Information Systems and Economics (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/infse.2025.060213
ISSN 2523-6407 Vol. 6 Num. 2

Research on Logistics Warehouse Site Selection Problem Based on Improved Particle Swarm Optimization Algorithm

Kan Liu[a], Hailan Pan[b,*], Yanjiang Li[c]

School of Economics and Management, Shanghai Polytechnic University, Shanghai, 201209, China
[a]2021057204@qq.com, [b]panhailan@sspu.edu.cn, [c]20231530011@sspu.edu.cn
[*]Corresponding author

Keywords: Warehouse Siting; Improved Particle Swarm Algorithm; Total Cost Minimization; Carbon Emission Minimization; Genetic Algorithm

Abstract: In recent years, the logistics industry continues to grow, distribution center as a key node in the logistics system, in reducing logistics and distribution costs, improving the efficiency of service response has a very important strategic value. Therefore, the reasonable planning of logistics warehouse site selection is to establish a mathematical model that considers the minimum total cost and the minimum carbon emission, and use an improved particle swarm algorithm to make the decision. This algorithm incorporates genetic algorithms based on traditional particle swarm algorithms, and adds chaotic mappings to the solution process to obtain a more uniform initial distribution and random perturbations, thereby enhancing global search capability. The experimental results show that the improved particle swarm algorithm provides a basis for helping enterprises to reduce logistics costs, carbon emissions in various links and to improve the efficiency of logistics management. The experimental results show that the improved particle swarm algorithm provides a basis for helping enterprises to reduce logistics costs, carbon emissions in various links and environmental issues management by selecting the optimal warehouse nodes.

1. Introduction

With the booming development of modern logistics industry and e-commerce industry, the market demand for all kinds of commodities shows a continuous growth trend. As a key node in the logistics system, logistics warehouses play a pivotal role in the flow of goods. In the traditional distribution network mode, due to the lack of logistics warehouses for the coordination of scheduling, often derives from the distribution of high cost, inventory backlog, environmental pollution and other issues, the increase in these hidden costs will undoubtedly reduce the operating efficiency of enterprises. Scientific and reasonable warehouse location can not only significantly reduce the distribution costs, but also reduce carbon dioxide emissions in the transportation process. As a classic problem in operations research, the warehouse siting problem refers to determining the number and location of logistics nodes in the logistics system to rationally plan the logistics network structure under the premise of determining the object of siting, the target area of siting, the cost function and the constraints, and with the goal of minimizing the total logistics cost or the optimal total service level or maximizing the social benefits.

In recent years, many scholars have researched the problem of site selection of logistics warehouses in the present time. Han Hao et al. proposed a model for solving the optimization problem of node siting in logistics network, which transforms the problem of logistics node siting into the problem of finding the optimal solution for the minimum value of a function[1]. Liu Shanqiu, Fan Bingpeng take the minimum cost as the objective and use a genetic algorithm to solve the problem of express logistics distribution center siting; Huang Kaiming et al. intended to solve the multilevel facility siting-path planning problem of logistics network, established a mathematical model of mixed-integer planning, and proposed a dual-intelligent algorithm integrating the solution scheme of the synergy of the quantum evolutionary algorithm and genetic algorithm; Zhou Yuyang, Zhang Huizhen used the improved immuno-optimization algorithm to establish a logistics warehouse siting model with capacity constraints; Lv Weibin, He Lili used the improved cuckoo algorithm to effectively reduce the cost of logistics and distribution, and improve the number of inventory turnover[2-5]. The above method provides ideas for the algorithm improvement in this paper, but with the greenhouse effect constantly highlighted, green logistics requirements must accelerate the green and low-carbon transformation of logistics in the context of the study of the warehouse siting problem based on the comprehensive consideration of the carbon emissions of the problem of the study is less.

