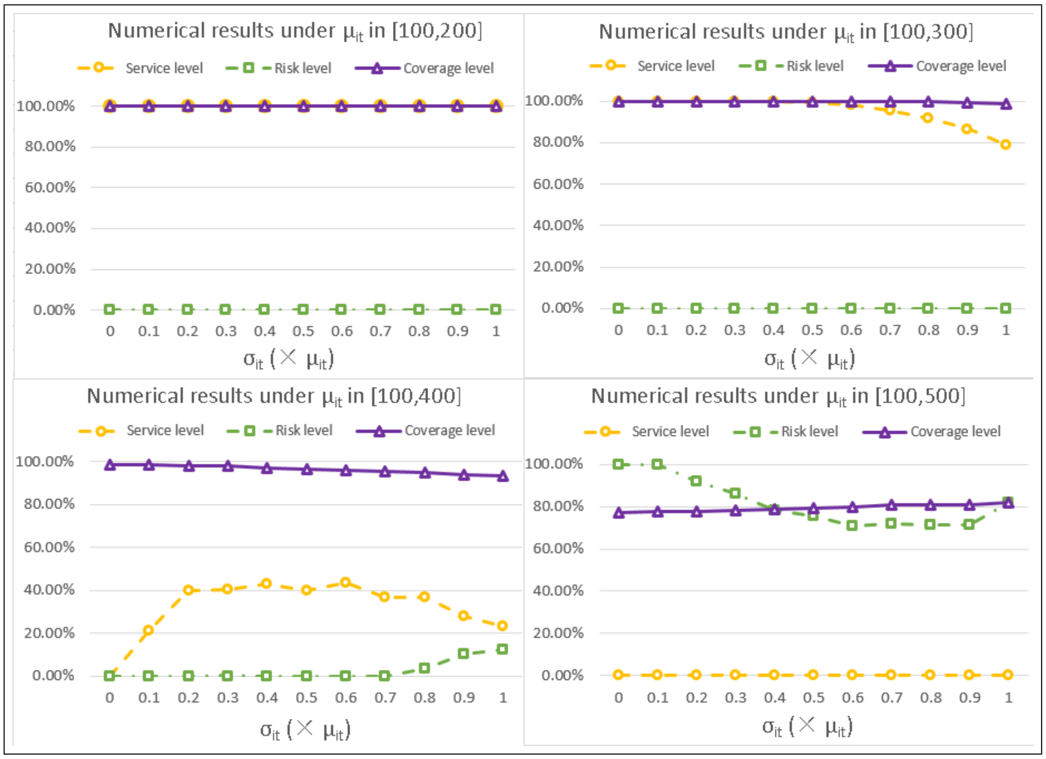

Figure 3: The impact of standard deviation under different mean demands

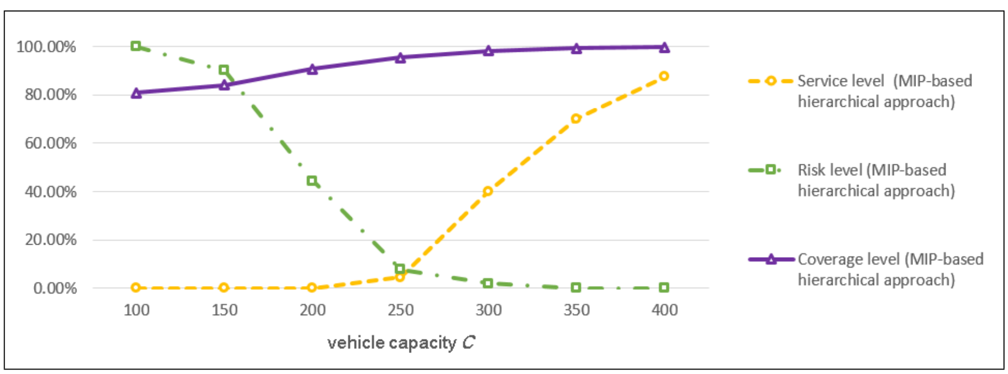

Figure 4: The impact of vehicle capacity

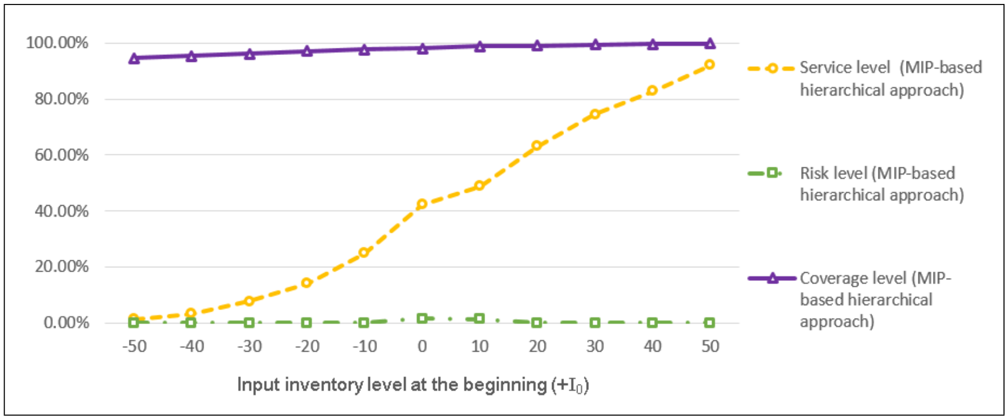

Figure 5: The impact of the inventory level at the beginning

risk level and coverage level; (ii) the computational time of the SAA is much larger than that of the MIP-based hierarchical approach; (iii) with the increase of the capital budget, the service and coverage levels increase and the risk level decreases; (iv) when the standard deviations of customer demands increase, the risk level increases; (v) under the given parameter setting of mean customer demand μ_{it}, with the increase of the standard deviation of each customer demand, the service and coverage levels increase at first and then decrease slightly; (vi) with the increase of the vehicle capacity, the service level and coverage level increase, and the risk level decreases; (vii) with the increase of the input inventory level of each customer at the beginning, the service and coverage levels increase, and the risk level does not change very much; (viii) we recommend the MIP-based hierarchical approach as solution method due to its efficiency.

