J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 5049--5067

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Research on green logistics network planning strategy combining machine learning and carbon emission constraints

Yanru Li[1,✉]

[1] *International Business College, Chengdu Polytechnic, Chengdu, Sichuan, 610041, China*

ABSTRACT

The planning of green logistics networks has gradually become the focus of attention in both academic and business circles, as it has been increasingly emphasized on environmental protection. This study aims to explore how to combine machine learning and carbon emission constraints to construct a more efficient and environmentally friendly green logistics network planning strategy. A machine learning-based logistics demand forecasting model is constructed by Support Vector Regression (SVR) machine, and the model parameters are optimized using genetic algorithm to improve the model accuracy. Analyze the sources of carbon emissions in the logistics network and establish a carbon emission calculation model. Construct a green logistics network planning model considering carbon emission constraints, and analyze the feasibility of the model through practical examples. The method of this paper can effectively measure the carbon emissions in the transportation and storage phases of the logistics network. Under the condition of considering carbon emission constraints, positioning the upper limit of carbon emission below 270,000 can realize a stable balance of economic and environmental benefits.

Keywords: machine learning; SVR; genetic algorithm; carbon emission constraint; green logistics network

1. Introduction

As global climate change becomes increasingly severe, controlling carbon emissions has become a common challenge for the international community. The "dual-carbon" goal aims to promote the country's transition to low-carbon development [1]. Urban logistics, as an important part of urban infrastructure, has a particularly prominent carbon emission problem [2]. The traditional logistics system is no longer adapted to the current needs of environmental protection due to its high energy

✉ Corresponding author.
E-mail address: Zhuanyong1218@126.com (Y. Li).
Received 20 March 2024; Revised 05 June 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-285

consumption and high emission characteristics [3]. Therefore, optimizing the logistics and distribution network and developing green logistics have become the necessary path to achieve sustainable urban development [4].

The green and low-carbon development of urban logistics and distribution is an important initiative aimed at optimizing the urban logistics system and reducing the negative impact on the environment, which not only helps to improve the transportation efficiency, but also helps to reduce carbon emissions and traffic congestion, mainly through the planning of efficient distribution networks, optimization of logistics paths, the promotion of renewable energy sources, and the introduction of intelligent technology to enhance the sustainability of distribution [5-7]. Urban green logistics and distribution not only helps to realize the goal of “double carbon”, but also provides urban residents with more efficient and environmentally friendly logistics services, and promotes the development of urban transportation, the environment, and the economy in a sustainable direction [8-9].

The greening and low-carbon development of urban logistics and distribution is an important initiative aimed at optimizing the urban logistics system and reducing the adverse impact on the environment, which not only helps to improve transportation efficiency, but also helps to reduce carbon emissions and traffic congestion. Firstly, the concept of green logistics is introduced, and the literature [10] compares and discusses the concepts and forms of traditional logistics and green logistics, and points out that the logistics process has an impact on the environment mainly in the form of greenhouse gases, pollution, noise, vibration and packaging waste. A large number of related studies in this field have been reviewed, and the main theme of the research is to explore and test the path of green logistics. Literature [11] uses time-discrete transport centralized network flow planning modeling in resource-sharing state-space-time (SST) networks to solve large eco-package transportation, and solves small eco-package pickup and delivery based on cost-minimizing synchronous guided positional routing model, and confirms the effectiveness of the proposed strategy in simulation experiments. Literature [12] describes the positive environmental contribution of freight trams compared to fuel heavy trucks and proposes a logistics model for planning small shipments by taking logistics uncertainty and risk into account. Literature [13] attempts to achieve the design and planning of green forward and reverse logistics networks based on a mixed integer linear programming model, which is characterized by the consideration of multiple objectives such as minimizing operating costs, process costs, transportation costs and fixed costs. Literature [14] introduced a two-stage stochastic planning model in biodiesel green supply chain architecture and tested it with real cases to confirm the feasibility of the proposed model. Literature [15] conceptualizes a location routing strategy that takes into account the charging location of logistics and transportation trams and vehicle paths, which can be used to solve problems related to uncertain customer patterns, and introduces a (parallelized) adaptive large neighborhood search method to cope with large instances. Literature [16] systematically reviews the related research on green logistics, builds a green logistics structure based on MCDM technology, and looks forward to the potential future development space of this structure. Literature [17] proposed the use of sustainable-oriented innovation methods to explore the green path of logistics 5.0, pointing out that artificial intelligence can be used to optimize logistics routes, and reduce logistics packaging waste as well as improve warehouse layout. Literature [18] combines interviews with rooted theory to clarify that the factors affecting the effectiveness of GLP include the degree of perfection of the GLP system, the government's green governance capability, the GLP awareness of logistics enterprises, the government's green governance capability and the degree of logistics development, which makes a positive contribution to the improvement of logistics performance and the sustainable development of logistics. Literature [19] aims to reduce logistics carbon emissions and conceptualizes a network and route planning framework for forward/reverse logistics applicable to fresh produce logistics, and the total logistics cost as well as the transportation process cost minimization are considered in the first and second phases of the analytical framework, respectively.

In this paper, a logistics demand forecasting model is constructed using the support vector regression mechanism under the optimization of genetic algorithm, and the model forecasting accuracy is examined by using the mean absolute error and the mean square error ratio. By analyzing the main sources of carbon emissions in the logistics network, the carbon emission calculation model is established by selecting the transportation link and the storage link, and the carbon emissions in the logistics network are measured. Based on the results of logistics demand prediction and carbon emission calculation, the planning and design of green logistics network under the constraint of carbon emission is carried out based on the uncertainty of fuzzy mathematics. Experiments are designed to demonstrate the feasibility of this paper's logistics forecasting model and green logistics network planning model in practical application, and in the process, the carbon emission cap that can realize the balance between economic and ecological benefits of logistics network is explored.

2. Machine learning-based logistics demand forecasting

The prediction of logistics demand can intuitively reveal the dynamics of logistics and transportation, thus providing data support for the calculation of carbon emissions in the logistics and transportation process. In this section, a logistics demand prediction model is constructed by using a support vector regression mechanism based on machine learning method, and the model is optimized by using genetic algorithm, which in turn improves the model prediction performance.

2.1 Support vector regression machine

2.1.1. Overview of the methodology. When the support vector machine (SVM) is applied to regression fitting analysis, the basic idea is no longer to find an optimal classification surface to make the two classes of samples separate, but to find an optimal classification surface to minimize the error of all training samples from that optimal classification surface. The support vector regression machine (SVR) model was established by introducing a γ-insensitive loss function [20].