In view of this, this paper introduces the improved PSO, that is, the chaotic projection initialization and genetic algorithms, to avoid premature convergence of the algorithm, enhance the global optimization ability, while combining the concept of low-carbon environmental protection will be the monitoring of carbon dioxide emissions into the indicator system to establish a dual-objective model[6]. Finally, the example simulation shows that the improved particle swarm algorithm in this paper has feasibility and applicability, and provides a reference basis for enterprises to reduce logistics costs and carbon dioxide emissions.

2. Problem Description and Modeling

2.1 Problem Description

The logistics warehouse siting problem studied in this paper considers the capacity constraints of the facilities, the total logistics cost and carbon dioxide emissions, and the model established needs to screen out the optimal layout scheme from the set of candidate facilities to provide cargo transportation services for each demand point. The core constraint is to ensure that each demand point is served and the total demand carried by a single facility point must not exceed its service capacity limit. In the siting problem, there is a contradiction between minimizing total cost and minimizing CO_2 emissions. Building fewer warehouses saves more logistics costs, but increases the frequency of transportation, thus increasing carbon emissions and vehicle costs, and vice versa. Therefore, constructing an appropriate warehouse location model can help to balance the logistics cost and carbon dioxide emissions[7].

2.2 Model Parameterization

I : denotes the set of manufacturing sites, $i \in I$;
J : denotes the set of potential warehouse location points, $j \in J$;
K : the set of customer demand points, $k \in K$;
d_{ij}: the distance from the manufacturing department i to the warehouse location j;

d_{jk}: The distance from the warehouse site location j to the customer demand point k;
q_{ij} : the volume of goods from the manufacturing department to the warehouse j;
q_{jk} : the volume of goods demanded from the warehouse to the customer's point of demand k;
c_{ij} : the cost of transportation per unit from the manufacturing department i to the warehouse location j;
c_{jk}: the cost of transportation per unit from the warehouse location j to the customer's point of demand k;
m_i : annual warehousing costs at the manufacturing department i;
h_j : annual warehousing costs at the warehouse site j;
f_j : the fixed cost of building a warehouse at the warehouse site j;
R : the average return rate of goods;
$S(d_{jk})$: Customer satisfaction function;
S_{min}: Satisfaction threshold;
P: customer satisfaction penalty factor;
e_t: Carbon dioxide emission factor;
x_{ij}If goods are transported from manufacturing to warehouse then$x_{ij}=1$, otherwise $x_{ij}=0$
y_{jk}: If goods are transported from warehouse to point of demand then $y_{jk}=1$, else $y_{jk}=0$
w_i: If Manufacturing i is selected then $w_i=1$, else $w_i=0$
u_j: $u_j=1$ if warehouse j is selected, else $u_j=0$
v_k: If demand point k is selected then $v_k=1$, else $v_k=0$

2.3 Objective Function

In this paper, there are two particle swarm optimization objectives, one is to minimize the total cost of logistics, and the other is to minimize the carbon dioxide emissions, with these two objectives as the conditions for optimization, and then select the optimal warehouse address from the alternative overseas warehouses to establish the warehouse location model takes into account the cost of transportation at the facility, warehousing costs, construction costs, the cost of returns and exchanges, and the cost of customer satisfaction penalties as the total cost of logistics content. Therefore, the logistics distribution center location problem is described as the following objective function:

$$Z_1=\sum_{i=1}^{I}\sum_{j=1}^{J} c_{ij}\, d_{ij} x_{ij} q_{ij}+\sum_{j=1}^{J}\sum_{k=1}^{K} c_{jk}\, d_{jk} y_{jk} q_{jk}+\sum_{j=1}^{J} f_j u_j+\sum_{i=1}^{I} w_i m_i+\sum_{j=1}^{J} u_j h_j+\sum_{j=1}^{J}\sum_{k=1}^{K} R u_j y_{jk} c_{jk} d_{jk} q_{jk}+\sum_{j=1}^{J} max\left(0, S_{min}-S(d_{jk})\right)\cdot q_{jk}\cdot P \quad (1)$$

$$Z_2=\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij} d_{ij} e_t+\sum_{j=1}^{J}\sum_{k=1}^{K} y_{jk}\, d_{jk} e_t \quad (2)$$