In view of the above observation, we have made the following proposals for the practitioners and managers:

(1) Trade-offs between the capital budget, the inventory level at the beginning and the service level are remarkably important in supporting the planning process for decision makers. A larger inventory level at the beginning leads to a higher service level, while it also results in a larger holding cost, thus a capital budget increase is needed.

(2) When the mean customer demands are quite small, there is no need to increase the budget, even if the standard deviations are getting larger. However, when there are very large mean customer demands, an

increase of budget is needed, regardless the standard deviations. In other situations, the standard deviation is a significant factor affecting decisions.

(3) The vehicle capacity largely impacts the service level, the coverage level and the risk level, i.e., under large vehicle capacity the same number of vehicle delivery tours can satisfy more customer demands, leading to high coverage level and service level. Thus the vehicle fleet should be hired reasonably.

6. Conclusion

This paper studies a stochastic IRP with various alternative handling modules and limited capital budget, to maximize the service level. Customer demands are assumed to be stochastic, and only partial knowledge on the probability distribution, i.e., the mean and covariance matrix, is known. A novel distributionally robust chance constraint formulation is proposed, in which the objective function is probabilistic and the total system cost is controlled by a chance constraint. As the formulation cannot be directly and optimally solved under partial distributional information, a equivalent transformation model is proposed and approximation methods are applied. Based on two approximation methods, the SAA and the MIP-based hierarchical approach based on the problem characteristics, are developed. Numerical experiments are conducted to evaluate the applicability and the performance of the two methods. We also make some proposals for the practitioners and managers who wish to attain a high service level.

Future research directions may include: (1) to design meta-heuristics which can solve the problem more efficiently, (2) to find the approximation approach to the distributionally robust chance constrained formulation with higher accuracy, and (3) to develop algorithms to calculate an upper bound on the objective function, i.e., the customer service level maximization, to measure the quality of solution approaches (i.e., relative gap).

Acknowledgements

The authors would like to thank the anonymous referees for their constructive comments. This work was supported by the National Natural Science Foundation of China (NSFC) under Grants 71531011, 71771048 and 71571134.

Appendix A: Approximation of chance constrained total cost

Due to the non-linear expression, i.e., $\max\left\{I_{i,0}+\sum\limits_{s=1}^{t}q_{is}-\sum\limits_{s=1}^{t}\xi_{is},0\right\}$ denoted by I_{it}^{+}, Constraint (7) cannot be well addressed. It is still difficult to solve model [**P4**]. Therefore, in this part, we develop a conservative approximation for Constraint (7). For ease of exposition, we use $I_{it}=\left(I_{i,0}+\sum\limits_{s=1}^{t}q_{is}-\sum\limits_{s=1}^{t}\xi_{is}\right)$ in the following. Hereafter, the discussion is for any probability distribution $\mathbb{P}$, and thus we omit $\mathbb{P}$ for ease of exposition. Firstly, a new continuous and nonnegative variable, $\psi_{it}\geq 0,\ \forall i\in N\backslash\{0\},t\in T$, is introduced. That is, $\text{Prob}\left(\psi_{it}\geq 0,\ \forall i\in N\backslash\{0\},t\in T\right)=1$. For a better exposition, suppose $\text{Prob}\left(\psi_{it}\geq I_{it},\ \forall i\in N\backslash\{0\},t\in T\right)=p_1$, where $p_1\in[0,1]$. Thus we have the following relations:

$$
\begin{aligned}
&\text{Prob}\left(\psi_{it}\geq I_{it}^{+},\ \forall i\in N\backslash\{0\},t\in T\right)\\
=&\text{Prob}\left(\psi_{it}\geq I_{it},\psi_{it}\geq 0,\ \forall i\in N\backslash\{0\},t\in T\right)\\
=&\text{Prob}\left(\{\psi_{it}\geq I_{it},\ \forall i\in\mathcal{N},t\in T\}\cap\{\psi_{it}\geq 0,\ \forall i\in N\backslash\{0\},t\in T\}\right)\\
=&\text{Prob}\left(\varphi_{it}\geq I_{it},\ \forall i\in N\backslash\{0\},t\in T\right)\\
&-\text{Prob}\left(\{\varphi_{it}\geq I_{it},\ \forall i\in N\backslash\{0\},t\in T\}\cap\{\varphi_{it}<0,\ \exists i\in N\backslash\{0\},t\in T\}\right)\\
=&\text{Prob}\left(\varphi_{it}\geq I_{it},\ \forall i\in N\backslash\{0\},t\in T\right)\\
=&p_1.
\end{aligned}
$$

