International Journal of Production Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tprs20

Multi-period distribution network design with boundedly rational customers for the service-oriented manufacturing supply chain: a 4PL perspective

Yuxin Zhang, Zheming Gao, Min Huang, Songchen Jiang, Mingqiang Yin & Shu-Cherng Fang

INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH
https://doi.org/10.1080/00207543.2022.2140220

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Multi-period distribution network design with boundedly rational customers for the service-oriented manufacturing supply chain: a 4PL perspective

Yuxin Zhang[a], Zheming Gao[a], Min Huang[a], Songchen Jiang[a], Mingqiang Yin[a,b] and Shu-Cherng Fang[c]

[a]College of Information Science and Engineering, Northeastern University, Shenyang, People's Republic of China; [b]College of Information and Control Engineering, Liaoning Petrochemical University, Fushun, People's Republic of China; [c]Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC, USA

ABSTRACT
With the service level playing an increasingly essential role in the service-oriented manufacturing (SOM) supply chain, the distribution network design can be heavily affected by the customer behaviour based on their satisfaction of services. In this paper, we consider the service level for service time and delivery quantity separately. A novel mixed integer non-linear programming model is proposed to design the multi-period distribution network from a fourth-party logistics (4PL) perspective. The customer satisfaction based on prospect theory is maximised while considering the investment budget and the service level. A scenario-based linear reformulation is proposed to find the optimal solution when the problem scale is small. For a large-scale problem, we propose an individual-driven Q-learning based memetic particle swarm optimisation algorithm. Numerical experiments are conducted to demonstrate the effectiveness and efficiency of the proposed algorithm. Furthermore, the impact of service modes, different customer behaviour, and customer satisfaction evaluation periods on distribution network is investigated. We find that the length of evaluation periods leads to differences in customer satisfaction due to different perceptions of 'small loss' and 'big gain' by boundedly rational customers.

ARTICLE HISTORY
Received 31 May 2022
Accepted 18 October 2022

KEYWORDS
SOM supply chain network design; 4PL; bounded rationality behaviour; individual-driven Q-learning; memetic particle swarm optimisation algorithm

1. Introduction

Service-oriented manufacturing (SOM) has gradually become an emerging business paradigm that integrates service and manufacturing (Huang et al. 2020; Jiang et al. 2022; Rymaszewska, Helo, and Gunasekaran 2017). Nowadays, besides selling the products, manufacturers are making more profit by providing services (He et al. 2021). For example, as of 1 April 2020, the Toyota listed on its website that Toyota offers parts logistics services in different countries around the world, which has helped Toyota capture the absolute aftermarket. As of 10 May 2021, the Toyota listed on its website that Toyota outsources its logistics services to the third-party logistics (3PL) to focus on their core product business. A 3PL provider offers transportation services from plants to customers, however, is not able to offer integrated solutions for the SOM supply chain. Instead of relying on the 3PL providers, some large manufacturers (e.g. Haier [Lee 2016] and Hisense [Cheng 2021]) have recently cooperated with Cainiao Logistics (the largest fourth-party logistics [4PL] provider in China) in supply chain solutions.

The concept of 4PL was originally brought up by Accenture as an integrator that provides comprehensive supply chain solutions by integrating the resources, technologies and capabilities of its own and other organisations (Gattorna and Jones 1998). A 4PL provider can strategically integrate multiple 3PL providers and offer comprehensive supply chain solutions with low cost, high customer satisfaction and strong reliability through reasonable decisions on logistics facilities, product distribution and transportation providers (Huang, Tu, et al. 2019). Compared with a single 3PL provider with limited capabilities, a 4PL provider is able to offer the logistics services of higher quality, so is more preferred by the SOM supply chain. 4PL has three operating modes: synergy plus, solution integrator and industry integrator. Among them, as the core mode, solution integrator mode can provide the complete supply chain solution for a single company in a certain industry (Büyüközkan, Feyzioğlu, and Ersoy 2009). As one of the most important factors in the SOM supply chain management, network design has a great impact on strategic, tactical, and operational decisions. In this problem, the products

CONTACT Min Huang mhuang@mail.neu.edu.cn College of Information Science and Engineering, Northeastern University, Shenyang 110819, Liaoning, People's Republic of China

are transported by plants to customers via distribution centres (DCs).

In today's fast-paced competitive environment, many parameters such as potential market demands, costs, and capabilities of DCs and transportation routes may change. There would be a potentially catastrophic weakness if the network could not be adjusted in time (Janjevic, Winkenbach, and Rice 2020). Therefore, it is necessary to consider regular adjustments to the distribution network. In this case, there is usually a set of discrete time points, which divide the distribution network design into several different periods. This is known as the multi-period distribution network design. It solves the problem that the update period is long and cannot well adapt to the actual dynamic scenarios.

In addition, high customer satisfaction and low cost in the distribution network are urgently needed for supply chain management (Lim et al. 2006). For manufacturing companies that provide both products and services, there are more likely to be financial constraints (Jiang, Feng, and Yi 2021). Therefore, it is essential to evaluate the distribution network design in terms of both operating costs and customer satisfaction. Marketing91 classifies customer satisfaction on 1–5 levels, with level 1 representing the lowest customer satisfaction. These customers are likely to leave the company and even denigrate the company. While the other levels of customers will not guarantee a positive evaluation, but there will be no malicious slander (Bhasin 2020). Low-satisfied customers are particularly important to companies. At the same time, different companies have different evaluation periods for customer satisfaction. Some large companies measure overall customer satisfaction over a long period. For small and medium-sized companies, customer satisfaction may be assessed on a shorter period to avoid customer churn. In the logistics industry, customers are the bounded rational agents (Wang et al. 2021). People are often influenced by various factors to make 'irrational' decisions that are inconsistent with rational standards. Some people are risk neutral, while others may be risk averse. Therefore, it is very important to consider bounded rationality behaviour. The prospect theory (PT) value function can reflect the customers psychological criteria for satisfaction through a functional relationship between the actual service level and the service level reference point.

The 4PL network design (4PLND) problem is NP-hard, while the PT value function is S-shaped and non-convex (Wang et al. 2021). Obtaining the optimal solution to such complex, large-scale problem in a reasonable computational time is extremely difficult. Over the past few decades, meta-heuristics have received much attention due to their advantage and efficiency in solving large-scale real-world supply chain network design (SCND) problem. For the single-objective SCND problem, the particle swarm optimisation (PSO), genetic algorithm (GA) and differential evolution (DE) algorithm are widely used (Faramarzi-Oghani et al. 2022). Among them, PSO algorithm has good applicability in large-scale and complex networks (Soleimani and Kannan 2015; Huang, Lv, et al. 2019; Zhang et al. 2020; Wang et al. 2021). However, the PSO algorithm has the disadvantages of being prone to premature convergence and falling into the local optimum solution (Zhu and Zhang 2008). In this study, reinforcement learning is introduced to guide the local search strategy to help the PSO algorithm to jump out of the local optimum. Q-learning-based reinforcement learning can learn the optimal local search strategy by interacting with the environment through a trial-and-error procedure, and guarantee the optimal search state for the evolutionary process through reward/penalty techniques. More rigorous supports for the proposed algorithm will be provided in the numerical experiments in our study.

This study mainly focuses on the SOM supply chain network design problem, i.e. determining a multi-period distribution network while considering the boundedly rational behaviour of customers from a 4PL perspective, so as to maximise customer satisfaction under the constraints of limited investment budget and service levels.

The main contributions of this paper are summarised as the following: (i) A novel multi-period distribution network design problem with boundedly rational customers for the SOM supply chain from a 4PL perspective is proposed. We consider the constraints of investment budget and service level for service time and delivery quantity separately. (ii) By transforming the non-linear objective function into an equivalent novel model through a scenario-based linear reformulation, CPLEX can efficiently solve small and medium scale instances. (iii) The individual-driven Q-learning based memetic particle swarm optimisation (IDQLMPSO) algorithm is proposed to effectively and efficiently solve the proposed model when the computational complexity increases fast along with demands, periods, and network scales. (âĚş) From the perspective of 4PL provider mode, some management insights are given for investors on how to design a multi-period distribution network under the limited investment budget with different customer behaviour. The length of evaluation periods leads to differences in customer satisfaction due to different perceptions of 'small loss' and 'big gain' by boundedly rational customers.

The remainder of this paper is organised as follows. In Section 2, the related literature reviews are presented. Section 3 proposes the multi-period distribution network design model. The IDQLMPSO algorithm is proposed in Section 4. Numerical experiments are given in Section 5 to demonstrate the effectiveness and efficiency of the proposed algorithm, and related problem analysis is performed. Section 6 concludes the paper.

2. Literature review

Network design is one of the most important parts of the SOM supply chain management. In the past, researchers have studied network design problem in manufacturing supply chains under various names and terms, such as facility location (FL), SCND, production-distribution network design (PDND), distribution network design, 4PLND and so on. A comprehensive summary of research related to FL in supply chain management is reviewed (Melo, Nickel, and Saldanha-Da-Gama 2009). Govindan et al. reviewed the SCND problem under uncertainty from different perspectives (Govindan, Fattahi, and Keyvanshokooh 2017). A detailed review of the distribution network design problem is provided (Ambrosino and Scutella 2005). Yin et al. studied the 4PLND problem under uncertain environment. The 4PLND problem considers not only the decisions of logistics facilities and delivery quantity, but also the selection decisions of 3PL providers (Yin et al. 2022). Among them, SCND and PDND consider the entire production process, including the supply of raw materials. There are significant differences in distribution network design and 4PLND, which focus on the logistical service process in the downstream supply chain. For the distribution network design problem from a 4PL perspective, in addition to the traditional decisions (Hiremath, Sahu, and Tiwari 2013), it is also necessary to decide on 3PL providers and the corresponding delivery quantity between different locations. Therefore, it can be viewed as a mode in 4PLND.

The multi-period network design problem is clearly more complicated and has received more attention than single period. Research on multi-period network design problem usually considers the dynamic adjustment of facilities in different periods. Hinojosa et al. considered that the opening/closing of facilities can be changed in each period (Hinojosa, Puerto, and Fernández 2000). In recent years, more scholars have been studying multi-period network design problem with more practical significance. Hasani et al. considered multi-period and multi-product in a robust closed-loop SCND problem for perishable goods in agile manufacturing (Hasani, Zegordi, and Nikbakhsh 2012). Arampantzi et al. considered the selection of suppliers, the establishment of plants and DCs, and the choice of transportation methods and routes (Arampantzi, Minis, and Dikas 2019). Zhang et al. considered multi-period pricing decisions in a dynamic 4PLND problem under stochastic demand (Zhang, Huang, and Yin 2021). Currently, studies on multi-period network design problem are relatively comprehensive, but there are still relatively few studies from the 4PL perspective.

The cost and profit are usually used as key performance indicators to evaluate the effectiveness of network design. In recent years, more scholars take customer satisfaction as an indicator to measure network performance. Shen and Daskin considered the trade-offs between service level and cost in the integrated SCND problem (Shen and Daskin 2005). Lin studied the single-source capacitated FL problem and considered service level constraint (Lin 2009). Not only the location of DCs, but also the overall service level was considered in the distribution network design problem (Wang, Gunasekaran, and Ngai 2018). Sabouhi et al. minimised the expected total cost while ensuring that the minimum customer service level was achieved (Sabouhi et al. 2020). Huang et al. considered the psychological behaviour of customers from 4PL perspective and maximised demand satisfaction (Huang et al. 2021). Wang et al. maximised suppliers and customers satisfaction based on delivery quantity in a single-period 4PLND problem (Wang et al. 2021). Yin et al. minimised the 4PL network total cost under the service time constraint (Yin et al. 2021). In the current study, few studies consider bounded rationality behaviour when characterising customer satisfaction. At the same time, service level is only characterised from a single perspective.