Customer satisfaction $S(d_{jk})$ is a segmented function based on the distanced_{jk} :

$$S(d_{jk})=\begin{cases} 1, & d_{jk}\leq 50 \\ 1-\dfrac{d_{jk}-50}{200-50}, & 50< d_{jk} <200 \\ 0, & d_{jk}\geq 200 \end{cases} \quad (3)$$

If the customer satisfaction $S(d_{jk})$ is lower than the thresholdS_{min} , the penalty cost will be incurred:

$$\sum_{j=1}^{J} max\big(0\ ,S_{min}\text{-}S(d_{jk})\big)\ \cdot q_{jk}\cdot P \tag{4}$$

Meanwhile, the single-objective planning problem is easy to find the optimal solution, while the multi-objective planning problem is very complex and difficult to find the optimal solution. Therefore, in this paper, we refer to the method of BRONFMAN et al. who transformed the bi-objective problem into a single-objective problem and then solved it in the model solution, and transformed the bi-objective problem of this paper into a single-objective problem as well:.

$$Z=min\left(\omega_1\frac{Z_{1max}\text{-}Z_1}{Z_{1max}\text{-}Z_{1min}}+\omega_2\frac{Z_{2max}\text{-}Z_2}{Z_{2max}\text{-}Z_{2min}}\right) \tag{5}$$

s.t. Eqs. (3)~(14)

$$\omega_1+\omega_2=1 \tag{6}$$

Where Z_{1max} and Z_{1min} denote the maximum and minimum values of the objective functionZ_1 respectively;Z_{2max} andZ_{2min} denote the maximum and minimum values of the objective functionZ_2 respectively;ω_1 andω_2 are the weighting factors.

Constraints:

$$\sum_{j=1}^{J} y_{jk}=1 \qquad \forall k\in K \tag{7}$$

Each customer demand point k must be and can only be served by a unique repository j. This constraint ensures that all demand points are covered, avoiding unmet demand and ensuring the integrity of the logistics network.

$$x_{ij}\leq u_j \qquad \forall i\in I, \forall j\in J \tag{8}$$

$$y_{ij}\leq u_j \qquad \forall j\in J, \forall k\in K \tag{9}$$

If Manufacturing i transports goods to Warehouse j (x_{ij}=1), then Warehouse j must have been selected (u_j=1);

If warehouse j delivers goods to demand point k (y_{jk}=1), warehouse j must have been selected (u_j=1). This constraint ensures that only selected warehouses can participate in the transit and distribution of goods, avoiding the need for unbuilt warehouses to take on logistics functions.

$$x_{ij}\leq w_i \qquad \forall i\in I, \forall j\in J \tag{10}$$

If the manufacturing department i transports goods to the warehouse j (x_{ij}=1), the manufacturing department i must be selected (w_j=1). This constraint ensures that the transportation activity is initiated only by the operational manufacturing department.

$$\sum_{i=1}^{I} x_{ij}q_{ij}=\sum_{k=1}^{K} y_{jk}q_{jk} \qquad \forall j\in J \tag{11}$$

The amount of goods from Manufacturing to Warehouse must be equal to the amount of goods from Warehouse to Demand Point. This constraint ensures that the inventory in the warehouse is balanced, avoiding backlogs or shortages of goods and ensuring the rationalization of the logistics chain.

$$x_{ij},y_{jk},w_i,u_j \in\{0,1\} \quad \forall i\in I,\ \forall j\in J,\ \forall k\in K \tag{12}$$

The decision variables are all binary variables that explicitly represent discrete decisions such as "whether to choose" and "whether to transport", which is in line with the combinatorial optimization characteristics of the warehouse location problem.

$$d_{ij},\ d_{jk},c_{ij},\ c_{jk},\ q_{ij},\ q_{jk} \geq 0 \tag{13}$$

The parameters such as distance, cost and cargo volume are all non-negative values, which are in line with the physical meaning and economic logic of the actual logistics scenario.