Assume that the training set sample pair containing l training samples is $\{(x_i, y_i), i=1,2,\cdots,l\}$, where $x_i\left(x_i \in R^d\right)$ is the input column vector of the i th training sample, $x_i=\left[x_i^1, x_i^2, \cdots x_i^d\right]^T$, $y_i \in R$ are the corresponding output values. where $\phi(x)$ is the nonlinear mapping function.

Let the linear regression function in the high latitude feature space be established as:

$$f(x)=w\Phi(x)+b \tag{1}$$

Define the linear γ-insensitive loss function:

$$L\left(f(x), y, \gamma\right)=\begin{cases}0, |y-f(x)| \leq \gamma \\ |y-f(x)|-\gamma, |y-f(x)|>\gamma\end{cases} \tag{2}$$

where $f(x)$ is the predicted value returned by the regression function. y is the corresponding true value, if the difference between $f(x)$ and y is less than or equal to γ, the loss is equal to 0. If the $f(x)$ and y. The difference is greater than or equal to γ, the loss is equal to $|y-f(x)|-\gamma$.

Introduce the slack variables ξ_i, ξ_i^* and describe the above problem of finding w, b in mathematical terms, i.e.:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\left(\xi_i + \xi_i^*\right)$$
$$s.t.\begin{cases} y_i - w\phi(x) - b \le \gamma + \xi_i \\ -y_i + w\phi(x) + b \le \gamma + \xi_i^* \\ \xi_i \ge 0, \xi_i^* \ge 0, i = 1,2,\cdots l \end{cases} \tag{3}$$

where C is the penalty factor, C larger means larger penalty for samples with training error greater than γ, γ specifies the error requirement of the regression function, γ smaller means smaller error of the regression function.

In order to solve (3), the same Largrange function is introduced and converted to dyadic form:

$$\max\left[-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\left(a_i - a_i^*\right)\left(a_j - a_j^*\right)K\left(x_i, x_j\right) - \sum_{i=1}^{l}\left(a_i + a_i^*\right)\gamma + \sum_{i=1}^{l}\left(a_i - a_i^*\right)y_i\right]$$
$$s.t.\begin{cases} \sum_{i=1}^{l}\left(a_i - a_i^*\right) = 0 \\ 0 \le a_i \le C \\ 0 \le a_i^* \le C \end{cases} \tag{4}$$

where $K\left(x_i, x_j\right) = \Phi\left(x_i\right)\Phi\left(x_j\right)$ is the kernel function.

Let the optimal solution obtained by solving (4) be $a = \left[a_1, a_2, \cdots, a_l\right]$, $a* = \left[a_1^*, a_2^*, \cdots, a_l\right]$, then we have:

$$w^* = \sum_{i=1}^{l}\left(a_i - a_i^*\right)\Phi\left(x_i\right) \tag{5}$$

$$b^* = \frac{1}{N_{nsv}}\left\{\sum_{i=1}^{l}\left[y_i - \sum_{x_i \in w}\left(a_i - a_i^*\right)K\left(x_{i,} x_j\right) - \gamma\right] + \sum_{0<a_i<C}^{l}\left[y_i - \sum_{x_i \in sw}\left(a_i - a_i^*\right)K\left(x_i x_j\right) + \gamma\right]\right\} \tag{6}$$

Where N_{nsv} is the support. The number of vectors. Thus, the regression function is:

$$f(x) = w^*\Phi(x) + b^* = \sum_{i=1}^{t}\left(a_i - a_i^*\right)\Phi\left(x_i\right)\Phi(x) + b^* = \sum_{i=1}^{l}\left(a_i - a_i^*\right)K\left(x_i, x_j\right) + b^* \tag{7}$$

where only some of the parameters $\left(a_i - a_i^*\right)$ are non-zero and their corresponding samples x_i are the support vectors in the problem.

2.1.2. Kernel function selection. $K\left(x_i, x_j\right)$ as a kernel function, the following kernel functions are commonly used, linear kernel function $K\left(x_i, x_j\right) = kx_i x_j$. polynomial kernel function $K\left(x_{i,} x_j\right) = \left(kx_i x_j + s\right)^d$, d represents the rank of the polynomial kernel. Radial quantity RBF kernel function $K\left(x_i,x_j\right)=\exp\left(-\frac{\left\|x_i-x\right\|^2}{2\sigma^2}\right)$, σ represents the frequency width of the radial quantity kernel. The sigmoid kernel function $K\left(x_i,x_j\right)=\tanh\left(kx_ix_j+\theta\right)$ with parameters κ and θ. In this paper, the radial volume RBF is finally selected as the kernel function of the model through comparative experiments.

2.2 Predictive modeling

2.2.1. Genetic Algorithm Optimization. In general, when choosing parameters, most researchers follow the trial-and-error method, which first builds an SVR model based on different parameters and then obtains the optimized parameters on the basis of the test set, but this processing is both time-consuming and requires luck.

Compared with the trial-and-error method of SVR parameters mentioned above, this paper uses genetic algorithm [21] to optimize the kernel function parameters of support vector regression machine, and this method can optimize the parameters of SVR at the same time, and construct the genetic algorithm-support vector machine regression (GA-SVR) prediction model. Genetic algorithm is a stochastic search algorithm that draws on natural selection and natural genetic mechanisms in biology, and is suitable for dealing with complex and nonlinear planning problems.

2.2.2. Construction process. Logistics demand forecasting model construction process.
1) Carry out data selection and data preprocessing.
2) Optimize the parameters of genetic algorithm, train linear kernel function, polynomial kernel function, radial volume kernel function, sig-moid kernel function support vector machine, and select the optimal parameters to establish support vector regression model.
3) The accuracy of the support vector regression model was checked by MAPE and MSE indexes, and the optimal model was selected for data prediction.

2.2.3. Model Accuracy Test. The accuracy of the model was tested by two indicators: the mean absolute error, and the ratio of mean squared deviations:

$$MAPE=\frac{\sum_{i=1}^{l}\left|a_i-p_i\right|/a_i}{l}100\% \tag{8}$$

$$MSE=\sqrt{\frac{\sum_{i=1}^{l}\left[\varepsilon_i-\bar{\varepsilon}\right]^2}{\sum_{i=1}^{l}\left(a_i-\bar{a}_i\right)^2}} \tag{9}$$

where a_i denotes . True value, $\bar{a}_i$ is true. The average of the values, p_i indicates the prediction. Value, $\bar{\varepsilon}$ is the average of the residual series ε_i. MSE and MAPE are often used to measure the deviation of the true value from the predicted value, and the smaller the MSE and MAPE, the closer the predicted value is to the true value.