Note that as φ_{it} is nonnegative by the definition, i.e., $\varphi_{it}\geq 0,\ \forall i\in N\backslash\{0\},t\in T$ and $\text{Prob}\left(\{\varphi_{it}<0,\ \exists i\in N\backslash\{0\},t\in T\}\right)=0$, thus

$$\text{Prob}\left(\{\varphi_{it}\geq I_{it},\ \forall i\in N\backslash\{0\},t\in T\}\cap\{\varphi_{it}<0,\ \exists i\in N\backslash\{0\},t\in T\}\right)=0.$$

That is, given $\psi_{it}\geq 0,\ \forall i\in N\backslash\{0\},t\in T$, inequality $\psi_{it}\geq I_{it}^{+},\ \forall i\in N\backslash\{0\},t\in T$ is equivalent to inequality $\psi_{it}\geq I_{it},\ \forall i\in N\backslash\{0\},t\in T$. It is understanding that if inequality $\psi_{it}\geq I_{it}^{+},\ \forall i\in N\backslash\{0\},t\in T$ holds, $\sum\limits_{i\in N\backslash\{0\}}\sum\limits_{t\in T}c_{it}^{inv}\cdot\psi_{it}\geq\sum\limits_{i\in N\backslash\{0\}}\sum\limits_{t\in T}c_{it}^{inv}\cdot I_{it}^{+}$ must be satisfied, while the converse may not be true. Suppose $\text{Prob}\left(\sum\limits_{i\in N\backslash\{0\}}\sum\limits_{t\in T}c_{it}^{inv}\cdot\psi_{it}\geq\sum\limits_{i\in N\backslash\{0\}}\sum\limits_{t\in T}c_{it}^{inv}\cdot I_{it}^{+}\right)=$

p_2, where $p_2 \in [0,1]$, we have

$$\text{Prob}\left(\psi_{it} \geq I_{it}^{+},\ \forall i \in N\backslash\{0\}, t \in T\right) \leq \text{Prob}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot\psi_{it} \geq \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot I_{it}^{+}\right),$$

that is, $p_1 \leq p_2$.

In the following, for ease of exposition, we use a new notation $cost' = \left(\sum_{t\in T}\sum_{i\in N}\sum_{j\in N, j\neq i} c_{ij}y_{ijt} + \sum_{t\in T}\sum_{i\in N\backslash\{0\}} f_t x_{i0t} + \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{i0}^{b}x_{i0t} + \sum_{i\in N\backslash\{0\}}\sum_{h\in H} c_h^{a}z_{ih}\right)$. It can be observed that if the following two inequalities, i.e., $\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot\psi_{it} \geq \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot I_{it}^{+}$ and $\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot\psi_{it}+cost' \leq B$ hold simultaneously, inequality $\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot I_{it}^{+}+cost' \leq B$ must be satisfied, but the converse may not be true. Suppose that $\text{Prob}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot I_{it}^{+} + cost' \leq B\right) = p_3$, where $p_3 \in [0,1]$, and $\text{Prob}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot\psi_{it} + cost' \leq B\right) = 1$, we have

$$\begin{aligned}
&\text{Prob}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\cdot I_{it}^{+} + cost' \leq B\right) = p_3\\
\geq&\text{Prob}\left(\left\{\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\psi_{it} \geq \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}I_{it}^{+}\right\}\bigcap\left\{\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\psi_{it} + cost' \leq B\right\}\right)\\
=&\text{Prob}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\psi_{it} \geq \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}I_{it}^{+}\right) -\\
&\text{Prob}\left(\left\{\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\psi_{it} \geq \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}I_{it}^{+}\right\}\bigcap\left\{\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\psi_{it} + cost' > B\right\}\right)\\
=&\text{Prob}\left(\sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}\psi_{it} \geq \sum_{i\in N\backslash\{0\}}\sum_{t\in T} c_{it}^{inv}I_{it}^{+}\right)\\
=&p_2,
\end{aligned}$$

that is, $p_2 \leq p_3$ thus $p_1 \leq p_2 \leq p_3$. Based on the above statement, it can be obtained that if the following inequalities hold, i.e.,

$$\begin{cases} \text{Prob}\left(\psi_{it} \geq I_{it},\ \forall i \in N\backslash\{0\}, t \in T\right) = p_1, \\ \text{Prob}\left(\psi_{it} \geq 0,\ \forall i \in N\backslash\{0\}, t \in T\right) = 1, \\ \text{Prob}\left(\sum\limits_{i \in N\backslash\{0\}} \sum\limits_{t \in T} c_{it}^{inv} \cdot \psi_{it} + cost' \leq B\right) = 1, \\ p_1 \geq 1 - \alpha, \end{cases}$$

inequality $\text{Prob}\left\{\sum\limits_{i \in N\backslash\{0\}} \sum\limits_{t \in T} c_{it}^{inv} \cdot I_{it}^{+} + cost' \leq B\right\} = p_3 \geq 1 - \alpha$ must be satisfied.