Some scholars have used exact algorithms with the help of commercial software such as CPLEX to solve the 4PLND problem (Huang et al. 2021; Yin et al. 2021, 2022; Zhang, Huang, and Yin 2021). However, for large-scale 4PLND problem, the above algorithms may not provide an efficient solution. The PSO algorithm has the excellent application in solving the 4PLND problem. Many variants have been developed to improve its performance. The existing improved algorithms have mainly focused on the following five aspects: modified-based, hybrid-based, cooperative-based, micro-based, and memetic-based algorithms (Samma, Lim, and Saleh 2016). Among them, the memetic particle swarm optimisation (MPSO) algorithm has a good solution effect because it can combine the global search strength of PSO algorithm and the refinement ability of local search strategy. For the MPSO algorithm, the most critical challenges include when to call the local search strategy, how often to call the local search strategy, and which local search strategy should be used for each individual. Reinforcement learning, especially the Q-learning algorithm, can adaptively select an appropriate local search strategy at each step of the search process based on the search state and the historical performance of the operator. Therefore, the Q-learning algorithm is considered as a more advanced method than the traditional methods based on sequential selection and random selection (Karimi-Mamaghan et al. 2023). Samma et al. proposed a new reinforcement learning-based MPSO algorithm that addresses the first two challenges mentioned above. However, the above study adopts the same local search strategy for all individuals in the population and still does not address the third challenge. Therefore, a IDQLMPSO algorithm is proposed in this paper. The suitable local search strategy for each individual in the population is automatically selected.

In this study, we propose a model that integrates 3PL decisions from a 4PL perspective under the constraints of investment budget and service level. The IDQLMPSO algorithm is proposed to effectively and efficiently solve the above model with real-scale problems.

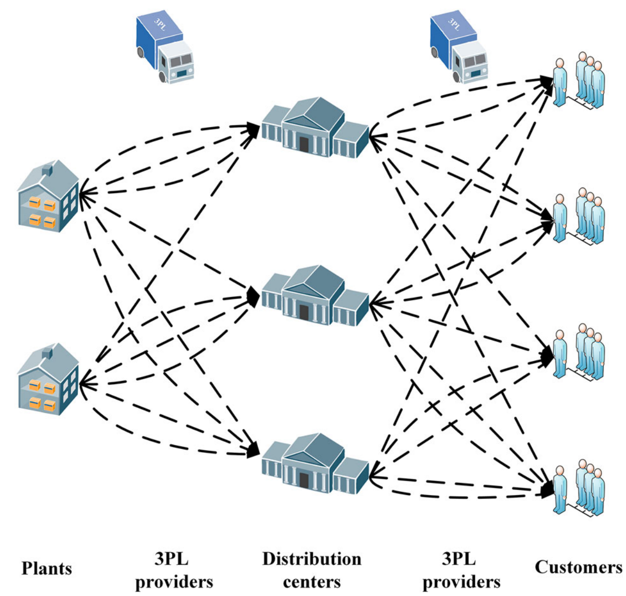


Figure 1. The potential distribution network.

3. Problem description and formulation

In this section, the multi-period distribution network design problem with boundedly rational customers from a 4PL perspective is described, and the notations used throughout the paper are introduced. In the first model, the minimum total cost of a multi-period distribution network that fully satisfies customer demands is calculated. When the investment budget is less than the minimum total cost, the second model is to establish a multi-period distribution network that maximises the minimum bounded rationality customer satisfaction. Finally, the mathematical model is linearly reformulated.

A manufacturing company (such as Haier) wants to invest in designing a multi-period distribution network to deliver its products and services from plants to customers through DCs and 3PL providers to reduce costs and improve customer satisfaction. As a result, it employs a 4PL provider to offer a comprehensive supply chain solution. Specifically, manufacturing companies are investors. A 4PL provider needs to help investors integrate 3PL providers, select the number and location of DCs, and complete the distribution of product flows from plants to customers in different periods. The potential distribution network is shown in Figure 1.

The main assumptions of the model are as follows:

(1) DCs and 3PL providers have four attributes, which are the fixed construction/ cooperation cost, unit product processing/ transportation cost, processing/ transportation capacity and processing/ transportation time.
(2) The capacity and cost of each potential DC and 3PL provider are predefined.
(3) For 3PL providers, the higher the unit transportation cost, the shorter the transportation time. At the same time, the unit transportation cost and transportation time are proportional to the distance.
(4) It is not necessary to satisfy all customer demands.
(5) The investment budget in each period is determined based on the minimum network cost.

The notations used throughout the paper are as follows:

Sets:

V	Set of all nodes in the network, including plants, DCs, and customers
P	Set of plants, $P \subset V$
D	Set of DCs, $D \subset V$
C	Set of customers, $C \subset V$
T	Set of periods, $t \in T, t = 1, 2, \ldots, NT$. NT is the maximum number of periods.
K_{ij}	Set of 3PL providers between $i \in V$ and $j \in V$

Parameters:

SA_i^t	Capacity of plant $i \in P$ in period $t \in T$
F_j^t	Fixed construction cost of DC $j \in D$ in period $t \in T$
U_j^t	Unit processing cost of DC $j \in D$ in period $t \in T$
Q_j^t	Processing capacity of DC $j \in D$ in period $t \in T$
PT_j^t	Processing time of DC $j \in D$ in period $t \in T$
f_{ijk}^t	Fixed cooperation cost of 3PL provider K_{ij} in period $t \in T$

u_{ijk}^t Unit transportation cost of 3PL provider K_{ij} in period $t \in T$
q_{ijk}^t Transportation capacity of 3PL provider K_{ij} in period $t \in T$
T_{ijk}^t Transportation time of 3PL provider K_{ij} in period $t \in T$
M^t Minimum network cost in period $t \in T$
I^t Investment budget in period $t \in T$
δ Investment budget coefficient
E_j^t Expected delivery time of customer $j \in C$ in period $t \in T$
G_j^t Service level of customer $j \in C$ in period $t \in T$
G_j^r Required service level of customer $j \in C$
$G_j^{\min}$ Service level limitation of customer $j \in C$

Decision variables:

x_{ijk}^t 1 if the 3PL provider K_{ij} is selected in period $t \in T$; otherwise, 0
y_j^t 1 if the DC $j \in D$ is selected in period $t \in T$; otherwise, 0
z_{ijk}^t The delivery quantity of the 3PL provider K_{ij} in period $t \in T$

3.1. Multi-period distribution network minimum cost

The multi-period distribution network design model that minimises the total cost when customer demands are fully satisfied is established as follows:

$$\min \sum_{t\in T}\left(\sum_{i\in D} F_i^t y_i^t(1-H_i^t) + \sum_{i\in V}\sum_{j\in V}\sum_{k\in K_{ij}} f_{ijk}^t x_{ijk}^t(1-h_{ijk}^t) + \sum_{i\in V}\sum_{j\in D}\sum_{k\in K_{ij}} U_j^t z_{ijk}^t + \sum_{i\in V}\sum_{j\in V}\sum_{k\in K_{ij}} u_{ijk}^t z_{ijk}^t\right) \tag{1}$$

$$H_i^t = \begin{cases} 1, \sum_{m=1}^{t-1} y_i^m > 0, \forall i \in D, t \in \mathrm{T}\backslash\{1\} \\ 0, \sum_{m=1}^{t-1} y_i^m = 0, \forall i \in D, t \in \mathrm{T}\backslash\{1\} \\ 0, \forall i \in D, t = 1 \end{cases} \tag{2}$$

$$h_{ijk}^t = \begin{cases} 1, \sum_{m=1}^{t-1} x_{ijk}^m > 0, \forall i \in V, j \in V, k \in K_{ij}, t \in \mathrm{T}\backslash\{1\} \\ 0, \sum_{m=1}^{t-1} x_{ijk}^m = 0, \forall i \in V, j \in V, k \in K_{ij}, t \in \mathrm{T}\backslash\{1\} \\ 0, \forall i \in V, j \in V, k \in K_{ij}, t = 1 \end{cases} \tag{3}$$

$$\sum_{j\in V}\sum_{k\in K_{ij}} z_{ijk}^t \le SA_i^t, \forall i \in P, t \in T \tag{4}$$

$$\sum_{j\in V}\sum_{k\in K_{ij}} z_{ijk}^t \le y_i^t Q_i^t, \forall i \in D, t \in T \tag{5}$$

$$z_{ijk}^t \le x_{ijk}^t q_{ijk}^t, \forall i \in V, j \in V, k \in K_{ij}, t \in T \tag{6}$$

$$\sum_{j\in V}\sum_{k\in K_{ij}} z_{ijk}^t = \sum_{j\in V}\sum_{k\in K_{ji}} z_{jik}^t, \forall i \in D, t \in T \tag{7}$$

$$x_{ijk}^t \le y_i^t, \forall i \in D, j \in V, k \in K_{ij}, t \in T \tag{8}$$

$$x_{jik}^t \le y_i^t, \forall i \in D, j \in V, k \in K_{ji}, t \in T \tag{9}$$

$$\sum_{j\in V}\sum_{k\in K_{ij}} x_{ijk}^t \ge y_i^t, \forall i \in D, t \in T \tag{10}$$

$$\sum_{j\in V}\sum_{k\in K_{ji}} x_{jik}^t \ge y_i^t, \forall i \in D, t \in T \tag{11}$$

$$\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t = D_j^t, \forall j \in C, t \in T \tag{12}$$

$$z_{ijk}^t \in N, \forall i \in V, j \in V, k \in K_{ij}, t \in T \tag{13}$$

$$x_{ijk}^t \in \{0,1\}, \forall i \in V, j \in V, k \in K_{ij}, t \in T \tag{14}$$

$$y_i^t \in \{0,1\}, \forall i \in D, t \in T \tag{15}$$

The objective function (1) is to minimise the total cost of the distribution network. Among them, the first two items are the total fixed costs; the last two items are the total processing and transportation costs. For constraints (2) and (3), the variable is 0 when $t = 1$. For $t \in T\backslash\{1\}$, as long as there is any period from 1 to $(t-1)$ when DCs and 3PL providers are selected, the variable is 1; otherwise, it is 0. Constraints (4)–(6) represent the capability limitations of plants, DCs and 3PL providers, respectively. Constraint (7) is the flow balance. Constraints (8) and (9) indicate that if a DC is not selected, 3PL providers that can provide services for it cannot be selected. Constraints (10) and (11) indicate that when a DC is selected, there must be at least one 3PL provider for it to deliver products. Constraint (12) indicates that customer demands must be fully satisfied. Constraint (13) indicates that the delivery quantity of 3PL providers is the non-negative integer. Finally, constraints (14) and (15) indicate that decision variables are both 0–1 variable.

3.2. The service level

Due to the limitation of investment budget, it is not possible to fully satisfy customers. We characterise the service level from the perspective of service time and delivery quantity, respectively.

Nowadays, service time has become an important factor for manufacturing company. T_{pdk}^t and T_{dck}^t are the transportation time from plants to DCs, and DCs to customers, respectively. The total service time $Time_c^t$ of customer c in period t is as follows:

$$Time_c^t = \max_{p\in P, d\in D, k\in K_{pd}\cup K_{dc}} \{T_{pdk}^t x_{pdk}^t + PT_d^t z_{pdk}^t + T_{dck}^t x_{dck}^t\}, \forall c \in C, t \in T \tag{16}$$

We use delivery quantity to describe the service level that manufacturing company can provide to customers. The specific expression is as follows:

$$G_j^t = \frac{\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t}{D_j^t}, \forall j \in C, t \in T \tag{17}$$

3.3. Customer satisfaction

In fact, customer satisfaction with logistics service depends not only on the service time and delivery quantity, but also on their bounded rationality behaviours. To characterise the impact of irrationality on customer satisfaction, this paper uses the value function in PT to express customer satisfaction (Huang et al. 2021; Wang et al. 2021).

The service level provided by manufacturing company to customers is higher than the required service level, which is called revenue; otherwise, it's called loss. The customer satisfaction can be expressed as follows:

$$S_j = \begin{cases} \left(\frac{1}{NT}\sum_{t=1}^{NT} G_j^t - G_j^r\right)^\alpha, & \frac{1}{NT}\sum_{t=1}^{NT} G_j^t \geq G_j^r \\ -\lambda\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT} G_j^t\right)^\beta, & \frac{1}{NT}\sum_{t=1}^{NT} G_j^t < G_j^r \end{cases}, \quad \forall j \in C \tag{18}$$

3.4. Multi-period distribution network design considering bounded rationality behaviour

When the investment budget is lower than the minimum total cost calculated above, the distribution network cannot fully satisfy customer demands. A 4PL provider needs to design network based on bounded rationality behaviour to maximise customer satisfaction. According to the description of customer satisfaction, the multi-period distribution network design mathematical model that maximises the minimum value of customer satisfaction is established as follows:

$$\max\left(\min_{j\in C} S_j\right) \tag{19}$$

s.t.

$$\sum_{i\in D} F_i^t y_i^t (1 - H_i^t) + \sum_{i\in V}\sum_{j\in V}\sum_{k\in K_{ij}} f_{ijk}^t x_{ijk}^t (1 - h_{ijk}^t) + \sum_{i\in V}\sum_{j\in D}\sum_{k\in K_{ij}} U_j^t z_{ijk}^t + \sum_{i\in V}\sum_{j\in V}\sum_{k\in K_{ij}} u_{ijk}^t z_{ijk}^t \leq I^t, \forall t \in T \tag{20}$$

(2) –(11)

$$\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t \leq D_j^t, \forall j \in C, t \in T \tag{21}$$

$$G_j^t \geq G_j^{\min}, \forall j \in C, t \in T \tag{22}$$

$$Time_j^t \leq E_j^t, \forall j \in C, t \in T \tag{23}$$

(13)–(15)

Formula (19) is the objective function, which maximises the minimum customer satisfaction in the multi-period distribution network. Constraint (20) indicates that the total network cost cannot exceed the investment budget in each period. Constraint (21) indicates that the total amount of products actually arrived cannot exceed demand, but the shortages are allowed. Constraint (22) indicates that the service level of each customer in each period must at least reach the service level limitation. Constraint (23) indicates that the service time cannot exceed the expected delivery time in each period.