3. Carbon emission calculation models

3.1 Calculation of carbon emissions from transportation

The carbon emissions of a vehicle are directly proportional to its fuel consumption, so carbon emissions can be determined by estimating the fuel consumption of a vehicle. The fuel consumption of a vehicle is related to many factors, such as the speed of the vehicle, the load of the vehicle, the distance traveled by the vehicle, the type of vehicle, and the road conditions. There are many factors that affect fuel consumption, and the model for measuring fuel consumption can be built based on the relationship between different factors and fuel consumption.

3.1.1. Standard driving conditions. First establish a mathematical expression for the carbon emissions of a vehicle in the standard state. The standard driving state of the vehicle is that the vehicle has zero load and is traveling in a road with zero gradient. When the vehicle is in the standard driving condition, the rate of carbon emission per unit of the vehicle is shown in equation (10):

$$\varepsilon^{empty}(v) = K + av + bv^2 + cv^3 + d/v + e/v^2 + f/v^3 \tag{10}$$

Of these:
is the unit emission rate of a vehicle traveling at a 0% grade with no load, in units of g/km.
K is a constant.
$a-f$ is a coefficient.
v is the average speed at which the vehicle is traveling, in units of km/h.

3.1.2. Slope correction factors. For each vehicle class, slope and CO2 emission rate, the slope correction factor can be calculated as a polynomial function of the average vehicle speed:

$$GC(v) = A_6 v_6 + A_5 v^5 + A_4 v^4 + A_3 v^3 + A_2 v^2 + A_1 v + A_0 \tag{11}$$

where:
$GC(v)$ denotes the slope correction factor.
$A_0, A_1, \ldots, A_6$ is the coefficient.
v is the average speed at which the vehicle is traveling in km/h.

3.1.3. Load correction factor. The driving resistance of a vehicle is influenced by the vehicle mass, i.e. a higher vehicle mass requires a higher power from the engine during driving, especially in acceleration mode. It is well known that emissions and fuel consumption are proportional to the engine power and therefore the calculation must in principle take into account the vehicle load. Also in real distribution, it is essential to take the load factor into account in the model since the vehicle is fully loaded at the beginning of the distribution and empty on the return trip.

In the case of heavy duty vehicles, vehicle loads have a significant impact on emissions and fuel consumption because loads can significantly affect the total weight of the vehicle. The load correction factor is shown in equation (12):

$$LC(\gamma, v) = k + n\gamma + p\gamma^2 + q\gamma^3 + rv + sv^2 + tv^3 + \frac{u}{v} \tag{12}$$

where:
$LC(\gamma, v)$ denotes the load correction factor.
γ denotes the road slope.
k denotes the constant.
$n-u$ denotes the coefficient.

Thus the rate of carbon dioxide emission from the vehicle under full load is:

$$\varepsilon^{full}(v) = \varepsilon^{empty}(v) \cdot LC(\gamma, v) \tag{13}$$

where:
$\varepsilon^{full}(v)$ is the unit emission rate of carbon dioxide when the vehicle is traveling at 0% grade with no load, unit g/km.

$\varepsilon^{empty}(v)$ is the CO2 unit emission rate when the vehicle is traveling at 0% grade in unloaded condition, unit g/km.

$LC(\gamma,v)$ 5 represents the load correction factor.

In the actual distribution of transportation links, the vehicle is usually not empty or fully loaded two extreme states, more often in the non-empty state. The carbon emission in the unloaded state can be regarded as linearly related to the load of the vehicle. Therefore, the carbon dioxide emission rate of the vehicle in the non-full load state is obtained as:

$$\varepsilon_{ij}\left(X_{ij},W_1,v_1\right)=\left(\frac{\varepsilon_{ij}^{full}\left(v_1\right)-\varepsilon_{ij}^{empty}\left(v_1\right)}{W_1}\right)X_{ij}+\varepsilon_{ij}^{empty}\left(v_1\right) \tag{14}$$

$$\varepsilon_{jk}\left(X_{jk},W_2,v_2\right)=\left(\frac{\varepsilon_{jk}^{full}\left(v_2\right)-\varepsilon_{jk}^{empty}\left(v_2\right)}{W_2}\right)X_{jk}+\varepsilon_{jk}^{empty}\left(v_2\right) \tag{15}$$

$\varepsilon_{ij}\left(X_{ij},W_1,v_1\right)$ is the rate of carbon dioxide emissions in the transportation chain of the vehicle from the supplier to the distribution center, in units of g/km.

$\varepsilon_{jk}\left(X_{jk},W_2,v_2\right)$ is the rate of carbon dioxide emissions in the transportation chain of the vehicle from the distribution center to the point of demand, in units of g/km.

W_1 is the maximum load for a vehicle traveling from the supplier to the distribution center, in t.

W_2 is the maximum load of the vehicle from the distribution center to the point of demand in t.

X_{ij} is the distribution volume from the supplier to the distribution center, in t.

X_{jk} is the distribution volume from the distribution center to the point of demand, in t.

v_1, v_2 is the average speed of the respective vehicle operation, in km/h.

Then, at that time, is the rate of carbon dioxide emissions in the fully loaded state.

3.2 *Calculation of carbon emissions from warehousing*

3.2.1. Lighting consumption. Lighting is an essential part of any kind of warehouse, and lighting accounts for a relatively high proportion of energy consumption, so when considering the carbon emissions of the warehouse chain, the carbon emissions of lighting activities need to be measured. Firstly, the energy consumption required for lighting activities in warehouses is calculated, and the energy consumption of lighting activities is measured according to the method of calculating the power demand per unit area.

Generally speaking, the unit power area demand of the outdoor area is proportional to the lighting intensity, if the outdoor area area is S_1 and the required lighting intensity is x_1, then the power consumption E_1 of the lighting in time T_1 can be roughly expressed by the following equation:

$$E_1=S_1\cdot x_1\cdot\frac{10}{500}\cdot T_1=\frac{x_1S_1T_1}{50} \tag{16}$$

However, for warehouses, the energy consumption caused by lighting is not only related to the strength of the lighting tools, but also affected by the layout of the warehouse, such as the width of the corridor and the vertical height of the warehouse. The power demand per unit area is analyzed by multivariate first regression analysis, and the expression of power demand per unit area in the warehouse room can be obtained as follows:

$$x_2=0.333y_1+0.4y_2+0.049y_3 \tag{17}$$

Where, x_2 is the power demand per unit area in the warehouse interior, y_1 denotes the corridor width, y_2 denotes the vertical height, y_3 denotes the lighting intensity demand, and let the power consumption of a warehouse of area S_2 be in time T_2, then:

$$E_2 = x_2 \cdot S_2 \cdot T_2 = \left(0.333y_1 + 0.4y_2 + 0.049y_3\right) \cdot S_2 \cdot T_2 \tag{18}$$