Therefore, Constraint (7) can be approximated by the following three inequalities:

$$\begin{cases} \psi_{it} \geq 0,\ \forall i \in N\backslash\{0\}, t \in T \\ \inf_{\mathbb{P} \in \mathcal{P}} \text{Prob}_{\mathbb{P}}\left(\psi_{it} \geq I_{i,0} + \sum\limits_{s=1}^{t} q_{is} - \sum\limits_{s=1}^{t} \xi_{is},\ \forall i \in N\backslash\{0\}, t \in T\right) \geq 1 - \alpha \\ \sum\limits_{i \in N} \sum\limits_{t \in T} c_{it}^{inv} \cdot \psi_{it} + \sum\limits_{t \in T} \sum\limits_{i \in N} \sum\limits_{j \in N, j \neq i} c_{ij} \cdot y_{ijt} + \sum\limits_{t \in T} \sum\limits_{i \in N\backslash\{0\}} f_t \cdot x_{i0t} + \\ \quad \sum\limits_{i \in N\backslash\{0\}} \sum\limits_{t \in T} c_{i0}^{b} x_{i0t} + \sum\limits_{i \in N\backslash\{0\}} \sum\limits_{h \in H} c_h^{a} \cdot z_{ih} \leq B \end{cases}$$

Appendix B: Approximation of distributionally robust chance constraints

By introducing a vector $\boldsymbol{\delta}_{it} = [1, ..., 1, 0, ..., 0]^{\mathsf{T}},\ \forall i \in N\backslash\{0\}, t \in T$, where the first t elements in this vector are equal to 1, Constraint (27) can be rewritten as following:

$$\inf_{\mathbb{P} \in \mathcal{P}} \text{Prob}_{\mathbb{P}}\left(\sum_{s=1}^{t} \xi_{is} \leq I_{i,0} + \sum_{s=1}^{t} q_{is}\right) \geq 1 - \epsilon_{it}^{1}, \quad \forall i \in N\backslash\{0\}, t \in T,$$

or

$$\inf_{\mathbb{P} \in \mathcal{P}} \text{Prob}_{\mathbb{P}}\left(\boldsymbol{\delta}_{it}{}^{\mathsf{T}} \boldsymbol{\xi}_i \leq I_{i,0} + \sum_{s=1}^{t} q_{is}\right) \geq 1 - \epsilon_{it}^{1}, \quad \forall i \in N\backslash\{0\}, t \in T,$$

which can be conservatively approximated by the method introduced by Zhang et al. (2017) as:

$$\sqrt{\frac{1}{1-a-b}}\left(1+\sqrt{\frac{\epsilon_{it}^1\cdot b}{1-\epsilon_{it}^1}}\right)\sqrt{\boldsymbol{\delta}_{it}{}^{\mathsf{T}}\boldsymbol{\Sigma}_i\boldsymbol{\delta}_{it}} \leq \quad \sqrt{\frac{\epsilon_{it}^1}{1-\epsilon_{it}^1}}\left(I_{i,0}+\sum_{s=1}^{t} q_{is}-\boldsymbol{\mu}_i{}^{\mathsf{T}}\boldsymbol{\delta}_{it}\right), \quad \forall i \in N\backslash\{0\}, t \in T$$

where the values of a, b and γ_1, γ_2 are restricted by the following equalities:

$$\gamma_1 = \frac{b}{1-a-b}, \quad \gamma_2 = \frac{1+b}{1-a-b}$$

Similarly, a new coefficient column vector $\boldsymbol{\pi}_{it} = [-1, ..., -1, 0, ..., 0]^{\mathsf{T}}$, $\forall i \in N\backslash\{0\}, t \in T$ is introduced, in which the first t elements are -1. Constraint (28) can be rewritten as:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathrm{Prob}_{\mathbb{P}}\left(-\sum_{s=1}^{t}\xi_{is} \leq -I_{i,0}-\sum_{s=1}^{t} q_{is}+\sum_{h\in H} W_h\cdot z_{ih}\right) \geq 1-\epsilon_{it}^2, \ \forall i \in N\backslash\{0\}, t \in T,$$

or

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathrm{Prob}_{\mathbb{P}}\left(\boldsymbol{\pi}_{it}^{\mathsf{T}}\boldsymbol{\xi}_i \leq -I_{i,0}-\sum_{s=1}^{t} q_{is}+\sum_{h\in H} W_h\cdot z_{ih}\right) \geq 1-\epsilon_{it}^2, \quad \forall i \in N\backslash\{0\}, t \in T.$$

which can be approximated by the following inequality:

$$\sqrt{\frac{1}{1-a-b}}\left(1+\sqrt{\frac{\epsilon_{it}^2\cdot b}{1-\epsilon_{it}^2}}\right)\sqrt{\boldsymbol{\pi}_{it}{}^{\mathsf{T}}\boldsymbol{\Sigma}_i\boldsymbol{\pi}_{it}}$$
$$\leq\sqrt{\frac{\epsilon_{it}^2}{1-\epsilon_{it}^2}}\left(-I_{i,0}-\sum_{s=1}^{t} q_{is}+\sum_{h\in H} W_h\cdot z_{ih}-\boldsymbol{\mu}_i{}^{\mathsf{T}}\boldsymbol{\pi}_{it}\right), \quad \forall i \in N\backslash\{0\}, t \in T$$