3.5. The scenario-based linear reformulation

The multi-period distribution network design problem considering bounded rationality behaviour has been modelled. In the above model, the objective function (19) is a non-linear function, the constraint (20) is the quadratic constraint programming (QCP), and the rest of constraints are linear constraints. Therefore, this section transforms the non-linear objective into an equivalent linear objective (Huang, Zhang, et al. 2019).

The delivery quantity provided by the distribution network is an integer. If the delivery to customer j in period t is w, then the service level is $G_j^t = w/D_j^t, \forall j \in C, t \in T$. Constraint (22) limits the minimum delivery. Therefore, the range of delivery quantity to each customer in each period is $[G_j^{\min} \times D_j^t, D_j^t]$. There are $\lfloor D_j^t \times (1 - G_j^{\min}) + 1 \rfloor$ potential scenarios for the delivery quantity for each customer in each period. We regard the delivery quantity for the same customer in different periods as one scenario, so there are $L = \prod_{j=1}^{NC}\prod_{t=1}^{NT} \lfloor D_j^t \times (1 - G_j^{\min}) + 1 \rfloor$ scenarios in total. For customer $j \in C$, the delivery quantity in each period

t under scenario l is $w_{lt}, \forall t \in T, l = \lceil G_j^{\min} \times D_j^t \rceil \ldots D_j^t$. Figure 2 presents an example of a scenario tree for one customer with two scenarios and three periods. To decide the delivery quantity of each customer in different periods, we design the decision variable $g_{jl} \in \{0, 1\}, \forall j \in C, l = \lceil G_j^{\min} \times D_j^t \rceil \ldots D_j^t$. When $g_{jl} = 1$, it indicates that the delivery quantity for customer j in each period t is w_{lt}. We can calculate customer satisfaction $S_{jl}, \forall j \in C, l = \lceil G_j^{\min} \times D_j^t \rceil \ldots D_j^t$ according to formula (24).

$$S_{jl} = \begin{cases} \left(\frac{1}{NT} \sum_{t=1}^{NT} \frac{w_{lt}}{D_j^t} - G_j^r \right)^{\alpha}, & \frac{1}{NT} \sum_{t=1}^{NT} \frac{w_{lt}}{D_j^t} \geq G_j^r \\ -\lambda \left(G_j^r - \frac{1}{NT} \sum_{t=1}^{NT} \frac{w_{lt}}{D_j^t} \right)^{\beta}, & \frac{1}{NT} \sum_{t=1}^{NT} \frac{w_{lt}}{D_j^t} < G_j^r \end{cases},$$
$$\forall j \in C \tag{24}$$

Then, we transform the original model to the following model:

$$\max S^{\min} \tag{25}$$

s.t. (2)–(11) (20)

$$\sum_{l=1}^{L} g_{jl} \times S_{jl} \geq S^{\min}, \forall j \in C \tag{26}$$

$$\sum_{l=1}^{L} g_{jl} = 1, \forall j \in C \tag{27}$$

$$\sum_{i \in D} \sum_{k \in K_{ij}} z_{ijk}^t = \sum_{l=1}^{L} g_{jl} w_{lt}, \forall j \in C, t \in T \tag{28}$$

$$T_{pdk}^t x_{pdk}^t + PT_d^t z_{pdk}^t + T_{dck}^t x_{dck}^t \leq E_d^t, \forall p \in P, d \in D, c \in C, k \in K_{pd} \cup K_{dc}, t \in T \tag{29}$$

(13)–(15)

The objective function (25) and constraints (26)–(28) are equivalent to the original objective function (19) and constraint (22). Constraint (29) and constraint (23) are equivalent. After the above transformation, the new objective function is linear and the constraints are QCP and linear, so the model can be solved by CPLEX.

Theorem 3.1: *The multi-period distribution network design model considering bounded rationality behaviour is equivalent to the model after the scenario-based linear reformulation.*

Proof: See Appendix 1. ■

The CPLEX solver can solve the equivalent model after the scenario-based linear reconstruction and provide the optimal solution on the small-scale instance. However, as customer demands, periods, and network scales increase, the scenarios of the problem grow exponentially. For the large-scale problem, the CPLEX solver cannot obtain a feasible solution in a limited time even with the linear reformulation. Therefore, in Section 4, we propose the corresponding solution algorithm to deal with the above problem.

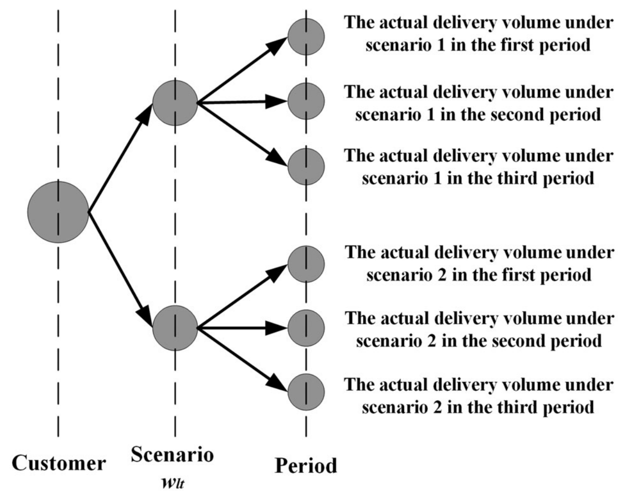


Figure 2. An example of scenario tree.

4. The individual-driven Q-learning based memetic particle swarm optimisation algorithm

Since the network design problem is NP-hard (Shen, Zhan, and Zhang 2011) and the scenarios in the transformed model grow exponentially, a meta-heuristic algorithm is proposed to solve the problem. The MPSO algorithm, which integrates the global search ability of the PSO algorithm and the refinement ability of the local search strategy, has been successfully applied to solve various optimisation problems such as continuous nonlinear and discrete optimisation. The individual-driven Q-learning (IDQL) algorithm is proposed to address the challenge of which local search strategy to use for each individual faced by the MPSO algorithm. Furthermore, the minimum cost maximum flow algorithm is designed in Section 4.3 to obtain the maximum network delivery quantity. Meanwhile, four different types of repair strategies are designed in Section 4.6. In this section, the proposed IDQLMPSO algorithm is introduced in detail.

4.1. Encoding and decoding strategies

Encoding Strategy: We adopt a one-dimensional integer encoding method and each node in the distribution network is labelled in the order of plants, DCs, and customers. An adjacency matrix is established according

$$
\begin{array}{c} \\ 1' \\ 2' \\ 3' \\ 4' \\ 5' \end{array}
\begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{bmatrix} 0 & 0 & 3 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 2 \end{bmatrix} \end{array}
\Longrightarrow
\begin{array}{c} \\ 1' \\ 2' \\ 3' \\ 4' \\ 5' \end{array}
\begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{bmatrix} [0,0] & [0,0] & [0,7] & [0,1] & [0,3] & [0,0] & [0,0] & [0,0] & [0,0] \\ [0,0] & [0,0] & [0,1] & [0,7] & [0,3] & [0,0] & [0,0] & [0,0] & [0,0] \\ [0,0] & [0,0] & [0,0] & [0,0] & [0,0] & [0,3] & [0,3] & [0,1] & [0,1] \\ [0,0] & [0,0] & [0,0] & [0,0] & [0,0] & [0,1] & [0,3] & [0,1] & [0,3] \\ [0,0] & [0,0] & [0,0] & [0,0] & [0,0] & [0,1] & [0,0] & [0,3] & [0,3] \end{bmatrix} \end{array}
$$

Figure 3. An example of the encoding strategy.

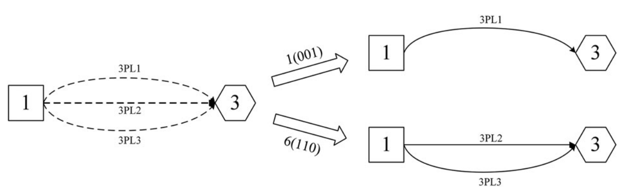


Figure 4. An example of the decoding strategy.

to 3PL providers between two nodes. Then according to the number of optional arcs between two nodes in the adjacency matrix, the integer encoding range of the particle is determined. Specifically, assuming that there are n 3PL providers between two nodes, the position encoding range is $[0, 2^n - 1]$, and the velocity encoding range is $[-(2^n - 1), 2^n - 1]$. For example, the potential distribution network shown in Figure 1 can be transformed into the adjacency matrix in Figure 3, which also gives the corresponding particle position encoding range.

The length of encoding, that is, the spatial dimension of particle is determined by the product of number of non-zero elements in the adjacency matrix and period T. We take the adjacency matrix in Figure 3 as an example, assuming that $T = 3$, the encoding length is $3*17 = 51$.

Decoding Strategy: The binary decoding is adopted to convert integer encoding to 0–1 binary, the selection of 3PL providers between any two nodes can be judged in order from the lowest to the highest. At the same time, the decisions of DCs can be judged according to 3PL providers. Taking node 1 and 3 in Figure 3 as an example, the situation where the decision variables are obtained through the decoding strategy is shown in Figure 4.

4.2. Population initialisation

Population initialisation is to randomly initialise the position and velocity of each particle in each dimension by choosing a random integer in the range. Therefore, the decision variables $x^t_{ijk}, \forall i \in V, j \in V, k \in K_{ij}, t \in T$ and $y^t_j, \forall j \in D, t \in T$ can be obtained.

4.3. Minimum cost maximum flow algorithm

When the DCs and 3PL providers is determined, the multi-period distribution network design problem is transformed into the problem of allocating delivery quantity. The minimum cost maximum flow (MCMF) algorithm is used to solve the above problem. Since the MCMF algorithm is suitable for solving network problems from a single source node to a single sink node while the intermediate nodes have no cost and capacity restrictions, we respectively add a virtual source node and a virtual sink node. We connect and translate the capability and cost of plants and customer demands through virtual arcs. Similarly, we add virtual intermediate nodes and virtual arcs to represent the cost, capability, and time properties of DCs. The MCMF algorithm is used to calculate the delivery quantity of 3PL providers to ensure that the constraints (4)–(7), (13) and (18) are satisfied. However, the result of flow assignment may violate (20), (22) or (23), which will be dealt with in the solution repair process.

4.4. Fitness function

In this paper, maximising the minimum customer satisfaction in the multi-period distribution network is taken as the fitness function. The objective is to find the particle that maximises the minimum value of all customer satisfaction among all particles that satisfy the constraints.

4.5. Update formula

The update methods of the particle velocity and position in the population are shown in equations (30) and (31):

$$v^{n+1}_{iD} = wv^n_{iD} + c_1\xi(p^n_{iD} - x^n_{iD}) + c_2\eta(p^n_{gD} - x^n_{iD}) \quad (30)$$

$$x^{n+1}_{iD} = x^n_{iD} + v^{n+1}_{iD} \quad (31)$$

where v_{iD}^{n+1} and x_{iD}^{n+1} represent the velocity and position of particle i in $(n+1)$th generation, both of which are integers. $p_i^n = (p_{i1}^n, p_{i2}^n, \ldots, p_{iD}^n)$ is the best historical point of the individual particle search in nth generation. $p_g^n = (p_{g1}^n, p_{g2}^n, \ldots, p_{gD}^n)$ is the best historical point of all particles in the population found in nth generation. c_1 and c_2 are learning factors, $\xi, \eta \in U[0, 1]$. In order to better balance the exploration and development capabilities of the PSO algorithm, this paper adopts the time-varying inertia weight w, as shown in the following formula:

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{N} \times n \tag{32}$$

4.6. Repair strategies

According to the encoding, decoding strategy and MCMF algorithm, the definite distribution network can be obtained. For the above network, four types of infeasible solutions may appear. First, the updated velocity and position values may exceed their encoding range. Second, the distribution network may be disconnected. Third, the total cost may exceed the investment budget. Finally, service time and service level may not satisfy limits. For any infeasible solution, it needs to be repaired.