The total energy consumption of the warehouse for lighting is:

$$E_a = E_1 + E_2 = \frac{x_1 S_1 T_1}{50} + \left(0.333y_1 + 0.4y_2 + 0.049y_3\right) \cdot S_2 \cdot T_2 \tag{19}$$

3.2.2. Thermostatic consumption. In the energy consumption of the constant temperature link, first of all, based on the nature of the characteristics of the stored goods, set the target temperature to be maintained for t_0, but also need to know the real-time temperature of the warehouse in the constant temperature link for t, the area of the warehouse for S, the energy consumption per unit area of the warehouse to change the temperature of 1 ℃ for m_l, where l represents the type of fuel, refrigeration or heating process of the efficiency of the energy consumption for the ρ_l, then the energy consumption to maintain a constant temperature of the warehouse function is The energy consumption function for keeping the warehouse at a constant temperature is:

$$E_b = \sum_l m_l S \left| t - t_0 \right| \rho_l \tag{20}$$

3.2.3. Consumption for loading and unloading. In warehouse operations, it is difficult to avoid the need to transfer the position of goods or stacking and other activities. For automated three-dimensional warehouse, its main body consists of shelves, alley stacking cranes, in (out) warehouse workstations and automatic transportation (out) and operation control system. Alleyway machines travel through the aisles between shelves to complete the work of storage and retrieval of goods. However, for warehouses without automation equipment, heavy forklifts or front reach electric stackers are used to stack goods through standard pallets.

For automated equipment, the main energy consumption is related to its operating time:

$$E_c = Pt \tag{21}$$

where denotes the energy consumption in the warehouse generated by the loading and unloading handling equipment, P denotes the power of the automated equipment, and t denotes the operating time of the automated equipment.

3.2.4. Carbon emissions from warehousing. Through the above measurement of energy consumption in the storage link, the carbon dioxide emission model can be used to measure carbon emissions in the storage link. At the same time, with reference to the reference coefficient of various energy sources discounting standard coal and the carbon emission coefficients of major energy sources in the China Energy Statistical Yearbook, the carbon emission is calculated by the formula:

$$f_b = \sum_{l=1}^{n} E_l \cdot \tau_l \cdot e_l l = 1, 2, \ldots, n \tag{22}$$

Of which.

f_b is the total carbon emissions from storage.

l represents the type of fossil fuel.

E_l represents the consumption of the first energy source.

τ_l, represents the standard coal conversion factor of the energy source in the th.

e_l represents the carbon emission coefficient of the energy source of the th.

4. Green logistics network planning model considering carbon emission constraints

4.1 Problem Description and Assumptions

By analyzing and exploring the multi-level green logistics network of raw material suppliers [22]. It is understood that the structure of this network is shown in Fig. 1, which mainly contains three levels: raw material suppliers, distribution centers and factories. In this process, a number of suppliers are selected from candidate supplier $\{A_1, A_2, A_3, \cdots\}$ and a number of nodes are selected from alternative node $\{B_1, B_2, B_3, \cdots\}$ of the distribution center for the construction of the corresponding facilities, so as to ultimately achieve the goals of minimum logistics cost and minimum carbon dioxide emission.

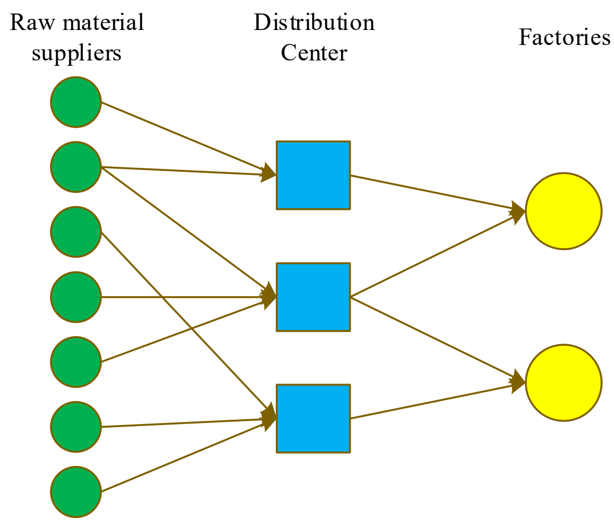


Fig. 1. Logistics network structure

4.2 Model parameters and decision variables

According to the description of the problem, based on the uncertainty of fuzzy mathematics to carry out the design and planning of the green logistics network, in the process of design and planning, should choose to use the scientific and advanced, and has a better logistics capacity of logistics facilities, the use of scientific green logistics technology for transportation. The utility of the green logistics network should be fully realized to contribute to the protection of the environment and the improvement of economic efficiency, so that the whole logistics network meets the needs of social development.

4.2.1. Collections. Hypothesis $J = 1, \cdots,$ is the location of the distribution center. k is the location area of the factory, Assumption $k = 1, \cdots, K$; m is the raw material supplier supply capability level, Assumption $m = 1, \cdots, M$. n is the inventory management capability level of the distribution center, Assumption $n = 1, \cdots, N$; Z is the management technology used by the supplier, Assumption $z = 1, \cdots, Z$. p is the transportation mode used, Assumption $p = 1, \cdots, P$.

4.2.2. Parameters. Suppose d_k is the quantity of raw materials required by the factory k to use them, f_m is the fixed cost of a supplier i with energy level m. g_j^{n1} is the fixed cost of the distribution center, and Φ_n^j refers to the energy level of n that is built at the j location, and the inventory management technique used is l. c^p refers to the unit transportation cost per unit of time, and the mode of transportation used is p. p_j^1 is the unit inventory cost at the distribution center j and the inventory management technique used is l. T_{ij}^p refers to the transportation time, which is the time used to transport the product j from the distribution center to the plant k if the mode of transportation p is used. $\overline{\tau}_i^m$ refers to supply capacity; $\overline{s}_{ij}^p$ refers to CO2 emissions, i.e., the amount of CO2 emissions that would be generated if transportation p methods were used.

4.2.3. Decision-making variables. u_{ij}^p refers to the amount of raw materials. v_{jk}^p refers to the amount of raw materials that have been processed using inventory management techniques l if transportation mode p is used. $x_i^m \in \{0,1\}$, $x_i^m = 1$ refers to a supplier with a capacity level of m, otherwise $x_i^m = 0$, $y_j^{nl} \in \{0,1\}$, $y_j^{nl} = 1$ refers to a distribution center with a capacity level of n and using inventory management techniques l in J place, otherwise $y_j^{nl} = 0$.