2.4 Model Solution Method

2.4.1 Basic Particle Swarm Algorithm

Particle swarm algorithm, firstly proposed by Kennedy and Eberhart, is a kind of heuristic algorithm and random search algorithm. The initial population starts from a random solution, and then iterates and searches for the optimal solution through group collaboration. The particle swarm algorithm originated from the simulation of the foraging behaviour of bird flocks, using the information sharing mechanism between individuals in the flock, so that the foraging behaviour. changes from disordered to orderly, and gradually obtain the optimal foraging route.
The updating equations of particle velocity and position are as follows:

$$v_{ij}(t+1)= v_{ij}(t)+c_1r_1(t)(\mathrm{p}_{ij}(t)x_{ij}(t))+c_2r_2(t)((t)x(t)) \tag{14}$$

$$x_{ij}(t+1)=x_{ij}(t)+v_{ij}(t+1) \tag{15}$$

Velocity indicates the direction and distance of particle movement in the next iteration, and position is a solution to the problem being solved. Particle swarm algorithms have the advantages of fast convergence speed, few parameters, and simple and easy-to-implement algorithms. However, they also have the problem of getting stuck in local optima, so scholars often combine them with other algorithms to solve problems.

2.4.2 Improved Particle Swarm Algorithm Flow

In order to avoid the basic particle swarm algorithm to fall into the local optimal situation too early in the process of iteration, which leads to the bias of the warehouse siting decision, this paper will introduce the chaotic projection mechanism to increase the randomness and uniformity of the particle distribution in the initialization stage of the population. At the same time, the selection, crossover and mutation operations of the genetic algorithm are integrated into the particle iteration process so that the screening of the individual optimum is guaranteed to a certain extent, and the operations are as follows:

Initialization: Generate the initial population through Tent mapping, set the particle velocity, individual optimum and global optimum:

$$x_{n+1}=\begin{cases}2x_n, & 0\leq x_n<0.5\\ 2(1-x_n), & 0.5\leq x_n<1\end{cases} \tag{16}$$

This operation maps chaotic sequences into candidate solutions for warehouse location: if the chaotic valuex_j>0.5, then the first j demand point is selected as the warehouse candidate location, and finally the number of warehouses is ensured to meet the constraints through random addition or deletion.

Iterative optimization:

a: For each particle, update the velocity and position according to the PSO formula;

b: Perform chaotic perturbation to enhance the local search capability;

c: Every 10 generations, a tournament selection strategy was used to randomly select k individuals from the population and select the best adapted individuals into the next generation. A two-point crossover is used for binary-coded particles (1 means selected as a depot and 0 means not selected). Finally, the particle genes are mutated with probability p_m=0.1;

d: Update the individual optimum and the global optimum;

Termination condition: after reaching the maximum number of iterations, output the global optimal solution, i.e., the optimal warehouse siting scheme.

3. Algorithm Analysis

According to the above mathematical model and algorithm analysis, an example is set to evaluate the feasibility and performance of the improved particle swarm algorithm in solving the logistics and distribution center[8]. Due to market expansion and the need to save logistics costs, in order to meet the balance of cost and environmental conditions, B logistics company needs to build new logistics warehouses to provide product distribution services for 30 customer demand points, of which the new logistics warehouses are selected from the above 30 demand points, and four new warehouses are proposed to be built.

Based on the research assumptions of this paper, the actual demand of company B is quantified and substituted into the model framework set in this paper for solving. This experimental case only studied the particle swarm algorithm to solve the logistics warehouse in the 30 demand points between the setup program. Manufacturing Department to the logistics warehouse, logistics warehouse to the demand point of each unit of transportation costs for 1 yuan, logistics warehouse construction costs for 500,000 yuan, manufacturing department and logistics warehouse annual warehousing costs for 20,000 yuan, the average return rate of 2%, the remaining part of the parameter settings as Table 1:

Table 1 Parameter settings

Population size	30
Maximum number of iterations	100
Inertia weight	0.8
Carbon dioxide emission factor	0.24
Satisfaction threshold	0.7
Penalty factor	100
Crossover probability	0.8
Mutation probability	0.1
Genetic operation interval	10

3.1 Data Preparation

The serial numbers, coordinates, and requirements of the demand points involved in this paper are shown in Table 2, and the coordinates of the location of the production and manufacturing department are shown in Table 3:

Table 2 Demand Point Coordinates and Demand Quantity

Serial number	Latitude and longitude coordinates (East longitude, North latitude)	Demand quantity (tons)
1	(120.21 , 30.25)	42
2	(117.86 , 31.77)	33
3	(121.46 , 31.25)	42
4	(121.42 , 28.66)	34
5	(119.42 , 32.19)	55
6	(119.98 , 32.53)	72
7	(116.48 , 33.93)	62
8	(120.31 , 31.49)	66
9	(118.43 , 31.35)	33
10	(120.59 , 31.30)	75
11	(119.97 , 31.81)	36
12	(117.09 , 31.89)	23
13	(119.33 , 34.77)	56
14	(117.99 , 29.22)	68
15	(118.24 , 29.56)	34
16	(120.89 , 32.87)	12
17	(121.47 , 31.23)	34
18	(118.78 , 32.07)	54
19	(120.15 , 30.28)	64
20	(118.02 , 31.32)	65
21	(118.55 , 29.88)	24
22	(119.16 , 34.59)	35
23	(117.70 , 28.50)	54
24	(117.18 , 34.26)	86
25	(122.20 , 30.00)	23
26	(120.86 , 32.01)	47
27	(120.09 , 30.89)	68
28	(120.15 , 33.35)	67
29	(119.64 , 29.12)	12
30	(121.40 , 31.62)	83

Table 3 Coordinates of Production and Manufacturing Department

Serial number	Latitude and longitude coordinates (East longitude, North latitude)
1	(118.74 , 32.00)
2	(120.09 , 30.13)
3	(117.13 , 28.98)
4	(117.21 , 31.85)

3.2 Operation Results

Table 4 Logistics distribution center site selection scheme

Warehouse ID	Longitude	Latitude
Warehouse 12	117.09	31.89
Warehouse 18	118.78	32.07
Warehouse 19	120.15	30.28
Warehouse 23	117.70	28.50

Substituting the parameter settings and related data described above into the improved particle swarm algorithm, the algorithm is implemented through Python code, and the final logistics distribution center site selection scheme is shown in Table 4. Finally, four logistics distribution centers are generated to provide logistics distribution services for 30 urban demand points, respectively:[12, 18, 19, 23], and its total logistics cost composition is shown in Table 5:

Table 5 Total logistics cost composition

Cost Type	Amount (Yuan)	Proportion
Construction Cost	2000000.00	83.6%
Warehousing Cost	160000.00	6.7%
Transportation Cost	163671.92	6.8%
Return and Exchange Cost	3273.44	0.1%
Satisfaction Penalty Cost	65848.57	2.8%
Total Cost	2392793.93	/
Carbon Dioxide Emissions(kg)	39281.26	/