Moreover, chance constraint (34) can be rewritten as:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathrm{Prob}_{\mathbb{P}}\left(-\sum_{s=1}^{t}\xi_{is} \leq \psi_{it}-I_{i,0}-\sum_{s=1}^{t} q_{is}\right) \geq 1-\alpha_{it}, \ \forall i \in N\backslash\{0\}, t \in T,$$

or

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathrm{Prob}_{\mathbb{P}} \left(\boldsymbol{\pi}_{it}^{\mathsf{T}} \boldsymbol{\xi}_i \leq \psi_{it} - I_{i,0} - \sum_{s=1}^{t} q_{is} \right) \geq 1 - \alpha_{it},\ \forall i \in N \backslash \{0\}, t \in T,$$

which can be approximated by:

$$\sqrt{\frac{1}{1-a-b}} \left(1 + \sqrt{\frac{\alpha_{it} \cdot b}{1-\alpha_{it}}} \right) \sqrt{\boldsymbol{\pi}_{it}{}^{\mathsf{T}} \boldsymbol{\Sigma}_i \boldsymbol{\pi}_{it}} \leq$$
$$\sqrt{\frac{\alpha_{it}}{1-\alpha_{it}}} \left(\psi_{it} - I_{i,0} - \sum_{s=1}^{t} q_{is} - \boldsymbol{\mu}_i{}^{\mathsf{T}} \boldsymbol{\pi}_{it} \right), \quad \forall i \in N \backslash \{0\}, t \in T$$

Appendix C: Preliminary analysis

Preliminary analyses (on the instance in Section 5.4) are conducted to adjust the input parameters for the SAA and the MIP-based hierarchical approach. The impact of θ_1, θ_2 and θ_3 on the SAA is first tested, and 20 combinations of θ_1, θ_2 and θ_3 are considered: $(1,1,1)$, $(1,1,2)$, $(1,1,3)$, $(1,1,4)$, $(1,1,5)$, $(1,1,6)$, $(1,1,7)$, $(1,1,8)$, $(1,1,9)$, $(1,1,10)$, $(2,2,1)$, $(3,3,1)$, $(4,4,1)$, $(5,5,1)$, $(6,6,1)$, $(7,7,1)$, $(8,8,1)$, $(9,9,1)$, $(10,10,1)$, $(20,20,1)$. The computational results are reported in Figure 6, where the numbers in the horizontal axis denote the above 20 combinations. We can observe that under the 20th combination $(20,20,1)$ the service level and coverage level are the highest, while the risk level is also the highest. Except for $(20,20,1)$, under the 15th combination $(6,6,1)$, the service level is the highest and the service level is the smallest and the coverage level is relatively high. Therefore, we set $\theta_1 = 6,\ \theta_2 = 6,\ \theta_3 = 1$ for the SAA.

The preliminary analysis on number of scenarios for data training, i.e., $|\Omega|$, is then examined on a instance with 17 customers, and the number of scenarios is set from 5, 10, ..., 50. The numerical results are reported in Figure 7, where the number of scenarios is denoted in the horizontal axis. It can be observed that the computational time increases with $|\Omega|$. The risk levels and the coverage levels are very similar under different values of $|\Omega|$. Besides, the service level obtained under $|\Omega| = 20$ is larger than those obtained under $|\Omega| = 5, 10, 15$ and similar to those obtained under other values of $|\Omega|$. Therefore, based on the tradeoff between the computational time and the solution quality, $|\Omega|$ is set to be 20.

To determine the combination of parameters γ_1 and γ_2 for the MIP-based hierarchical approach, 20 parameter combinations are tested. Results are shown in Table 9 and Figure 8. It can be observed that the MIP-based hierarchical approach under $(\gamma_1, \gamma_2) = (0.2, 0.4)$ performs better than that under other combinations. The service level, the risk level and the coverage level are 40.12%, 0.00% and 98.76%, respectively, which are underlined and in bold in Table 9 and marked with a red border in Figure 8. It can be obtained from Table 9 that the average values of the service level, the risk level and the coverage level are 38.28%, 0.01% and 98.53%, respectively. The largest and smallest values of the service level are 40.12% and 37.14%, and the largest and smallest values of the risk level are 0.03% and 0.00%, and the largest and smallest values of the coverage level are 98.84% and 98.17%. In sum, the service levels, the risk levels and the coverage levels are very close to their average values. Therefore, the trend of the curve in Figure 8 is not obvious.

Therefore, in the numerical experiments, the SAA is conducted under parameters $\theta_1 = 6,\ \theta_2 = 6,\ \theta_3 = 1$, and the MIP-based hierarchical approach is conducted under $\gamma_1 = 0.2$ and $\gamma_2 = 0.4$, to obtain solutions with high quality for each method and to make the two methods comparable.

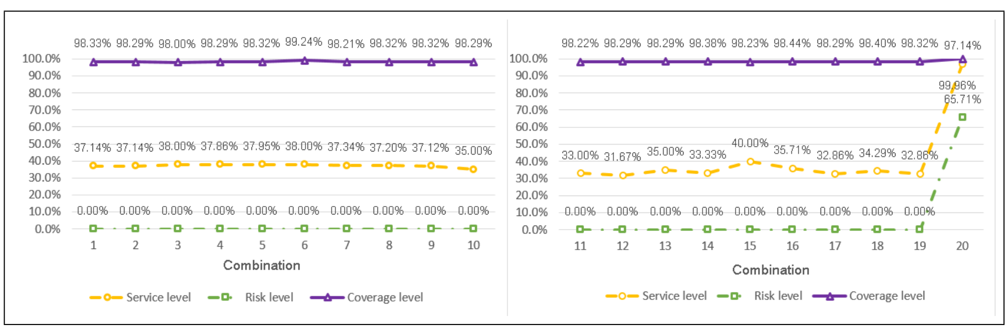

Figure 6: The impact of combinations of θ_1, θ_2 and θ_3 on the SAA

References

[1] Adulyasak, Y., Cordeau, J., and Jans, R. (2014). Formulations and branch-and-cut algorithms for multivehicle production and inventory routing problems. *INFORMS Journal on Computing*, 26(1):103–120.