4.2.4. Objective function. The objective function of the model is:

$$\min W = \omega_1 W_1 + \omega_2 W_2 \tag{23}$$

Among them:

$$W_1 = \sum_i \sum_m \tilde{f}_i^m x_i^m + \sum_j \sum_n \sum_l \tilde{g}_j^{nl} y_j^{nl} + \sum_i \sum_j \sum_p c^p \tilde{T}_{ij}^p u_{ij}^p + \sum_j \sum_k \sum_p \sum_l \left(\overline{p}_j^l + c^p \tilde{T}_{jk}^p \right) v_{jk}^{lp} \tag{24}$$

$$W_2 = \sum_i \sum_j \sum_p \overline{s}_{ij}^p u_{ij}^p + \sum_j \sum_k \sum_p \sum_l \overline{s}_{ij}^p v_{jk}^{lp} \tag{25}$$

4.2.5. Constraints. A perfect green logistics network model is constructed by calculating the following constraint equations:

$$\sum_j \sum_l \sum_p v_{jk}^{lp} \geq \tilde{d}_k, \forall k \tag{26}$$

$$\sum_i \sum_p u_{ij}^p = \sum_j \sum_l \sum_p v_{jk}^{lp}, \forall j \tag{27}$$

$$\sum_i \sum_p u_{ij}^p \leq \sum_i \sum_p x_i^m \overline{\tau}_i^m, \forall i \tag{28}$$

$$\sum_i \sum_p v_{jk}^{lp} \leq \sum_n y_j^{nl} \overline{\phi}_j^n, \forall j, l \tag{29}$$

$$\sum_m x_i^m \leq 1, \forall i; \sum_n \sum_l y_j^{nl} \leq 1, \forall j \tag{30}$$

$$x_i^m, y_j^{nl} \in \{0,1\}, \forall i, j, m, n, l \tag{31}$$

$$u_{ij}^p, y_{jk}^{lp} \geq 0, \forall i, j, k, p, l \tag{32}$$

5. Green logistics network planning experiment and analysis

Based on the series of methods proposed in the previous section, the validity of these methods is verified in this section. Firstly, the logistics demand in the study area is predicted by selecting logistics data. Second, on the basis of logistics demand prediction, the carbon emission in the logistics network is measured using the carbon emission calculation method proposed in this paper. Finally, the green logistics network planning model of this paper is solved according to the carbon emission constraints.

5.1 Logistics demand forecast

5.1.1. Data sources. The weekly consolidated freight volume of logistics companies from 2022 to 2023 by a city statistics bureau is selected as the quantitative index of logistics demand. A total of 200 data are organized and obtained and simulation experiments are carried out, the first 140 data are used as the training set, and the last 60 data are used as the test set of the model, and the original logistics demand samples are shown in Fig. 2, from which it can be seen that the logistics demand of the place shows a significant upward trend from 2022 to 2023, so the construction of a green logistics network that takes into account the constraints of carbon emission has an economic, social and ecological important value.

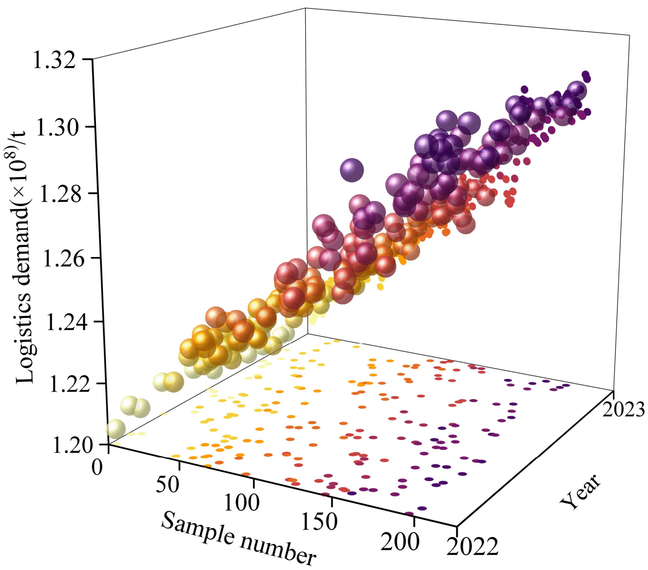


Fig. 2. Sample logistics requirements

5.1.2. Chaos analysis and phase space reconstruction of samples. Chaos analysis of the logistics demand data reveals that its maximum Lyapunov (Lyapunov) index is 1.473, which indicates that the data has a certain degree of chaotic nature, and it is necessary to find out the changing characteristics of it through phase space reconstruction. The phase space reconstruction theory can understand the geometrical properties of the phase space of nonlinear dynamical systems from a certain component in the measured time series, and recover chaotic attractors in the high-dimensional phase space. Phase space reconstruction is performed for a one-dimensional chaotic time series $x(i), i = 1, 2, \cdots, n$, and the reconstructed m -dimensional state vector can be expressed as:

$$X(i) = (x(i), x(i+\tau), \cdots, x(i+(m-1)\tau))^T, i = 1, 2, \cdots, n \tag{33}$$

where $N = n-(m-1)\tau$ is the number of phase points, m is the dimension of the embedding, and τ is the delay time.

In this paper, the mutual information method and correlation dimension are used to determine the optimal time delay $\tau = 8$ and embedding dimension $m = 8$, as shown in Fig. 3. The logistics demand data in Fig. 2 is reconstructed using $\tau = 8, m = 8$ to form a learning sample of logistics demand data for the support vector regression machine.

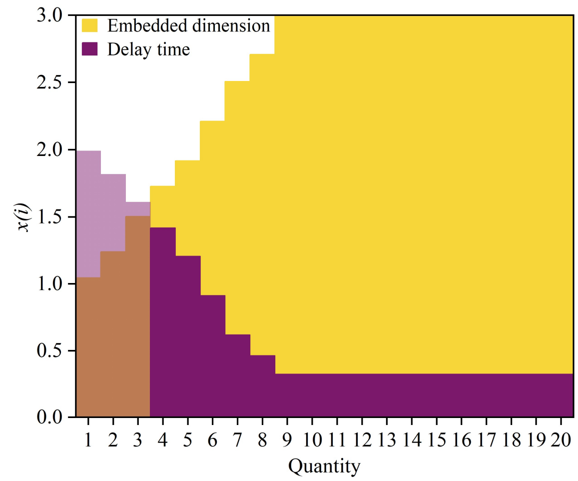
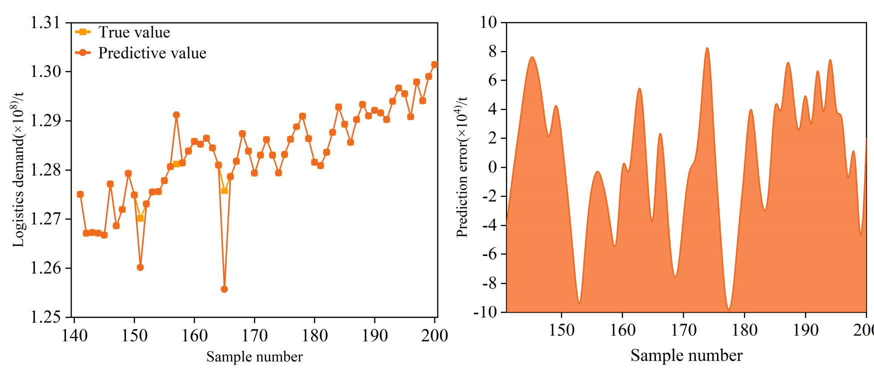