(1) *Encoding range violation*: For the case that the velocity/ position encoding value exceeds its encoding range, the repair strategy we adopt is to take the boundary value.

(2) *Network disconnection repair*: According to the network structure, disconnection can be roughly divided into three situations. (a) The customers are not connected to the network, and there are no 3PL providers to serve customers. (b) The plants are not connected to the network, resulting in products cannot reach customers through DCs. (c) The DCs are disconnected to the network, resulting in the network not meeting the constraint of flow balance. When the above situation appears, we randomly add one or more arcs at the corresponding disconnected nodes. The specific network disconnection repair is shown in Figure 5.

(3) *Cost adjustment strategy*: For the constraint (20), the total cost consists of two parts: fixed cost and transportation processing cost. In the first stage, a connected network is determined, and the fixed cost can be determined according to the selection of DCs and 3PL providers. According to the logistics network determined and the customer demands, the MCMF algorithm is used to obtain the delivery quantity and transportation processing cost in the second stage. If the fixed cost is higher than the investment budget, the delivery quantity is allocated 0. Otherwise, customer demands are adjusted according to the formula below.

$$\Delta D_j^t = D_j^t - D_j^t \times \frac{\Delta I^t}{I^t} \times \frac{1}{NC} \tag{33}$$

Among them, ΔI^t is the difference between the total network cost and the investment budget. NC is the number of customers, and ΔD_j^t is the demand of customer j in period t after cost adjustment. Then, the demand is rounded up as the current demand for the next calculation. We should check before and after each adjustment until the constraint (20) is satisfied.

(4) *The service level repair*: When actual service level is less than the service level limitation or the service time exceeds the expected delivery time, the delivery quantity of customer j in current period t is adjusted to 0.

There may be new infeasible solutions generated during the repair process, so it is necessary to repair distribution network several times until all constraints are satisfied.

4.7. Local search strategies

This section introduces four local search strategies combined with the PSO algorithm to generate the MPSO algorithm. The following four local search strategies are designed:

(1) Randomly select the position where the particle changes, and change the value of that position within the encoding range.
(2) Randomly select two positions of the particle and exchange the values of the two positions.
(3) A range of particle is randomly selected and swapped.
(4) Two positions of the particle are randomly selected, the values of these two positions are placed in the head of the new arrangement, and the values of the remaining positions are arranged in the original order.

4.8. Individual-driven Q-learning algorithm

The MPSO algorithm improves the performance of PSO algorithm to a certain extent. In this paper, the IDQL algorithm is proposed to determine when and which local search strategy to use for each individual. The definition of the state, action, reward function, and

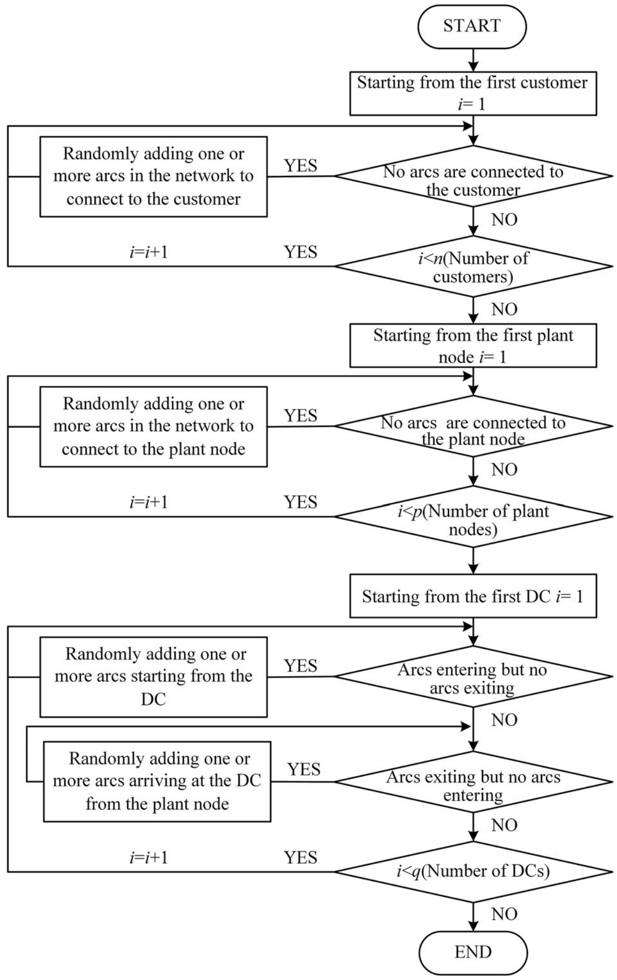


Figure 5. Network disconnection repair.

Q-value update are the key elements that affect the algorithm performance.

4.8.1. State

The states in Q-learning represent the characteristics of the external environment. In this paper, the fitness of each individual in the population is defined as the current state. When the maximum number of iterations is N, the number of individuals in the population is M, and the state space $S = \{s_{11}, s_{12}, \ldots, s_{NM}\}$.

4.8.2. Action

Four local search strategies are defined as the alternative actions, and the action space is $A = \{a_1, a_2, a_3, a_4\}$.

4.8.3. Reward function

The reward function reinforces the effective actions of each state by rewarding improving states while punishing degenerate states. We set the reward function for each individual in the population. Assuming that $fit(n)(m)$ is the fitness of the mth individual in the current nth generation, the reward function is set as follows:

$$reward(n)(m) = fit(n)(m) - fit(n-1)(m) \tag{34}$$

4.8.4. Individual-driven Q-value

The Q-value reflects the search ability under a certain state-action, that is, the utility of the local search strategy for different individuals. We design an individual-driven Q-Value, and the update is defined as:

$$\begin{aligned} Q(s_{(n+1)m}, a_i) = Q(s_{nm}, a_i) + \alpha_n[r_{nmi} \\ + \gamma \arg\max_{a \in A} Q(s_{(n+1)m}, a) \\ - Q(s_{nm}, a_i)], \\ \forall n \in N, m \in M, i \in A \end{aligned} \tag{35}$$

In this paper, we use the following dynamic adjustment α_n (Rakshit et al. 2013):

$$\alpha_n = 1 - \left(0.9 \times \frac{n}{N}\right) \tag{36}$$

The specific steps of the IDQL algorithm are shown in Algorithm 1.

4.9. An individual-driven Q-learning based memetic particle swarm optimisation algorithm

The PSO algorithm is combined with four local search strategies to form the MPSO algorithm. To select the most appropriate action to achieve the best performance based on the state of each individual in the population, we design an individual-driven Q-learning algorithm. Combining the IDQL algorithm with the MPSO algorithm can effectively solve the current challenges faced by MPSO algorithm. The specific steps of the IDQLMPSO algorithm are shown in Algorithm 2.

5. Numerical experiments

In this section, numerical experiments are designed to analyse the algorithm performance and significance of the multi-period manufacturing distribution network design considering 4PL provider and bounded rationality behaviour. Firstly, we present numerical experimental schemes and performance indexes. Secondly, we compare GA and DE algorithm that are also widely used in the SCND problem (Faramarzi-Oghani et al. 2022). At the same time, for each individual, the sequentially selected MPSO (SMPSO) and the randomly selected MPSO (RMPSO) (Karimi-Mamaghan et al. 2023) are compared. Therefore, the solution results of CPLEX, PSO, GA, DE, SMPSO, RMPSO, QLMPSO (Samma, Lim, and Saleh 2016), and IDQLMPSO algorithm are compared. Finally, we analyse the impact of related problem

Algorithm 1: IDQL algorithm.

Input: Current iteration: n; Population: M; $local$; $GBfit(n)$; Other parameters
Output: Q-value
$average = \sum_{i=1}^{M} (GBfit(n) - fit(X_i))/M$
if $local = 0$ **then**
  if $GBfit(n) - fit(X_i) \geq average$ **then**
    Update velocity and position of particle i according to (30) and (31);
  end if
  if $0 \leq GBfit(n) - fit(X_i) < 0.25 \times average$ **then**
    Perform action a_1;
  end if
  if $0.25 \times average \leq GBfit(n) - fit(X_i) < 0.5 \times average$ **then**
    Perform action a_2;
  end if
  if $0.5 \times average \leq GBfit(n) - fit(X_i) < 0.75 \times average$ **then**
    Perform action a_3;
  end if
  if $0.75 \times average \leq GBfit(n) - fit(X_i) < average$ **then**
    Perform action a_4;
  end if
end if
if $local \neq 0$ **then**
  if $GBfit(n) - fit(X_i) \geq average$ **then**
    Update velocity and position of particle i according to (30) and (31);
  end if
  if $GBfit(n) - fit(X_i) < average$ **then**
    for $j = 1$ to 4 **do**
    if $Q(s_{ni}, a) > Q(s_{ni}, a_j), a \in A, a \neq j$ **then**
      $best \leftarrow a$;
    end if
    end for
    Perform action a_{best};
  end if
end if
Calculate the corresponding $Reward(n)(i)(best)$;
Update $Q(s_{(n+1)i}, a_{best})$;
if $j \neq best$ **do**
  $Q(s_{(n+1)i}, a_j) \leftarrow Q(s_{ni}, a_j)$;
end if
return Q-value

Algorithm 2: IDQLMPSO algorithm.

Input: Iteration: N; Population: M; Other parameters
Output: $GBest, GBfit(N), LBfit(N)(M)$
for $i = 1$ to M **do**
  Random initialisation velocity and position of particle i;
  Allocate delivery quantity and repair the infeasible solutions;
  Evaluate particle $fit(X_i)$ and set $pBest_i = X_i, LBfit(0)(i) \leftarrow fit(X_i)$;
end for
$GBfit(0) = \max\{fit(X_a)\}, a \in M, GBest \leftarrow pBest_a, local = 0$, $judge(1) = 1$;
for $n = 1$ to N **do**
  if $judge(n) = 1$ **then**
    for $i = 1$ to M **do**
    Update velocity and position of particle i;
    Allocate delivery quantity and repair the infeasible solutions;
    Evaluate particle $fit(X_i)$;
    if $fit(X_i) > fit(pBest_i)$ **then**
      $LBfit(n)(i) \leftarrow fit(X_i), pBest_i = X_i$;
    end if
    if $fit(pBest_i) > GBfit(n-1)$ **then**
      $GBfit(n) = fit(pBest_i), GBest \leftarrow pBest_i, decision = 1$;
    end if
    end for
    if $decision = 1$ **then**
      $judge(n+1) = 1$;
    end if
    if $decision \neq 1$ **then**
      $GBfit(n) = GBfit(n-1), judge(n+1) = 2$;
    end if
      $Q(s_{(n+1)m}, a) \leftarrow Q(s_{nm}, a)$;
    end if
  if $judge(n) = 2$ **then**
    for $i = 1$ to M **do**
      Algorithm 1;
      Allocate delivery quantity and repair the infeasible solutions;
      Evaluate particle $fit(X_i)$;
      if $fit(X_i) > fit(pBest_i)$ **then**
        $LBfit(n)(i) \leftarrow fit(X_i), pBest_i = X_i$;
      end if
      if $fit(pBest_i) > GBfit(n-1)$ **then**
        $GBfit(n) = fit(pBest_i), GBest \leftarrow pBest_i, decision = 1$;
      end if
    end for
    if $decision = 1$ **then**
      $judge(n+1) = 2$;
    end if
    if $decision \neq 1$ **then**
      $GBfit(n) = GBfit(n-1), judge(n+1) = 1$;
    end if
  end if
$local = local + 1$;
end for
return $GBest, GBfit(N), LBfit(N)(M)$

parameters on distribution network design. The above algorithms are coded in C, the version of the CPLEX solver is IBM ILOG CPLEX 12.8, and the instances are solved using the Intel(R) Xeon(R) Gold 6146 CPU @ 3.20 GHz.