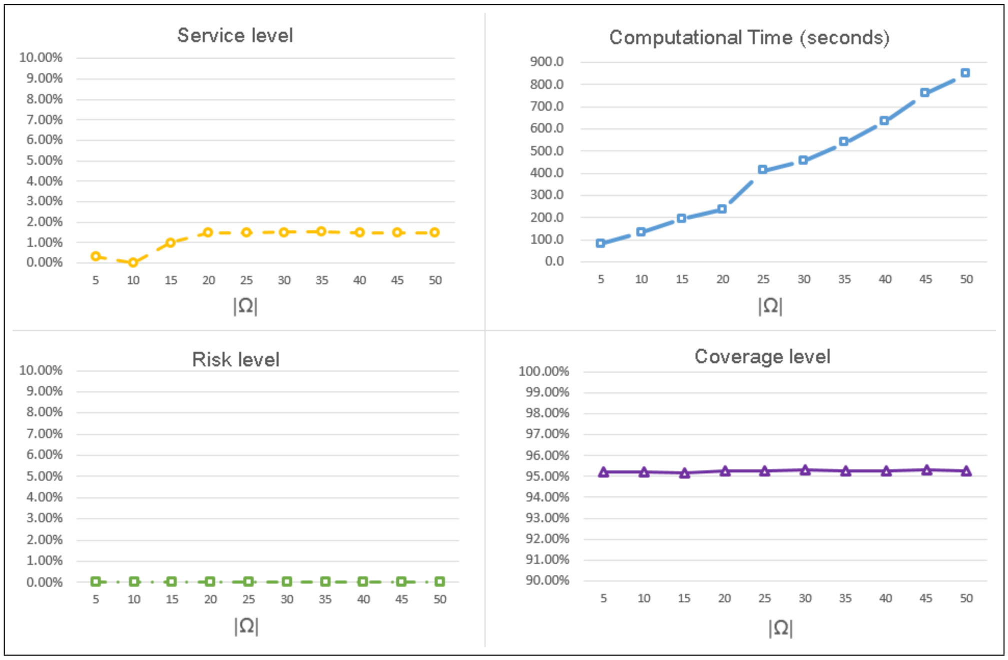

Figure 7: The impact of $|\Omega|$ on the SAA

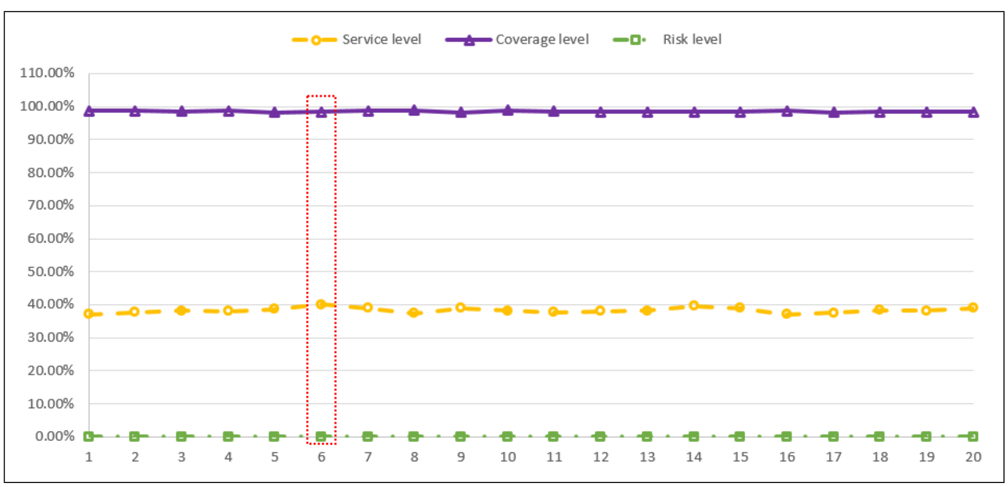

Figure 8: The impact of combinations of γ_1 and γ_2 on the MIP-based hierarchical approach

Table 9: Impact of parameters γ_1 and γ_2 on the MIP-based hierarchical approach

	(0.1,0.2)	(0.1,0.4)	(0.1,0.6)	(0.1,0.8)	(0.1,1)	**(0.2,0.4)**	(0.2,0.6)	(0.2,0.8)	(0.2,1)	(0.3,0.4)
Service level	37.14	37.68	38.12	38.03	38.67	**40.12**	38.98	37.35	38.98	38.24
Risk level	0.00	0.01	0.00	0.00	0.01	**0.00**	0.00	0.01	0.02	0.01
Coverage level	98.67	98.75	98.57	98.78	98.32	**98.46**	98.78	98.84	98.31	98.84
	(0.3,0.6)	(0.3,0.8)	(0.3,1)	(0.4,0.6)	(0.4,0.8)	(0.4,1)	(0.5,0.6)	(0.5,0.8)	(0.6,0.8)	(0.7,0.8)
Service level	37.65	37.95	38.12	39.54	38.98	37.14	37.59	38.28	38.11	38.97
Risk level	0.00	0.00	0.01	0.00	0.00	0.03	0.00	0.00	0.00	0.00
Coverage level	98.59	98.38	98.43	98.46	98.45	98.67	98.17	98.36	98.37	98.36

[2] Aghezzaf, E. H., Raa, B., and Landeghem, H. V. (2006). Modeling inventory routing problems in supply chains of high consumption products. *European Journal of Operational Research*, 169(3):1048–1063.