Fig. 3. Chaotic analysis of logistics demand data

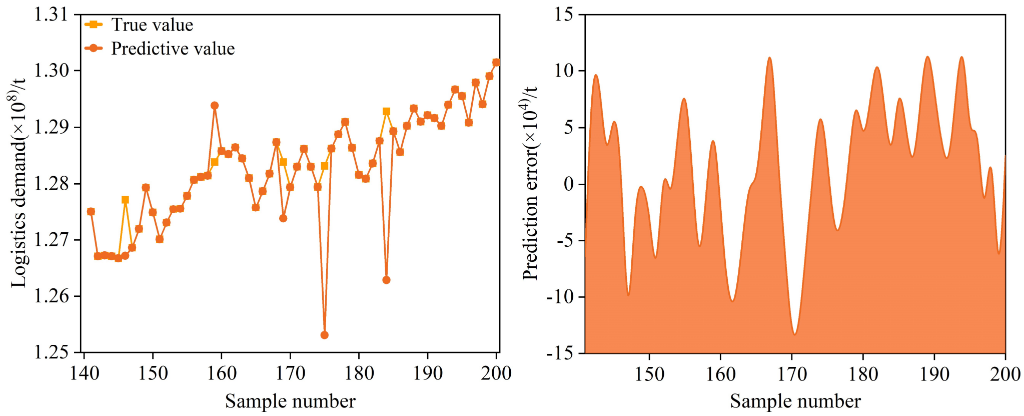
5.1.3. Analysis of forecast results. After optimizing the parameter (C,σ) of the support vector regression machine using genetic algorithm, the optimal parameter for one-step prediction of logistics demand is obtained as $C=834.72$, $\sigma=0.933$. Learning samples of logistics demand data are inputted into the logistics demand prediction model established by the support vector regression machine using the optimal parameter, and the one-step prediction results are obtained as shown in Fig. 4, and the fitting comparison between the real value and the predicted value as well as the deviation between the real and the predicted value are shown in (a) and (b), respectively. From the figure, it can be seen that the one-step prediction value of this paper's model is quite consistent with the real value of logistics demand data, and the deviation between the two is small, indicating that this paper's model can accurately describe the stochastic and time-varying nature of the logistics demand, and it is a feasible logistics demand prediction model with reliable results.



(a) Predictive result (b) Predictive error

Fig. 4. Prediction results of logistics requirements

Due to the complexity and variability of the logistics market, one-step prediction results do not have much practical reference value, for this reason, this paper carries out the logistics demand prediction in three steps in advance, using genetic algorithm to optimize the parameter (C,σ) of the support vector regression machine, to get the optimal parameter of one-step prediction is $C=16732, \sigma=1.247$. The results of the 3-step prediction of logistics demand are shown in Fig. 5, with (a) and (b) indicating the fitting comparison between the true value and the predicted value and the deviation between the true value and the predicted value, respectively. From the figure, it can be seen that the prediction accuracy of the model decreases slightly with the increase of the prediction step, but among the 60 samples in the test set, 55 samples are still predicted correctly, and the prediction accuracy remains above 90%. It shows that the model in this paper is still able to effectively fit the trend of logistics demand in the case of 3-step prediction, and the prediction results can provide data support for the calculation of carbon emissions and the planning of green logistics network.



(a) Predictive result (b) Predictive error

Fig. 5. Prediction results of logistics requirements

5.1.4. Comparative performance analysis. The current classic logistics demand forecasting models: logistic regression (LR), convolutional neural network (CNN), and time series method (ARIMA) are adopted, and we choose to use the mean square error (MSE) and the mean absolute error (MAPE) as evaluation indexes. The experimental results of comparing this paper's model (GA-SVR) with other logistics demand forecasting models are shown in Table 1. As can be seen from the table, the average MSE value of this paper's GA-SVR model is 3.99 and the average MAPE value is 3.16% in 20 experiments, which show obvious advantages compared with the other three models. The experimental results show that the performance of this paper's model in the practical application of logistics demand forecasting is significantly better than the other models, and can provide more accurate and reliable data support for the study.

Table 1. Model performance comparison results

Experimental frequency	LR		CNN		ARIMA		GA-SVR	
	MSE	MAPE/%	MSE	MAPE/%	MSE	MAPE/%	MSE	MAPE/%
1	18.37	11.05	14.99	8.97	10.19	8.92	4.49	3.53
2	19.34	11.22	15.20	8.94	12.82	6.03	4.24	3.80
3	19.81	10.92	15.50	9.65	11.33	7.09	4.95	3.57
4	18.98	10.91	15.65	8.35	11.56	7.56	4.15	3.01
5	18.48	11.21	15.49	9.87	12.04	8.17	4.83	3.22
6	19.94	11.16	14.53	9.52	12.08	6.56	3.28	3.51
7	19.29	10.56	15.63	9.13	11.45	8.89	3.35	3.02
8	18.15	11.15	15.35	8.51	10.13	6.97	3.51	3.49
9	19.33	11.61	14.32	9.10	10.69	8.98	4.83	3.81
10	18.04	10.70	15.58	8.51	12.33	8.06	3.30	3.12
11	18.70	10.94	14.44	8.72	12.43	6.97	3.90	3.34
12	18.85	11.56	14.01	8.57	10.82	8.96	3.97	3.40
13	18.64	10.52	15.98	8.24	10.31	6.34	3.53	3.09
14	19.64	10.88	15.12	9.99	12.31	7.15	4.52	3.12
15	18.84	10.25	14.16	8.64	10.39	7.74	3.11	3.23
16	18.79	11.27	15.52	9.38	10.20	7.31	4.08	3.14
17	18.16	10.83	14.64	8.62	11.12	6.77	3.41	3.09
18	18.31	11.96	15.46	8.07	10.25	7.79	3.55	3.04
19	18.11	11.77	15.90	9.11	11.02	6.37	4.27	3.16
20	19.16	11.16	14.59	8.02	12.09	7.41	4.60	3.30
Mean	18.85	11.08	15.10	8.90	11.28	7.50	3.99	3.30