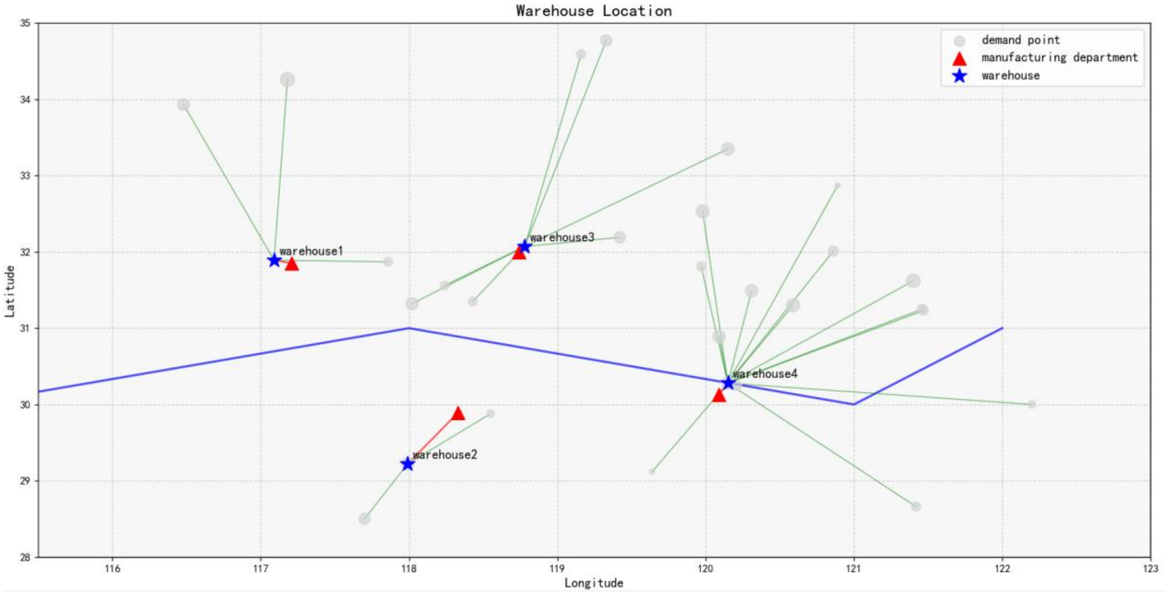

Fig. 1 Warehouse Location Selection Results

At the same time, in order to show the three-level logistics network layout of manufacturing department - warehouse - demand point more clearly, this paper provides a visualization image of logistics warehouse location as shown in Figure 1. It can be seen that the red triangles are the locations of four manufacturing departments, each serving a warehouse represented by a blue pentagram under the optimization algorithm, and finally the warehouse follows the dual objectives of total logistics costs and carbon emissions to distribute goods to 30 demand points.

3.3 Result Analysis

An analysis of logistics total cost components reveals that warehouse construction and storage costs are the primary cost drivers in logistics distribution center location decisions. However, their inherent necessity and relatively transparent pricing make them the hardest to reduce. Thus, cost control should focus on transport (e.g., vehicles, routes), product returns/exchanges, and customer satisfaction. Enterprises need to prioritize green, low-carbon transport, optimize routes, and deliver high-quality, satisfying services to enhance logistics cost control and competitiveness.

At the same time, in order to verify that the improved particle swarm algorithm has certain performance and advantages in solving the logistics distribution center site selection scheme, it is compared with the traditional particle swarm algorithm for the same instance, and the results are shown in Table 6:

Table 6 Comparison between improved particle swarm algorithm and particle swarm algorithm

Name: Improved Particle Swarm Algorithm	Improved particle swarm algorithm	Particle swarm algorithm
Total cost	2392793.93 Yuan	2432428.03 Yuan
Carbon Emission	39281.26 kg	49140.74 kg

From the above comparisons, it can be seen that compared with the traditional particle swarm algorithm, the improved particle swarm algorithm, which mixes the chaotic projection and genetic algorithm, not only converges to the optimal solution in the two objectives of the total cost and carbon emission, but also has some advantages in the running time, which can converge to the optimal solution faster and improves the speed of the solution. For the final total cost, the improved particle swarm algorithm saves $39,634.1 compared with the traditional particle swarm algorithm, with a saving ratio of 1.6%, and for the total carbon emission, the improved particle swarm algorithm reduces 9,859.48 kg compared with the traditional particle swarm algorithm, with a reduction ratio of 20.06%, which can be seen that the improved particle swarm algorithm is an effective algorithm for solving the problem of site selection of logistics and distribution centers and has a certain reliability and applicability.