[3] Agra, A., Requejo, C., and Rodrigues, F. (2018). An adjustable sample average approximation algorithm for the stochastic production-inventory-routing problem. *Networks*, 72(1):5–24.

[4] Bard, J. F. and Nananukul, N. (2010). A branch-and-price algorithm for an integrated production and inventory routing problem. *Computers & Operations Research*, 37(12):2202–2217.

[5] Bell, W. J., Dalberto, L. M., Fisher, M. L., Greenfield, A. J., Jaikumar, R., Kedia, P., Mack, R. G., and Prutzman, P. J. (1983). Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *Interfaces*, 13(6):4–23.

[6] Ben-Tal, A., Hertog, D. D., Waegenaere, A. D., Melenberg, B., and Rennen, G. (2013). Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357.

[7] Bertazzi, L., Bosco, A., Guerriero, F., and Lagan, D. (2013). A stochastic inventory routing problem with stock-out. *Tramsportation Research Part C*, 27(2):89–107.

[8] Bertazzi, L., Bosco, A., and Laganá, D. (2015). Managing stochastic demand in an inventory routing problem with transportation procurement. *Omega*, 56(12):112–121.

[9] Bertazzi, L., Paletta, G., and Speranza, M. G. (2002). Deterministic order-up-to level policies in an inventory routing problem. *Transportation Science*, 36(1):119–132.

[10] Bertsimas, D., Gupta, V., and Kallus, N. (2017). Robust sample average approximation. *Mathematical Programming*, 171:217–282.

[11] Bijvank, M. and Vis, I. F. A. (2012). Lost-sales inventory systems with a service level criterion. *European Journal of Operational Research*, 220(3):610–618.

[12] Birge, J. R. and Louveaux, F. (2011). *Introduction to Stochastic Programming*. Springer Science & Business Media, New York.

[13] Calafiore, G. C. and Ghaoui, L. E. (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory & Applications*, 130(1):1–22.

[14] Campbell, A. M. and Savelsbergh, M. W. P. (2004). A decomposition approach for the inventory-routing problem. *Transportation Science*, 38(4):488–502.

[15] Chen, L. G., Long, D. Z., and Sim, M. (2015). On dynamic decision making to meet consumption targets. *Operations Research*, 63(5):1117–1130.

[16] Chen, W., Sim, M., Sun, J., and Teo, C. P. (2010). From CVaR to uncertainty set: implications in joint chance-constrained optimization. *Operations Research*, 58(2):470–485.

[17] Cheng, J., Lisser, A., and Letournel, M. (2013). Distributionally robust stochastic shortest path problem. *Electronic Notes in Discrete Mathematics*, 41(41):511–518.

[18] Chung, B. D., Yao, T., and Zhang, B. (2012). Dynamic traffic assignment under uncertainty: A distributional robust chance-constrained approach. *Networks & Spatial Economics*, 12(1):167–181.

[19] Coelho, L. C., Cordeau, J. F., and Laporte, G. (2012). The inventory-routing problem with transshipment. *Computers & Operations Research*, 39(11):2537–2548.

[20] Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612.

[21] El Ghaoui, L., Oks, M., and Oustry, F. (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556.

[22] Esfahani, P. M. and Kuhn, D. (2018). Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):115–166.

[23] Hu, J., Homem-De-Mello, T., and Mehrotra, S. (2012). Sample average approximation of stochastic dominance constrained programs. *Mathematical Programming*, 133(1-2):171–201.

[24] Huang, S. H. and Lin, P. C. (2010). A modified ant colony optimization algorithm for multi-item inventory routing problems with demand uncertainty. *Transportation Research Part E*, 46(5):598–611.

[25] Irawan, C. A. and Jones, D. (2018). Formulation and solution of a two-stage capacitated facility location problem with multilevel capacities. *Annals of Operations Research*. https://doi.org/10.1007/s10479-017-2741-7.

[26] Jebali, A. and Diabat, A. (2017). A chance-constrained operating room planning with elective and emergency cases under downstream capacity constraints. *Computers & Industrial Engineering*, 114:329–344.

[27] Jiang, Y., Shang, J., Liu, Y., and May, J. (2015). Redesigning promotion strategy for e-commerce competitiveness through pricing and recommendation. *International Journal of Production Economics*, 167:257–270.

[28] Kleywegt, A. J., Shapiro, A., and Homem-De-Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *Siam Journal on Optimization*, 12(2):479–502.

[29] Li, J., Chu, F., Prins, C., and Zhu, Z. (2014). Lower and upper bounds for a two-stage capacitated facility location problem with handling costs. *European Journal of Operational Research*, 236(3):957–967.

[30] Li, J., Ren, C., Dong, J., and He, M. (2011). A lagrangean-based heuristic for a two-stage facility location problem with handling costs. In *IEEE*

International Conference on Service Operations, Logistics, and Informatics, pages 319–324.

[31] Liu, M., Liu, X., Chu, F., Zheng, F., and Chu, C. (2018). Service-oriented robust parallel machine scheduling. *International Journal of Production Economics*. DOI: 10.1080/00207543.2018.1497311.