5.2 Example Analysis of Green Logistics Network Planning

5.2.1. Description of the algorithm. A single-product inbound logistics network for a manufacturing company is known to contain 10 candidate suppliers and 6 candidate distribution centers to meet the demand for raw materials in 3 factories. The demand for raw materials from the factories in the logistics network, the logistics costs (including supplier selection costs, distribution center construction costs, transportation costs, etc.), and the environmental indicators (i.e., CO2 emissions) are all directly or indirectly given in the form of triangular fuzzy numbers. For example, let the factory set be $\{C_1, C_2, C_3\}$, according to the constructed logistics demand forecasting model to get its raw material logistics demand is $d_1 = (452, 478, 496), d_2 = (969, 974, 986), d_3 = (792, 813, 827)$ respectively.

Let the supplier selection set of the enterprise be $\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$, and other main parameters are shown in Table 2 and Table 3. Let the enterprise's distribution center selection set for $\{B_1, B_2, B_3, B_4, B_5, B_6\}$, each distribution center has 2 inventory management technology and inventory management capabilities. Since B_4 is a completed distribution center, its construction fixed costs compared to other candidates for distribution center farmland. In the transportation process between each logistics node, there are 2 different modes of transportation, and their carbon dioxide emissions in the transportation and warehousing phases are different from each other. According to the carbon emission measurement method proposed in this paper, the carbon emissions of the 6 distribution centers under the 2 modes of transportation are shown in Table 4.

Table 2. Candidate suppliers select cost and supply capability data

	m	A_1	$A2$
f_i^m	1	(492,508,519)	(710,725,748)
	2	(511,517,531)	(742,751.774)
τ_i^m	1	(772,843,869)	(973,982,1012)
	2	(842,886,903)	(1026,1033,1125)
	m	A_3	A_4
f_i^m	1	(441,462,495)	(482,519,537)
	2	(514,526,548)	(544,549,556)
τ_i^m	1	(771,829,897)	(701,713,716)
	2	(884,953,991)	(782,793,817)
	m	A_5	A_6
f_i^m	1	(433,445,448)	(502,528,533)
	2	(442,457,463)	(531,542,556)
τ_i^m	1	(559,568,582)	(778,801,816)
	2	(571,584,603)	(882,905,931)
	m	A_7	A_8
f_i^m	1	(493,519,531)	(600,616,627)
	2	(522,537,549)	(619,623,638)
τ_i^m	1	(753,782,810)	(1062,1103,1184)
	2	(774,815,843)	(1109,1113,1124)
	m	A_9	A_{10}
f_i^m	1	(532,537,543)	(468,479,482)
	2	(539,548,556)	(503,518,533)
τ_i^m	1	(671,723,769)	(579,586,593)
	2	(725,787,839)	(591,632,658)

Table 3. Candidate distribution center fixed cost and management capability data

	m	l	B_1	B_2	B_3
g	1	1	(443,446,457)	(641,654,669)	(388,449,482)
		2	(447,451,496)	(643,663,711)	(405,447,499)
	2	1	(651,752,773)	(667,723,735)	(441,482,536)
		2	(682,779,801)	(733,742,759)	(542,558,589)
Φ	1	—	(569,593,617)	(669,704,738)	(523,591,608)
	2	—	(688,702,783)	(843,862,879)	(632,667,695)
	m	l	B_4	B_5	B_6
g	1	1	(0,0,0)	(368,443,457)	(553,568,642)
		2	(124,147,153)	(508,526,557)	(574,598,657)
	2	1	(125,149,162)	(549,554,568)	(644,669,697)
		2	(166,205,227)	(551,559,596)	(656,673,682)
Φ	1	—	(541,568,595)	(523,552,571)	(579,586,598)
	2	—	(762,803,829)	(633,648,656)	(673,682,707)

Table 4. Carbon emission measurement

	Transport mode 1		Transport mode 2	
	Transport stage	Storage stage	Transport stage	Storage stage
B_1	2181.02	1381.16	2462.79	1398.71
B_2	2046.93	1320.68	2415.26	1584.34
B_3	2487.85	1215.42	2305.29	1387.92
B_4	2311.30	1240.61	2309.64	1592.54
B_5	2333.53	1337.21	2409.07	1389.12
B_6	2423.25	1443.94	2406.82	1478.02

5.2.2. Solving the example. In order to verify the feasibility of the green logistics network planning model under carbon emission constraints, LING11.0 software was used to solve the example. At the same time, the numerical experiments are analyzed from different fuzzy feasibility and weights at 2 levels, and the results of the algorithm are shown in Table 5. Among them, α is the minimum feasibility and ω_1 is the first objective function weight. W_1 and W_2 are the network logistics costs and CO2 emissions, respectively. X and Y are the information about the selected suppliers and distribution centers, respectively. $X(a,b)$ indicates that the a th supplier is selected, and its supply capacity is b. $Y(c,d,e)$ indicates that the c th distribution center is selected, and its supply capacity is d, and it adopts the inventory management technique e. It can be seen from the table that, in the process of changing the weights, the total cost increases correspondingly with the decrease of carbon emission, so it can be seen that the total cost and the CO2 emission are opposites of the equivalent objective functions.

Table 5. Experimental results

α	ω_1	W_1	W_2	$X(a,b)$	$Y(c,d,e)$
0.80	0.91	484599.78	338100.58	(2,2)(7,3)(9,4)	(2,2,1)(6,1,1)
	0.67	496619.69	324314.43	(6,3)(8,2)(10,4)	(2,1,2)(4,1,2)(5,2,1)
	0.66	536179.89	321627.27	(1,3)(8,4)(9,2)	(4,2,1)(3,1,2)(5,1,1)
	0.19	560503.09	308873.62	(1,1)(10,1)	(4,1,2)(3,1,2)
	0.03	583576.19	297848.11	(1,1)(10,6)	(4,2,1)(5,2,2)
0.60	0.97	503266.37	335813.48	(2,3)(8,5)(10,2)	(2,1,2)(4,1,1)
	0.82	551573.53	335281.05	(6,6)(7,3)(9,4)	(3,2,2)(4,2,1)(5,2,2)
	0.73	581553.17	328198.46	(1,5)(6,3)(9,6)	(2,1,2)(3,2,1)(6,1,1)
	0.42	582380.55	319729.17	(1,3)(10,4)	(2,2,1)(5,1,2)
	0.25	638764.83	292757.04	(1,2)(8,4)	(3,1,2)(5,2,1)

Based on the analysis of the characteristics of the 2 objective functions, the 1st objective (total cost minimization objective) tends to favor the selection of lower-cost inventory management technology 1 for distribution centers, while the 2nd objective (i.e., carbon emission minimization objective) tends to favor the selection of distribution centers that use greener inventory management technology 2 and a higher level of inventory handling capacity.