5.1. Experimental design

In the actual case, the Cainiao logistics network currently has seven DCs in China. Assuming that each province in China has one retailer (customer), the actual case scale is around 50 nodes (Wang et al. 2021). Therefore, combined with the actual case scale, we design numerical experiments with 9, 14, 24 and 51 nodes respectively. We set period $T = 3$, and there are 3 selectable 3PL providers in the distribution network. The transportation time in the dataset is calculated based on the actual distance between two nodes, and the unit transportation cost and fixed cooperation cost are inversely proportional to the transportation time. Some variable parameters are assumed to follow the uniform distributions (Huang et al. 2021; Shen, Zhan, and Zhang 2011). The capacity of each plant is taken from U [1000, 10,000], the unit processing cost of each DC is taken from U [30, 50], the processing capacity of each DC is taken from U [500, 1000], the processing time of each DC is taken from U [0.06, 0.1], the transportation capacity of each 3PL provider is taken from U [40, 180], the unit shortage cost of each customer is taken from U [250, 500], the expected delivery time of each customer is taken from U [120, 150], the demand of each customer is taken from U [50, 240], the required service level of each customer is 0.7, the service level limitation

Table 1. Data set.

Instances	P	D	C	The minimum cost $t = 1$	$t = 2$	$t = 3$
9 nodes	2	3	4	102,610	78,776	117,636
14 nodes	3	4	7	413,160	396,730	521,898
24 nodes	7	6	11	553,045	545,952	738,606
51 nodes	10	7	34	1,413,175	1,654,020	2,206,680

Table 2. Parameter combination schemes for four scale instances.

Instances	M	N	γ	w_{max}	w_{min}	c_1	c_2
9 nodes	120	100	0.9	1.1	0.3	3.5	0.5
14 nodes	130	110	0.8	1.0	0.3	2.5	1.5
24 nodes	150	130	0.8	0.9	0.3	2.5	1.5
51 nodes	180	150	0.8	0.9	0.3	2.5	1.5

is 0.5. According to previous research, both α and β are set to 0.88 and $\lambda = 2.25$ in PT (Huang et al. 2021; Wang et al. 2021). The distribution network structure and the corresponding minimum cost are shown in Table 1.

The performance indexes used in parameters setting are explained as follows: the best solution (*BS*), the worst solution (*WS*), the mean value (*Mean*), the sample standard error (*SE*) and the average time (*Time*) of R runs. Among them, *SE* is calculated as shown in equation (37).

$$SE = \frac{\sqrt{\frac{1}{R-1}\sum_{i=1}^{R}(x_i - Mean)^2}}{Mean} \times 100\% \tag{37}$$

5.2. Performance analysis

The parameters in the algorithm are adjusted under four different scale numerical experiments to find the most reasonable parameter combination scheme of the IDQLMPSO algorithm. In the process of adjustment, only one parameter is changed each time while keeping other parameters unchanged, and the number of runs is $R = 30$. The detailed parameter tuning process is shown in Appendix 2. The most reasonable parameter combinations of the IDQLMPSO algorithm under the four scale numerical experiments are shown in Table 2.

To demonstrate the effectiveness and efficiency of the IDQLMPSO algorithm, we compare the solution results of CPLEX, PSO, GA, DE, SMPSO, RMPSO and QLMPSO algorithm. Each algorithm runs 30 times under the respective optimal parameter combination scheme. Table 3 shows the comparative results of CPLEX and seven algorithms under four scales numerical experiments. For the three basic algorithms, both the GA and DE algorithms have longer solution times than the PSO algorithm. In most cases, the performance of PSO algorithm is the best among the three indicators of *BS*, *WS*, and *Mean*. The PSO algorithm may not perform well in terms of *SE* due to being easier to fall into the local optimal solution. We propose the IDQLMPSO algorithm to overcome its shortcomings. For different improvement strategies, the Q-learning algorithm performs better in most cases. At the same time, the method of randomly selecting the local search strategy for each individual also performs well. These demonstrate that the Q-learning algorithm and the improvement for each individual are effective.

The optimal solution provided by the CPLEX solver can be used as the benchmark solution to verify the performance of our proposed algorithm on the small-scale instance. It can be seen from the comparison that in the 9-node instance, IDQLMPSO algorithm can solve the same optimal solution as CPLEX in a shorter time. For the 14-node instance, CPLEX can obtain an efficient solution within a set time of 7200s. The IDQLMPSO algorithm can obtain a better *BS* in less time. For large-scale instances, CPLEX cannot obtain a feasible solution in a limited time. Indeed, the quality and efficiency of the CPLEX solver are highly dependent on the algorithm complexity. The algorithm complexity grows exponentially rapidly as the number of the constraints and decision variables increases. However, the heuristic algorithm is mainly affected by the population size and the number of iterations, the proposed IDQLMPSO algorithm is not greatly affected by the instance scale. Although the advantage is not obvious on the small-scale instance, the IDQLMPSO algorithm shows absolute superiority on the quality and running time as the scale increases.

From the above performance indexes, it can be seen that the proposed algorithm can not only effectively help the PSO algorithm to get out of the problem of falling into the local optimal solution, but also overcome the shortcoming that the local search strategy takes too long time to solve. The IDQLMPSO algorithm can obtain a better solution in a shorter time.

To compare the convergence and stability of the algorithms, we plot the convergence plots and boxplots for different scale instances in Figure 6. Among the seven algorithms, the IDQLMPSO algorithm is obviously more advantageous in terms of convergence speed and stability.

5.3. Problem analysis

This section discusses the impact of 4PL provider, investment budgets, different customer behaviour, and

Table 3. Numerical experimental results.

Instances	Algorithm	*BS*	*WS*	*Mean*	Gap_{BS}(%)	Gap_{Mean}(%)	*SE* (%)	*Time* (s)
9 nodes	CPLEX	**0.308242**	**0.308242**	**0.308242**	**0**	**0**	–	245.4
	PSO	0.308033	0.280093	0.296796	0.068	3.71	2.50572	**30.8574**
	GA	0.298829	0.272253	0.287515	3.05	6.72	2.67731	51.7469
	DE	0.308033	0.278035	0.295134	0.068	4.25	2.39251	47.4287
	SMPSO	**0.308242**	0.272951	0.298408	**0**	3.19	1.72414	43.5304
	RMPSO	**0.308242**	0.294213	0.299554	**0**	2.82	1.11962	44.5331
	QLMPSO	0.308033	0.298829	0.299625	0.068	2.8	**0.796984**	55.596
	IDQLMPSO	**0.308242**	0.298829	0.300273	**0**	2.59	1.07351	35.9623
14 nodes	CPLEX	0.241072	**0.241072**	**0.241072**	2.34	**2.34**	–	7200
	PSO	0.235425	0.223077	0.232222	4.63	5.93	1.2248	404.64
	GA	0.234967	0.219896	0.231803	4.81	6.09	1.17655	618.346
	DE	0.235425	0.223986	0.232057	4.63	5.99	1.20572	585.13
	SMPSO	0.235425	0.22781	0.232492	4.63	5.82	0.826711	486.437
	RMPSO	0.235425	0.22781	0.232876	4.63	5.66	0.872653	475.73
	QLMPSO	0.235425	0.22992	0.233046	4.63	5.59	0.810278	823.446
	IDQLMPSO	**0.241284**	0.233323	0.235093	**2.25**	4.76	**0.606722**	**382.587**
24 nodes	CPLEX	–	–	–	–	–	–	–
	PSO	0.294998	0.261813	0.279926	–	–	4.17207	1543.78
	GA	0.294998	0.26028	0.277469	–	–	3.72818	3202.59
	DE	0.296647	0.261562	0.279618	–	–	3.15239	3246.91
	SMPSO	0.296647	0.262374	0.286573	–	–	3.64702	1770.25
	RMPSO	0.296647	0.264194	0.28691	–	–	3.38868	1766.55
	QLMPSO	0.296647	0.264444	0.288833	–	–	2.89945	2586.82
	IDQLMPSO	**0.298683**	**0.285968**	**0.292088**	–	–	**1.19704**	**1380.66**
51 nodes	CPLEX	–	–	–	–	–	–	–
	PSO	0.212863	0.162188	0.190144	–	–	7.97728	7606.82
	GA	0.212495	0.158356	0.187456	–	–	6.67264	9912.42
	DE	0.212863	0.161033	0.188494	–	–	4.52608	10029.4
	SMPSO	0.213368	0.183657	0.201891	–	–	2.77783	8687.95
	RMPSO	0.213519	0.184147	0.20262	–	–	2.97145	8765.03
	QLMPSO	0.213932	0.185149	0.202992	–	–	3.20429	9176.95
	IDQLMPSO	**0.216309**	**0.199708**	**0.209389**	–	–	**2.12666**	**6869.71**

Note: The bold values represent the optimal values for each performance index.

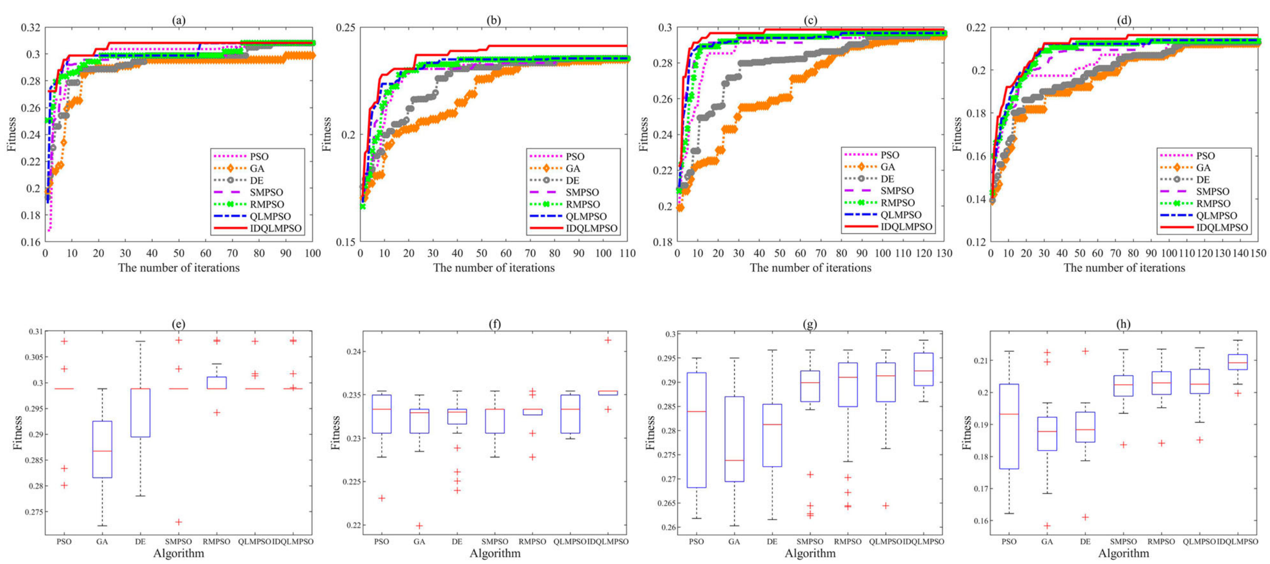


Figure 6. The convergence plots and boxplots for different scale instances.

customer satisfaction evaluation periods on the distribution network through the CPLEX solver in the 9-node instance and provides some management insights.

5.3.1. The influence of 4PL provider

To investigate the influence of 4PL provider on the multi-period distribution network, we compare 4PL provider

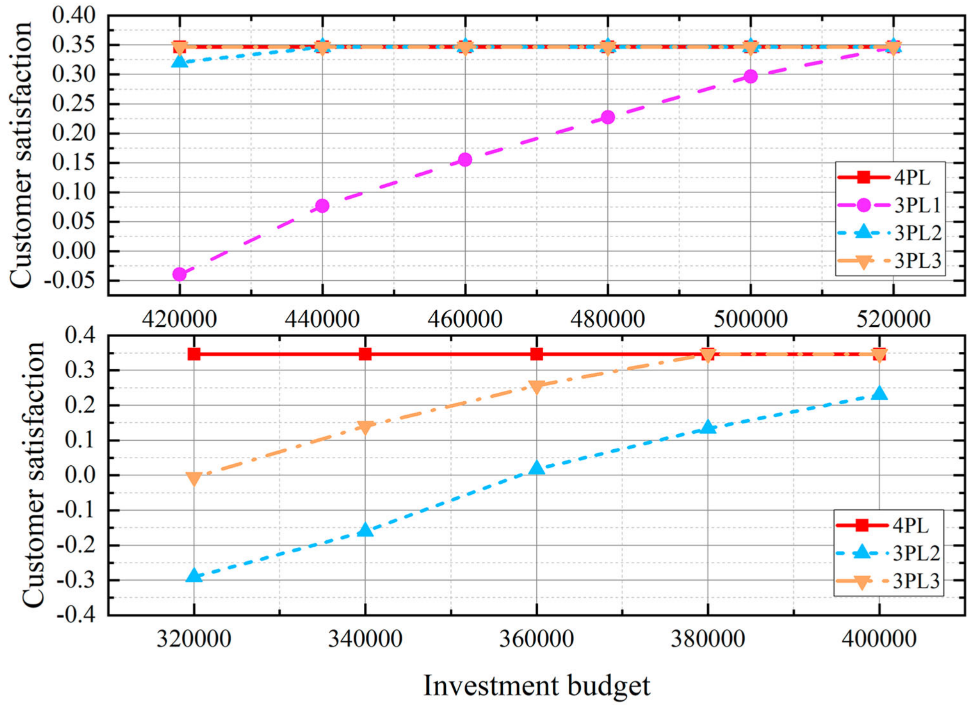


Figure 7. The influence of 4PL provider on customer satisfaction.