[32] Luedtke, J. and Ahmed, S. (2008). A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699.

[33] Michel, S. and Vanderbeck, F. (2012). A column generation based tactical planning method for inventory routing. *Operations Research*, 60(2):382–397.

[34] Mourad, N. and Reilly, J. P. (2010). Minimizing nonconvex functions for sparse vector reconstruction. *IEEE Transactions on Signal Processing*, 58(7):3485–3496.

[35] Nazareth, L. and Tseng, P. (2002). Gilding the lily: A variant of the nelder-mead algorithm based on golden-section search. *Computational Optimization & Application*, 22(1):133–144.

[36] Nemirovski, A. and Shapiro, A. (2006). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.

[37] Raa, B. and Aghezzaf, E. H. (2009). A practical solution approach for the cyclic inventory routing problem. *European Journal of Operational Research*, 192(2):429–441.

[38] Rahim, M. K. I. A., Zhong, Y., Aghezzaf, E. H., and Aouam, T. (2014). Modelling and solving the multiperiod inventory-routing problem with s-tochastic stationary demand rates. *International Journal of Production Research*, 52(14):4351–4363.

[39] Rahimi, M., Baboli, A., and Rekik, Y. (2017). Multi-objective inventory routing problem: A stochastic model to consider profit, service level and green criteria. *Transportation Research Part E*, 101:59–83.

[40] Savelsbergh, M. and Song, J. H. (2008). An optimization algorithm for the inventory routing problem with continuous moves. *Computers & Operations Research*, 35(7):2266–2282.

[41] Shukla, N., Tiwari, M., and Ceglarek, D. (2013). Genetic-algorithms-based algorithm portfolio for inventory routing problem with stochastic demand. *International Journal of Production Research*, 51(1):118–137.

[42] Singh, T., Arbogast, J. E., and Neagu, N. (2015). An incremental approach using local-search heuristic for inventory routing problem in industrial gases. *Computers & Chemical Engineering*, 80:199–210.

[43] Solyalı, O. and Süral, H. (2011). A branch-and-cut algorithm using a strong formulation and an a priori tour-based heuristic for an inventory-routing problem. *Transportation Science*, 45(3):335–345.

[44] Solyalı, O., Süral, H., Cordeau, J. F., and Laporte, G. (2012). Robust inventory routing under demand uncertainty. *Transportation Science*, 46(3):327–340.

[45] Soysal, M., Bloemhof-Ruwaard, J. M., Haijema, R., and van der Vorst, J. G. (2015). Modeling an inventory routing problem for perishable products with environmental considerations and demand uncertainty. *International Journal of Production Economics*, 164:118–133.

[46] Sun, H., Gao, Z., Szeto, W. Y., Long, J., and Zhao, F. (2014). A distributionally robust joint chance constrained optimization model for the dynamic network design problem under demand uncertainty. *Networks & Spatial Economics*, 14(3-4):409–433.

[47] Wagner, M. R. (2008). Stochastic 0-1 linear programming under limited distributional information. *Operations Research Letters*, 36(2):150–156.

[48] Xie, W. and Ahmed, S. (2018). Distributionally robust simple integer recourse. *Computational Management Science*, 15:351–367.

[49] Yu, Y., Chen, H., and Chu, F. (2008). A new model and hybrid approach for large scale inventory routing problems. *European Journal of Operational Research*, 189(3):1022–1040.

[50] Yu, Y., Chu, C., Chen, H., and Chu, F. (2012). Large scale stochastic inventory routing problems with split delivery and service level constraints. *Annals of Operations Research*, 197(1):135–158.

[51] Yu, Y., Chu, F., and Chen, H. (2007). A note on coordination of production and distribution planning. *European Journal of Operational Research*, 177(1):626–629.

[52] Yu, Y., Chu, F., and Chen, H. (2009a). A stackelberg game and its improvement in a vmi system with a manufacturing vendor. *European Journal of Operational Research*, 192(3):929–948.

[53] Yu, Y., Huang, G. Q., and Liang, L. (2009b). Stackelberg game-theoretic model for optimizing advertising, pricing and inventory policies in vendor managed inventory (vmi) production supply chains. *Computers & Industrial Engineering*, 57(1):368–382.

[54] Zhalechian, M., Tavakkoli-Moghaddam, R., Zahiri, B., and Mohammadi, M. (2016). Sustainable design of a closed-loop location-routing-inventory supply chain network under mixed uncertainty. *Transportation Research Part E*, 89:182–214.

[55] Zhang, Y., Jiang, R., and Shen, S. (2016). Distributionally robust chance-constrained bin packing. Available at http://www.optimizationonline.org/DB FILE/2016/09/5662.pdf.

[56] Zhang, Y., Jiang, R., and Shen, S. (2018). Ambiguous chance-constrained binary programs under mean-covariance information. *SIAM Journal on Optimization*, 28(4):2922–2944.

[57] Zhang, Y., Shen, S., and Erdogan, S. A. (2017). Solving 0-1 semidefinite programs for distributionally robust allocation of surgery blocks. *Optimization Letters*, 12(7):1503–1521.

[58] Zhang, Z., Denton, B., and Xie, X. (2015). Branch and price for chance constrained bin packing. *Available at Optimization-Online http://www.optimization-online.org/DB_HTML/2015/11/5217.html*.

[59] Zymler, S., Kuhn, D., and Rustem, B. (2013). Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198.