As shown in Figs. 2 and 3, the hybrid particle swarm algorithm with chaotic projection and genetic algorithm converges at about the 50th iteration of the algorithm. Compared with the basic particle swarm algorithm which does not achieve iterative optimization during the whole algorithm, the improved algorithm retains the fast convergence characteristics of PSO, and at the same time, it enhances global traversal through the chaotic mechanism, and enhances the diversity of the populations through the genetic operation, so that the results of the warehouse siting obtained are more reliable and feasible. The results are more reliable and feasible.

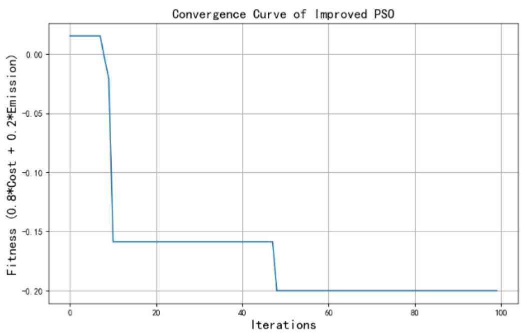

Fig. 2 Iteration Result of Improved PSO

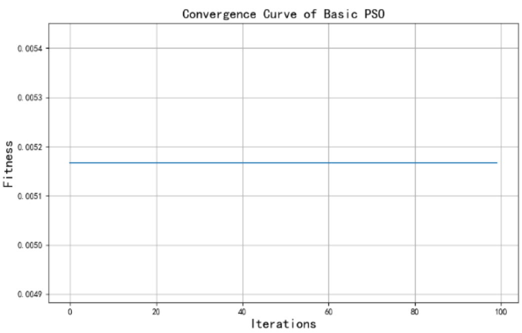

Fig. 3 Iteration Result of Basic PSO

4. Conclusion

This paper studies and analyzes the enterprise logistics warehouse siting problem, and puts forward an improved particle swarm algorithm, based on the traditional particle swarm algorithm in solving the problem is easy to precocious, easy to fall into the local optimum of the status quo, through the chaotic projection initialization, fusion of genetic algorithms, and other methods, to effectively improve the efficiency of the problem of logistics warehouse siting, to save the cost of logistics, and reduce the time of the solving process. Moreover, in the emphasis on enterprise ESG activities, many studies only focus on the minimization of logistics costs and ignore the environmental impact of the warehouse siting process, this paper incorporates carbon dioxide emissions into the objective function, and through the algorithm effectively reduces the carbon dioxide emissions, focusing on the combination of economic benefits and environmental benefits.

References

[1] Han Hao, Wang Suling. Modeling and Solution of Multilevel Logistics Node Siting Problem[J]. Journal of Shanghai Maritime University, 2009, 30(04): 30-35.

[2] Liu Shanqiu, FAN Bingpeng. Site selection of express logistics distribution center based on genetic algorithm[J]. Journal of Hunan University of Technology, 2021, 35(05): 70-76.

[3] Huang Kaiming, LU Caiwu, LIAN Minjie. Modeling and Algorithmic Research on Three-Level Facility Siting-Path Planning Problem[J]. Systems Engineering Theory and Practice, 2018, 38(03): 743-754.

[4] Zhou Yuyang, ZHANG Huizhen. Research on logistics warehouse siting problem based on improved immune optimization algorithm[J]. Software Guide, 2021, 20(03): 34-42.

[5] Lv Weibin, He Lili, Zheng Junhong. Study on the location of logistics distribution center based on improved cuckoo algorithm[J]. Logistics Engineering and Management, 2024, 46(04): 5-9.

[6] Cai Biaobing. Site Selection of Overseas Warehouses for Cross-border E-commerce in China[D]. Jiangxi University of Finance and Economics, 2023.

[7] Zhang Kuangda. Optimization of Overseas Warehouse Location Considering Green Logistics Performance[D]. Jiangsu University, 2022.

[8] Ren Xiaoling, ZHAO Juanjuan, REN Jiaoli. Optimization of logistics and distribution paths by hybrid adaptive cuckoo algorithm[J]. Computer Simulation, 2024, 41(05): 168-71+241.