Table 4. The influence of 4PL provider on cost.

	The minimum cost for each period			
Cases	$t = 1$	$t = 2$	$t = 3$	The minimum total cost
4PL	102,610	78,776	117,636	299,022
3PL1	218,300	118,420	177,720	514,440
3PL2	195,250	92,130	138,300	425,680
3PL3	182,350	77,550	116,340	376,240

with the single 3PL provider. The comparison results in Table 4 show that 4PL provider is effective in reducing the total cost of distribution network when the customer demands are fully satisfied.

When the investment budget is insufficient, we compare the 4PL provider mode with a single 3PL provider mode. We allocate the investment budget for each period according to the proportion of the minimum cost per period. From the results in Figure 7, we can find that a 4PL provider can help manufacturing companies improve customer satisfaction under the same investment budget. Therefore, it is verified by numerical experiments that a 4PL provider can reduce cost and improve customer satisfaction of the multi-period distribution network.

5.3.2. The influence of investment budget

The investment budget is an important factor affecting network structure and customer satisfaction. When the investment budget coefficient δ is less than 1, the network cannot fully satisfy customer demands. This section discusses the impact of the change in δ on customer satisfaction.

The result in Figure 8 shows that the change in customer satisfaction is more dramatic in the 'loss' state than in the 'gain' state when the investment budget coefficient changes equally. This indicates that customers are more sensitive to 'loss' than to 'gain'. This also verifies the risk-averse characteristics in the face of 'gain' and risk-seeking characteristics in the face of 'loss' for bounded rationality customers. When customer satisfaction is around 0, the service level is close to the required service level of customers. Around this point, that is, when the investment budget coefficient is between 0.65 and 0.7, increasing the unit investment budget changes the most in customer satisfaction. Therefore, when the investment budget is not sufficient, additional investment around the required service level makes the most sense.

5.3.3. The influence of different customer behaviour

In this section, we compare the effects of rational customers and boundedly rational customers with different behavioural characteristics on the problem. We assume that $\alpha = 1, \beta = 1, \lambda = 1$ represents the rational customer, that is, risk neutral. There are two types of boundedly rational customers, $\alpha = 0.88, \beta = 0.88, \lambda = 2.25$ for risk seeker (Tversky and Kahneman 1992), $\alpha = 0.52, \beta = 0.52, \lambda = 1.51$ for risk averse (Xu, Zhou, and Xu 2011).

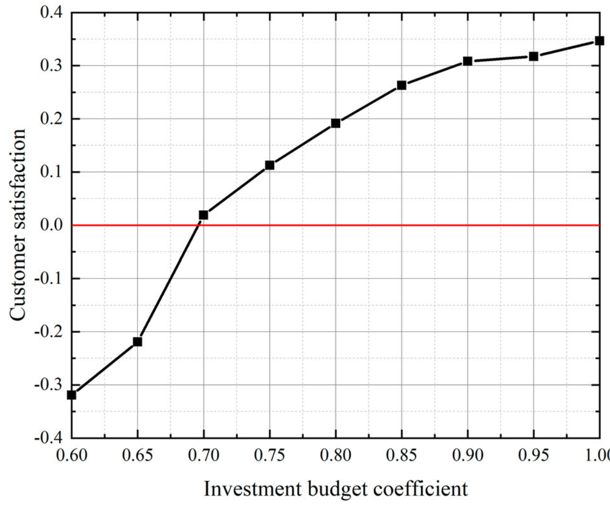


Figure 8. The influence of investment budget.

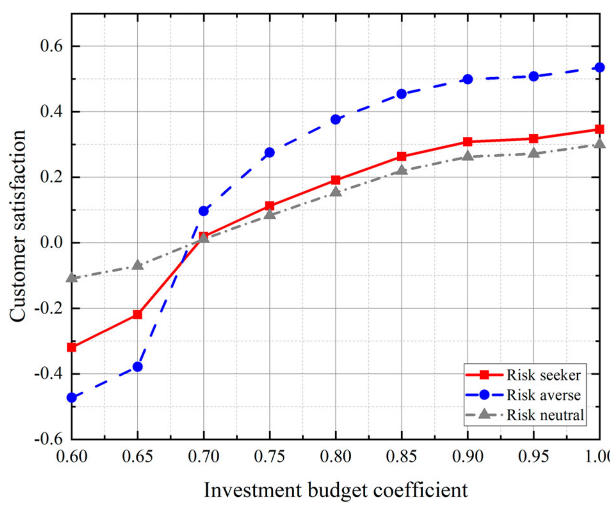


Figure 9. The influence of different customer behaviour.

The results in Figure 9 show that when actual service level reaches the required service level, that is, when customer is in the 'gain' state, considering bounded rationality behaviour can help manufacturing companies design a distribution network at a relatively low cost to improve customer satisfaction. Conversely, when the service level is low, considering bounded rationality behaviour can urge manufacturing companies to increase investment as soon as possible and rapidly improve customer satisfaction.

At the same time, it can be found that under the 'gain' state, when the investment budget is the same, the satisfaction of risk-averse customers is the highest, and the satisfaction of risk-seeking customers is the lowest. When the 'loss' state, the above situation is exactly the opposite. This indicates that risk-averse customers prefer to 'seek stability', showing more satisfaction when faced with 'gain' and the more dramatic decline in satisfaction when faced with 'loss'. However, risk-seekers prefer 'risk-taking behaviour' and expect to seek greater gains, so they are less satisfied than risk-averse customers when faced with 'gain'. They are more 'risk tolerant' and their satisfaction declines more slowly when faced with 'loss'. At the same time, manufacturing companies need more investment to be able to satisfy the required service level of risk-seeking customers.

5.3.4. The influence of customer satisfaction evaluation periods

In this paper, we characterise the long-period customer satisfaction by the average service level over all periods. As shown in equation (18), we call this overall long-period evaluation M1. In addition, as shown in equations (38)–(39), we can obtain customer satisfaction through the service level per period, and then obtain an average customer satisfaction. This short-period evaluation is called M2.

$$S_j^t = \begin{cases} (G_j^t - G_j^r)^\alpha, & G_j^t \geq G_j^r \\ -\lambda(G_j^r - G_j^t)^\beta, & G_j^t < G_j^r \end{cases}, \quad \forall j \in C, t \in T \tag{38}$$

$$S_j = \frac{1}{NT}\sum_{t=1}^{NT} S_j^t, \quad \forall j \in C \tag{39}$$

By comparison in Figure 10, it can be found that when the investment budget coefficient is 0.7, M1 has achieved the required service level, while the customer satisfaction of M2 is less than 0. This shows that the 'small loss' and 'big gain' can be integrated through the long-period customer satisfaction, and the negative emotion brought by the 'loss' can be offset by the pleasure brought by the 'gain'. At the same time, manufacturing companies can achieve the required service level at a lower cost by the long-period customer satisfaction. However, when the investment budget is insufficient and the service level is low, manufacturing companies can choose the short-period evaluation. Therefore, under different service levels, manufacturing companies can choose different periods to evaluate customer satisfaction so as to maximise customer satisfaction with the same investment.

6. Conclusion

In this paper, a novel multi-period distribution network design problem with boundedly rational customers for the SOM supply chain from a 4PL perspective is proposed.

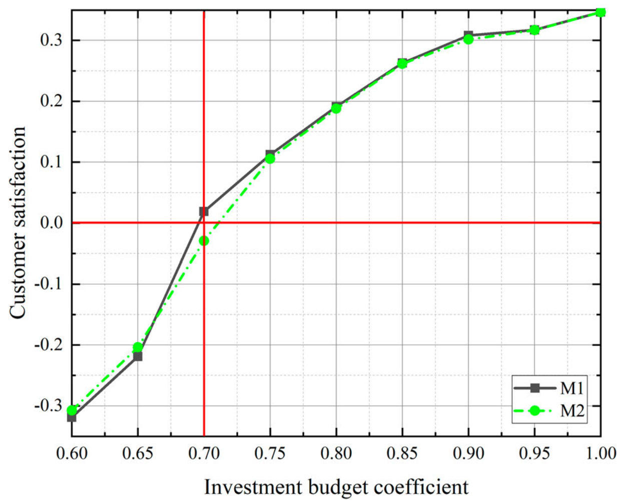


Figure 10. The influence of customer satisfaction evaluation periods.

- The service level is considered for service time and delivery quantity separately. A mixed integer non-linear programming model is established to obtain a network solution with maximum customer satisfaction based on PT under the constraints of investment budget and service level.
- To deal with the non-linear objective function, a scenario-based equivalent linear reformulation is proposed, and the small and medium scale distribution networks can be efficiently solved by the CPLEX solver.
- Aiming at the computation difficulty brought by a large number of scenarios caused by demands, periods and network scales, the IDQLMPSO algorithm is proposed. Numerical experiments at different scales show that the proposed algorithm outperforms the CPLEX solver in terms of solution time and solution quality in most cases.

The problem analysis shows that a 4PL provider can reduce total cost and improve customer satisfaction in a multi-period distribution network compared to the single 3PL provider. At the same time, the multi-period distribution network design will be affected by the boundedly rational customer behaviour, so the investment strategy of designing the distribution network can be adjusted according to the analysis.

Behavioural analysis shows that boundedly rational customers are risk-averse in the face of 'gain' and risk-seeking in the face of 'loss'. The manufacturing companies need more investment to satisfy the required service level of risk-seeking customers. Finally, by analysing the customer satisfaction evaluation periods, the long-period customer satisfaction evaluation can integrate 'small loss' and 'big gain' and achieve the required service level at a lower cost. At the same time, we present the short-period evaluation to help manufacturing companies choose different evaluation periods at different service levels, thereby maximising customer satisfaction with the same investment.

In the post-epidemic era, supply chain resilience plays a crucial role in protecting supply chains against the large-scale disruption (Rahman et al. 2022). How to measure resilience under stochastic disruptions is a big challenge we are facing. The simulation-based methods can be used to evaluate resiliency strategies impact on supply chain performance and select the best response plan before the next disruption (Xu et al. 2015; Moosavi and Hosseini 2021; Goodwin et al. 2022). Further, the proposed IDQLMPSO algorithm can be combined with the optimal computing budget allocation method (Chen et al. 2000; Chen and Lee 2011; Zhang et al. 2017) to solve the resilient supply chain network design problem with stochastic disruptions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the [National Key R&D Program of China] under Grant [number 2021YFB3300900]; [NSFC Major International (Regional) Joint Research Project] under Grant [number 71620107003]; [Fundamental Research Funds for State Key Laboratory of Synthetical Automation for Process Industries] under Grant [number 2013ZCX11]; [111 Project 2.0] under Grant [number B08015].

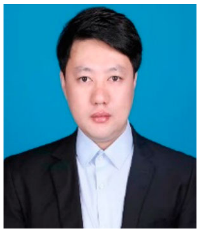
Notes on contributors

Yuxin Zhang received the B.Eng. degree in automatic and the M.Sc. degree in control science and engineering from Northeastern University, Shenyang, China, in 2019 and 2021, respectively. His current research interests include supply chain logistics network design, heuristic algorithms, and machine learning.

Zheming Gao received the Ph.D. degree in operations research from North Carolina State University in 2020. He is currently an Assistant Professor with Northeastern University (Shenyang, China). He has published many academic papers in journals such as *European Journal of Operational Research* and *Knowledge-Based Systems*, etc. His research interests include methodologies of machine learning, optimisation theories and numerical optimisation and algorithm design.

Min Huang received the B.Sc. degree in automatic instrument, the M.Sc. degree in systems engineering, and the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 1990, 1993, and 1999, respectively. She is currently a Professor with Northeastern University. She has published over 100 journal articles, books, and refereed conference papers. Her research interests include modeling and optimisation for logistics and supply chain.

Songchen Jiang received the B.Eng. degree in automatic and the M.Sc. degree in control science and engineering from Northeastern University, Shenyang, China, in 2019 and 2021, respectively. His current research interests include supply chain operation and management and data-driven optimisation technique.



Mingqiang Yin received the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China. His main research interests include logistics supply chain network design, optimisation theory and methods, etc. He has published more than 20 academic papers in journals and conferences such as *Computers & Industrial Engineering* and *Journal of Intelligent Manufacturing*, etc. He is a reviewer for the international journals *Computers & Industrial Engineering* and *Annals of Operations Research*.

Shu-Cherng Fang is currently a Professor with North Carolina State University Raleigh, North Carolina, USA. Professor Fang has published over two hundred refereed journal articles. He is currently serving on 22 editorial boards of scientific journals in the general area of Operations Research, Industrial/Systems Engineering and Management Science. He is also the Editor-in-Chief of *Fuzzy Optimization and Decision Making*. His research interests include Nonlinear Programming, Fuzzy Optimisation and Decision Making, Soft Computing, and Logistics and Supply Chain Management.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author, MH. The data are not publicly available due to their containing information that could compromise the privacy of research participants.

References

Ambrosino, Daniela, and Maria Grazia Scutella. 2005. "Distribution Network Design: New Problems and Related Models." *European Journal of Operational Research* 165 (3): 610–624. doi:10.1016/j.ejor.2003.04.009.

Arampantzi, Christina, Ioannis Minis, and Georgios Dikas. 2019. "A Strategic Model for Exact Supply Chain Network Design and Its Application to a Global Manufacturer." *International Journal of Production Research* 57 (5): 1371–1397. doi:10.1080/00207543.2018.1489155.

Bhasin, Hitesh. 2020. "Five Levels of Customer Satisfaction." *The Marketing91*, May 13. https://www.marketing91.com/five-levels-of-customer-satisfaction/#: ∼ :text = Customer%20satisfaction%20of%20a%20particular%20company%20can%20also,customer%20who%20is%20not%20satisfied%20with%20the%20company.

Büyüközkan, Gülçin, Orhan Feyzioğlu, and Mehmet Şakir Ersoy. 2009. "Evaluation of 4PL Operating Models: A Decision Making Approach Based on 2-Additive Choquet Integral." *International Journal of Production Economics* 121 (1): 112–120. doi:10.1016/j.ijpe.2008.03.013.

Chen, Chun-Hung, and Loo Hay Lee. 2011. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. Vol. 1. Singapore: World scientific.

Chen, Chun-Hung, Jianwu Lin, Enver Yücesan, and Stephen E. Chick. 2000. "Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization." *Discrete Event Dynamic Systems* 10 (3): 251–270. doi:10.1023/A:1008349927281.

Cheng, Evelyn. 2021. "'Made in China' Products Are Running into New Logistics Problems." *The CNBC*, August 3. https://www.cnbc.com/2021/08/04/shipping-made-in-china-products-run-into-new-logistics-problems.html.

Faramarzi-Oghani, Sohrab, Parisa Dolati Neghabadi, El-Ghazali Talbi, and Reza Tavakkoli-Moghaddam. 2022. "Meta-heuristics for Sustainable Supply Chain Management: A Review." *International Journal of Production Research*, 1–31. doi:10.1080/00207543.2022.2045377.

Gattorna, John, and Theresa Jones. 1998. *Strategic Supply Chain Alignment: Best Practice in Supply Chain Management*. Aldershot, UK: Gower Publishing.

Goodwin, Travis, Jie Xu, Nurcin Celik, and Chun-Hung Chen. 2022. "Real-Time Digital Twin-Based Optimization with Predictive Simulation Learning." *Journal of Simulation*, 1–18. doi:10.1080/17477778.2022.2046520.

Govindan, Kannan, Mohammad Fattahi, and Esmaeil Keyvanshokooh. 2017. "Supply Chain Network Design under Uncertainty: A Comprehensive Review and Future Research Directions." *European Journal of Operational Research* 263 (1): 108–141. doi:10.1016/j.ejor.2017.04.009.

Hasani, Aliakbar, Seyed Hessameddin Zegordi, and Ehsan Nikbakhsh. 2012. "Robust Closed-Loop Supply Chain Network Design for Perishable Goods in Agile Manufacturing under Uncertainty." *International Journal of Production Research* 50 (16): 4649–4669. doi:10.1080/00207543.2011.625051.

He, Na, Zhong-Zhong Jiang, Jian Wang, Minghe Sun, and Guanghe Xie. 2021. "Maintenance Optimisation and Coordination with Fairness Concerns for the Service-Oriented Manufacturing Supply Chain." *Enterprise Information Systems* 15 (5): 694–724. doi:10.1080/17517575.2020.1746406.

Hinojosa, Yolanda, Justo Puerto, and Francisco R. Fernández. 2000. "A Multiperiod Two-Echelon Multicommodity Capacitated Plant Location Problem." *European Journal of Operational Research* 123 (2): 271–291. doi:10.1016/S0377-2217(99)00256-8.

Hiremath, Nagayya C., Sadananda Sahu, and Manoj Kumar Tiwari. 2013. "Multi Objective Outbound Logistics Network Design for a Manufacturing Supply Chain." *Journal

of Intelligent Manufacturing 24 (6): 1071–1084. doi:10.1007/s10845-012-0635-8.
Huang, Fang, Juhong Chen, Linhui Sun, Yaqi Zhang, and Shujun Yao. 2020. “Value-based Contract for Smart Operation and Maintenance Service Based on Equitable Entropy.” *International Journal of Production Research* 58 (4): 1271–1284. doi:10.1080/00207543.2019.1617450.
Huang, Min, Liwei Dong, Hanbin Kuang, Zhong-Zhong Jiang, Loo Hay Lee, and Xingwei Wang. 2021. “Supply Chain Network Design Considering Customer Psychological Behavior-a 4PL Perspective.” *Computers & Industrial Engineering* 159: 107484. doi:10.1016/j.cie.2021.107484.
Huang, Han, Liang Lv, Shujin Ye, and Zhifeng Hao. 2019. “Particle Swarm Optimization with Convergence Speed Controller for Large-Scale Numerical Optimization.” *Soft Computing* 23 (12): 4421–4437. doi:10.1007/s00500-018-3098-9.
Huang, Min, Jun Tu, Xiuli Chao, and Delong Jin. 2019. “Quality Risk in Logistics Outsourcing: A Fourth Party Logistics Perspective.” *European Journal of Operational Research* 276 (3): 855–879. doi:10.1016/j.ejor.2019.01.049.
Huang, Min, Hui Zhang, Hanbin Kuang, Yang Yu, Loo Hay Lee, and Xingwei Wang. 2019. “Flexible Truckload Pickup and Delivery Problem Considering Reserved Orders and Fuel Consumption.” *Computers & Industrial Engineering* 138: 106117. doi:10.1016/j.cie.2019.106117.
Janjevic, Milena, Matthias Winkenbach, and James B. Rice Jr. 2020. “5 Reasons Why Your Supply Chain Design Must Change with the Times.” *The Supply Chain 24/7*, August 24. https://www.supplychain247.com/article/5_reasons_why_your_supply_chain_design_must_change_with_the_times.
Jiang, Zhong-Zhong, Guangqi Feng, and Zelong Yi. 2021. “How Should a Capital-Constrained Servicizing Manufacturer Search for Financing? The Impact of Supply Chain Leadership.” *Transportation Research Part E: Logistics and Transportation Review* 145: 102162. doi:10.1016/j.tre.2020.102162.
Jiang, Zhong-Zhong, Guangqi Feng, Zelong Yi, and Xiaolong Guo. 2022. “Service-Oriented Manufacturing: A Literature Review and Future Research Directions.” *Frontiers of Engineering Management* 9: 71–88. doi:10.1007/s42524-021-0171-3.
Karimi-Mamaghan, Maryam, Mehrdad Mohammadi, Bastien Pasdeloup, and Patrick Meyer. 2023. “Learning to Select Operators in Meta-Heuristics: An Integration of Q-Learning into the Iterated Greedy Algorithm for the Permutation Flowshop Scheduling Problem.” *European Journal of Operational Research* 304: 1296–1330. doi:10.1016/j.ejor.2022.03.054.
Lee, Emma. 2016. “Alibaba’s Cainiao Logistics Confirms First Financing At $7.7B Valuation.” *The technode*, March 15. https://technode.com/2016/03/15/alibaba-cainiao-funding/.
Lim, Seok Jin, Suk Jae Jeong, Kyung Sup Kim, and Myon Woong Park. 2006. “Hybrid Approach to Distribution Planning Reflecting a Stochastic Supply Chain.” The International Journal of Advanced Manufacturing Technology 28 (5): 618–625. doi:10.1007/s00170-004-2398-7.
Lin, C. K. Y. 2009. “Stochastic Single-Source Capacitated Facility Location Model with Service Level Requirements.” *International Journal of Production Economics* 117 (2): 439–451. doi:10.1016/j.ijpe.2008.11.009.
Melo, M. Teresa, Stefan Nickel, and Francisco Saldanha-Da-Gama. 2009. “Facility Location and Supply Chain Management – A Review.” *European Journal of Operational Research* 196 (2): 401–412. doi:10.1016/j.ejor.2008.05.007.
Moosavi, Javid, and Seyedmohsen Hosseini. 2021. “Simulation-based Assessment of Supply Chain Resilience with Consideration of Recovery Strategies in the COVID-19 Pandemic Context.” *Computers & Industrial Engineering* 160: 107593. doi:10.1016/j.cie.2021.107593.
Rahman, Towfique, Sanjoy Kumar Paul, Nagesh Shukla, Renu Agarwal, and Firouzeh Taghikhah. 2022. “Supply Chain Resilience Initiatives and Strategies: A Systematic Review.” *Computers & Industrial Engineering* 170: 108317. doi:10.1016/j.cie.2022.108317.
Rakshit, Pratyusha, Amit Konar, Pavel Bhowmik, Indrani Goswami, Swagatam Das, Lakhmi C. Jain, and Atulya K. Nagar. 2013. “Realization of an Adaptive Memetic Algorithm Using Differential Evolution and Q-Learning: A Case Study in Multirobot Path Planning.” *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43 (4): 814–831. doi:10.1109/TSMCA.2012.2226024.
Rymaszewska, Anna, Petri Helo, and Angappa Gunasekaran. 2017. “IoT Powered Servitization of Manufacturing – An Exploratory Case Study.” *International Journal of Production Economics* 192: 92–105. doi:10.1016/j.ijpe.2017.02.016.
Sabouhi, Fatemeh, Mohammad Saeed Jabalameli, Armin Jabbarzadeh, and Behnam Fahimnia. 2020. “A Multi-cut L-Shaped Method for Resilient and Responsive Supply Chain Network Design.” *International Journal of Production Research* 58 (24): 7353–7381. doi:10.1080/00207543.2020.1779369.
Samma, Hussein, Chee Peng Lim, and Junita Mohamad Saleh. 2016. “A New Reinforcement Learning-Based Memetic Particle Swarm Optimizer.” *Applied Soft Computing* 43: 276–297. doi:10.1016/j.asoc.2016.01.006.
Shen, Zuo-Jun Max, and Mark S. Daskin. 2005. “Trade-offs between Customer Service and Cost in Integrated Supply Chain Design.” *Manufacturing & Service Operations Management* 7 (3): 188–207. doi:10.1287/msom.1050.0083.
Shen, Zuo-Jun Max, Roger Lezhou Zhan, and Jiawei Zhang. 2011. “The Reliable Facility Location Problem: Formulations, Heuristics, and Approximation Algorithms.” *INFORMS Journal on Computing* 23 (3): 470–482. doi:10.1287/ijoc.1100.0414.
Soleimani, Hamed, and Govindan Kannan. 2015. “A Hybrid Particle Swarm Optimization and Genetic Algorithm for Closed-Loop Supply Chain Network Design in Large-Scale Networks.” *Applied Mathematical Modelling* 39 (14): 3990–4012. doi:10.1016/j.apm.2014.12.016.
Tversky, Amos, and Daniel Kahneman. 1992. “Advances in Prospect Theory: Cumulative Representation of Uncertainty.” *Journal of Risk and Uncertainty* 5 (4): 297–323. doi:10.1007/BF00122574.
Wang, Gang, Angappa Gunasekaran, and Eric W. T. Ngai. 2018. “Distribution Network Design with Big Data: Model and Analysis.” *Annals of Operations Research* 270 (1): 539–551. doi:10.1007/s10479-016-2263-8.
Wang, Huihui, Min Huang, W. H. Ip, and Xingwei Wang. 2021. “Network Design for Maximizing Service Satisfaction of Suppliers and Customers under Limited Budget for Industry Innovator Fourth-Party Logistics.” *Computers & Industrial Engineering* 158: 107404. doi:10.1016/j.cie.2021.107404.

Xu, Jie, Edward Huang, Chun-Hung Chen, and Loo Hay Lee. 2015. "Simulation Optimization: A Review and Exploration in the New Era of Cloud Computing and Big Data." *Asia-Pacific Journal of Operational Research* 32 (03): 1550019. doi:10.1142/S0217595915500190.

Xu, Hongli, Jing Zhou, and Wei Xu. 2011. "A Decision-Making Rule for Modeling Travelers' Route Choice Behavior Based on Cumulative Prospect Theory." *Transportation Research Part C: Emerging Technologies* 19 (2): 218–228. doi:10.1016/j.trc.2010.05.009.

Yin, Mingqiang, Min Huang, Xiaohu Qian, Dazhi Wang, Xingwei Wang, and Loo Hay Lee. 2021. "Fourth-party Logistics Network Design with Service Time Constraint under Stochastic Demand." *Journal of Intelligent Manufacturing*, 1–25. doi:10.1007/s10845-021-01843-7.

Yin, Mingqiang, Min Huang, Xingwei Wang, and Loo Hay Lee. 2022. "Fourth-Party Logistics Network Design under Uncertainty Environment." *Computers & Industrial Engineering* 167: 108002. doi:10.1016/j.cie.2022.108002.

Zhang, Xin, Ke-Jing Du, Zhi-Hui Zhan, Sam Kwong, Tian-Long Gu, and Jun Zhang. 2020. "Cooperative Coevolutionary Bare-Bones Particle Swarm Optimization with Function Independent Decomposition for Large-Scale Supply Chain Network Design with Uncertainties." *IEEE Transactions on Cybernetics* 50 (10): 4454–4468. doi:10.1109/TCYB.2019.2937565.

Zhang, Yuxin, Min Huang, and Mingqiang Yin. 2021. "Dynamic Network Design for Fourth Party Logistics Considering Multi-period Pricing under Stochastic Demand." Paper presented at the 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, May 22–24.

Zhang, Si, Jie Xu, Loo Hay Lee, Ek Peng Chew, Wai Peng Wong, and Chun-Hung Chen. 2017. "Optimal Computing Budget Allocation for Particle Swarm Optimization in Stochastic Optimization." *IEEE Transactions on Evolutionary Computation* 21 (2): 206–219. doi:10.1109/TEVC.2016.2592185.

Zhu, Guang-Yu, and Wei-Bo Zhang. 2008. "Drilling Path Optimization by the Particle Swarm Optimization Algorithm with Global Convergence Characteristics." *International Journal of Production Research* 46 (8): 2299–2311. doi:10.1080/00207540601042480.

Appendices

Appendix 1. Proof of Theorem 1

In this section, we prove that the objective function (19), the constraints (21)–(23) and the objective function (25), the constraints (26)–(29) are equivalent.

Through the constraints (21)–(22), it can be known that the delivery quantity $\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t, \forall j \in C, t \in T$ is an integer value between $G_j^{\min} \times D_j^t$ and D_j^t. The scenario $w_{lt}, \forall l \in L, t \in T$ ensures that the above constraints are satisfied. Constraint (28) represents the delivery quantity. Constraint (23) and Constraint (29) are equivalent. Therefore, the objective functions are equivalent to prove that the scenario-based linear reformulation is equivalent.

$$S_j = \begin{cases} \left(\frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t}{D_j^t} - G_j^r\right)^\alpha, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t}{D_j^t} \ge G_j^r \\ -\lambda\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t}{D_j^t}\right)^\beta, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{i\in D}\sum_{k\in K_{ij}} z_{ijk}^t}{D_j^t} < G_j^r \end{cases}$$

$$= \begin{cases} \left(\frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{l=1}^{L} g_{jl}\times w_{lt}}{D_j^t} - G_j^r\right)^\alpha, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{l=1}^{L} g_{jl}\times w_{lt}}{D_j^t} \ge G_j^r \\ -\lambda\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{l=1}^{L} g_{jl}\times w_{lt}}{D_j^t}\right)^\beta, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{\sum_{l=1}^{L} g_{jl}\times w_{lt}}{D_j^t} < G_j^r \end{cases}$$

$$= \begin{cases} \left(\frac{1}{NT}\sum_{t=1}^{NT}\sum_{l=1}^{L} g_{jl}\times \frac{w_{lt}}{D_j^t} - G_j^r\right)^\alpha, & \frac{1}{NT}\sum_{t=1}^{NT}\sum_{l=1}^{L} g_{jl}\times \frac{w_{lt}}{D_j^t} \ge G_j^r \\ -\lambda\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\sum_{l=1}^{L} g_{jl}\times \frac{w_{lt}}{D_j^t}\right)^\beta, & \frac{1}{NT}\sum_{t=1}^{NT}\sum_{l=1}^{L} g_{jl}\times \frac{w_{lt}}{D_j^t} < G_j^r \end{cases} \tag{A1}$$

Because of the decision variable $g_{jl} \in \{0,1\}, \forall j \in C, l = \lceil G_j^{\min} \times D_j^t \rceil \ldots D_j^t$ and $\sum_{l=1}^{L} g_{jl} = 1, \forall j \in C$, we can obtain:

$$\left(\frac{1}{NT}\sum_{t=1}^{NT}\sum_{l=1}^{L} g_{jl}\times\frac{w_{lt}}{D_j^t} - G_j^r\right)^\alpha = \left[\sum_{l=1}^{L} g_{jl}\times\left(\frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} - G_j^r\right)\right]^\alpha = \sum_{l=1}^{L} g_{jl}\times\left(\frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} - G_j^r\right)^\alpha \tag{A2}$$

$$-\lambda\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\sum_{l=1}^{L} g_{jl}\times\frac{w_{lt}}{D_j^t}\right)^\beta = -\lambda\sum_{l=1}^{L} g_{jl}\times\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t}\right)^\beta \tag{A3}$$

$$S_j = \begin{cases} \sum_{l=1}^{L} g_{jl}\times\left(\frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} - G_j^r\right)^\alpha, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} \ge G_j^r \\ -\lambda\sum_{l=1}^{L} g_{jl}\times\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t}\right)^\beta, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} < G_j^r \end{cases}$$

$$= \sum_{l=1}^{L} g_{jl}\times\begin{cases} \left(\frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} - G_j^r\right)^\alpha, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} \ge G_j^r \\ -\lambda\left(G_j^r - \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t}\right)^\beta, & \frac{1}{NT}\sum_{t=1}^{NT}\frac{w_{lt}}{D_j^t} < G_j^r \end{cases}$$

$$= \sum_{l=1}^{L} g_{jl}\times S_{jl} \tag{A4}$$

The objective function (19) can be equivalently transformed into objective function (25) and constraint (26). Therefore, the linearly reformulated model is equivalent to the original model.

Appendix 2. Parameters tuning

In the parameter adjustment process of the IDQLMPSO algorithm, only one parameter is adjusted each time and the other parameters are fixed. First, the 9-node instance experiment is presented. The analysis of the population size M is shown in Table A1. Through the comparison of the results based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different M values, it is found that it is better to set M equal to 120.

The analysis of the maximum number of iterations N is shown in Table A2. Through the comparison of the results

Table A1. Parameter tuning for the population size (M).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
60	60	0.9	1.1	0.3	3.5	1.0	0.298829	0.234011	0.281687	9.35647	10.9384
90	60	0.9	1.1	0.3	3.5	1.0	0.298829	0.234011	0.29332	5.7114	15.1145
120	60	0.9	1.1	0.3	3.5	1.0	0.308033	0.234011	0.296603	4.14869	20.6441
150	60	0.9	1.1	0.3	3.5	1.0	0.308033	0.234011	0.296031	4.07477	23.2845

Table A2. Parameter tuning for the maximum number of iterations (N).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
120	60	0.9	1.1	0.3	3.5	1.0	0.308033	0.234011	0.296603	4.14869	20.6441
120	80	0.9	1.1	0.3	3.5	1.0	0.308033	0.234011	0.296766	4.05421	27.9707
120	100	0.9	1.1	0.3	3.5	1.0	0.308242	0.280093	0.29946	1.63302	32.3064
120	120	0.9	1.1	0.3	3.5	1.0	0.308242	0.294213	0.299245	0.859415	40.1194

Table A3. Parameter tuning for the decay rate (γ).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
120	100	0.9	1.1	0.3	3.5	1.0	0.308242	0.280093	0.29946	1.63302	32.3064
120	100	0.8	1.1	0.3	3.5	1.0	0.308242	0.234011	0.298002	4.37571	33.8405
120	100	0.7	1.1	0.3	3.5	1.0	0.298829	0.234011	0.293807	5.64457	34.7489

Table A4. Parameter tuning for the maximal inertia weight (w_{max}).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
120	100	0.9	1.2	0.3	3.5	1.0	0.308242	0.283401	0.298259	1.00038	30.4202
120	100	0.9	1.1	0.3	3.5	1.0	0.308242	0.280093	0.29946	1.63302	32.3064
120	100	0.9	1.0	0.3	3.5	1.0	0.308033	0.234011	0.297121	4.16388	31.0123
120	100	0.9	0.9	0.3	3.5	1.0	0.308033	0.234011	0.28851	8.10975	30.3709

based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different N values, it is found that it is better to set N equal to 100.

The analysis of the decay rate γ is shown in Table A3. Through the comparison of the results based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different γ values, it is found that it is better to set γ equal to 0.9.

The analysis of the maximal inertia weight w_{max} is shown in Table A4. Through the comparison of the results based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different w_{max} values, it is found that it is better to set w_{max} equal to 1.1.

The analysis of the minimal inertia weight w_{min} is shown in Table A5. Through the comparison of the results based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different w_{min} values, it is found that it is better to set w_{min} equal to 0.3.

The analysis of the self-learning factor c_1 is shown in Table A6. Through the comparison of the results based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different c_1 values, it is found that it is better to set c_1 equal to 3.5.

The analysis of the social learning factor c_2 is shown in Table A7. Through the comparison of the results based on the values of *BS*, *WS*, *Mean*, *SE*, and *Time* with different c_2 values, it is found that it is better to set c_2 equal to 0.5.

The parameter tuning for the remaining three scale instances is based on the same process described above.

Table A5. Parameter tuning for the minimal inertia weight (w_{min}).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
120	100	0.9	1.1	0.4	3.5	1.0	0.308033	0.234011	0.295129	5.68555	30.5828
120	100	0.9	1.1	0.3	3.5	1.0	0.308242	0.280093	0.29946	1.63302	32.3064
120	100	0.9	1.1	0.2	3.5	1.0	0.308242	0.292532	0.298933	0.702598	31.6657

Table A6. Parameter tuning for the self-learning factor (c_1).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
120	100	0.9	1.1	0.3	3.5	1.0	0.308242	0.280093	0.29946	1.63302	32.3064
120	100	0.9	1.1	0.3	3.0	1.0	0.308033	0.234011	0.29481	5.1424	32.1458
120	100	0.9	1.1	0.3	2.5	1.0	0.308033	0.234011	0.296731	4.45186	30.9129
120	100	0.9	1.1	0.3	2.0	1.0	0.3049	0.234011	0.292352	5.59137	34.2503

Table A7. Parameter tuning for the social learning factor (c_2).

M	N	γ	w_{max}	w_{min}	c_1	c_2	*BS*	*WS*	*Mean*	*SE* (%)	*Time* (s)
120	100	0.9	1.1	0.3	3.5	2.0	0.308033	0.234011	0.285885	8.8057	31.0825
120	100	0.9	1.1	0.3	3.5	1.5	0.308242	0.278612	0.298516	1.72427	32.3064
120	100	0.9	1.1	0.3	3.5	1.0	0.308242	0.280093	0.29946	1.63302	32.3147
120	100	0.9	1.1	0.3	3.5	0.5	0.308242	0.298829	0.300273	1.07351	35.9623