By setting different CO2 emission ceilings, the impact on logistics costs and the location of related logistics facilities is explored, and the results are shown in Table 6. It can be seen from the table, but when the carbon dioxide emission cap is greater than 270,000, the logistics cost grows slowly and the change is not obvious. And when the upper limit of carbon emissions is less than 270000, the logistics cost changes rapidly. This shows that enterprises set strict carbon emission standards, will inevitably form the trend of logistics cost multiplication. From the viewpoint of this example, to control the carbon dioxide emissions of the logistics network under 270000, which is the economic and environmental benefits of both in a relatively stable state of decision-making.

Table 6. The logistics network of different carbon emissions caps

CO2 limit	W_1	$X(a,b)$	$Y(c,d,e)$
300000	470224.61	(2,2)(5,3)(8,2)(10,1)	(1,1,2)(3,1,2)(4,2,2)(5,1,2)
290000	479636.37	(2,3)(6,3)(7,3)(9,3)	(1,1,2)(4,1,2)(5,2,2)(6,1,2)
280000	479478.83	(3,4)(5,6)(7,6)(9,6)	(1,2,2)(3,2,2)(5,1,2)(6,1,2)
270000	525923.18	(2,2)(4,2)(7,2)(8,2)	(2,1,2)(3,1,2)(5,1,2)
260000	564778.23	(6,6)(9,6)(10,6)	(3,1,2)(4,2,2)(6,1,2)
250000	617838.31	(1,3)(5,4)(8,3)	(3,1,2)(5,1,2)(6,2,2)
240000	656935.35	(3,5)(5,5,)(8,5)	(2,2,2)(4,1,2)(6,1,2)
230000	745729.33	(5,1)(7,1)(10,1)	(5,1,2)(6,2,2)
220000	798998.45	(9,2)(10,2)	(4,2,2)(5,1,2)
210000	820018.07	(8,1)(10,1)	(3,2,2)(6,2,2)
200000	869637.25	(3,4)(7,4)	(1,1,2)(6,1,2)

6. Conclusion

This paper uses genetic algorithm to optimize support vector regression machine and constructs logistics demand forecasting model. It proposes a calculation model of carbon emissions in the transportation phase and warehousing phase of the logistics network, and constructs a green logistics network planning model under the constraint of carbon emissions.
1) The logistics demand prediction model constructed in this paper shows good prediction accuracy in both one-step and three-step prediction. The accuracy of the three-step prediction is slightly lower than that of the one-step prediction, but the prediction accuracy of more than 90% can be maintained. Meanwhile, the average MSE and average PAPE of this paper's prediction model are 3.99 and 3.16%, respectively, which have significant advantages over other logistics demand prediction models. It shows that the prediction model in this paper can effectively realize the accurate prediction of logistics demand.
2) The carbon emission calculation model in this paper can derive the carbon emissions in the transportation and storage phases under two different transportation modes. And according to the green logistics network planning model established in this paper, it can be concluded that the upper limit of carbon emission can be controlled below 270,000 in order to realize the balance and stability of economic and ecological benefits of logistics network operation.

References

[1] Govindan, K., & Gholizadeh, H. (2021). Robust network design for sustainable-resilient reverse logistics network using big data: A case study of end-of-life vehicles. Transportation Research Part E: Logistics and Transportation Review, 149, 102279.

[2] Trivellas, P., Malindretos, G., & Reklitis, P. (2020). Implications of green logistics management on sustainable business and supply chain performance: evidence from a survey in the greek agri-food sector. Sustainability, 12(24), 10515.

[3] Tan, B. Q., Wang, F., Liu, J., Kang, K., & Costa, F. (2020). A blockchain-based framework for green logistics in supply chains. Sustainability, 12(11), 4656.

[4] Chhabra, D., Garg, S. K., & Singh, R. K. (2017). Analyzing alternatives for green logistics in an Indian

automotive organization: A case study. Journal of cleaner production, 167, 962-969.

[5] Zhang, D., Zhan, Q., Chen, Y., & Li, S. (2018). Joint optimization of logistics infrastructure investments and subsidies in a regional logistics network with CO2 emission reduction targets. Transportation Research Part D: Transport and Environment, 60, 174-190.

[6] Nikseresht, A., Golmohammadi, D., & Zandieh, M. (2024). Sustainable green logistics and remanufacturing: a bibliometric analysis and future research directions. The International Journal of Logistics Management, 35(3), 755-803.

[7] Bal, A., & Satoglu, S. I. (2018). A goal programming model for sustainable reverse logistics operations planning and an application. Journal of cleaner production, 201, 1081-1091.

[8] Yingfei, Y., Mengze, Z., Zeyu, L., Ki-Hyung, B., Avotra, A. A. R. N., & Nawaz, A. (2022). Green logistics performance and infrastructure on service trade and environment-measuring firm's performance and service quality. Journal of King Saud University-Science, 34(1), 101683.

[9] Dev, N. K., Shankar, R., & Swami, S. (2020). Diffusion of green products in industry 4.0: Reverse logistics issues during design of inventory and production planning system. International Journal of Production Economics, 223, 107519.

[10] Blanco, E. E., & Sheffi, Y. (2024). Green logistics. In Sustainable supply chains: a research-based textbook on operations and strategy (pp. 101-141). Cham: Springer International Publishing.

[11] Wang, Y., Peng, S., Zhou, X., Mahmoudi, M., & Zhen, L. (2020). Green logistics location-routing problem with eco-packages. Transportation Research Part E: Logistics and Transportation Review, 143, 102118.

[12] Shramenko, N., Merkisz-Guranowska, A., Kiciński, M., & Shramenko, V. (2022). Model of operational planning of freight transportation by tram as part of a green logistics system. Archives of Transport, 63(3), 113-122.

[13] Zarbakhshnia, N., Soleimani, H., Goh, M., & Razavi, S. S. (2019). A novel multi-objective model for green forward and reverse logistics network design. Journal of cleaner production, 208, 1304-1316.

[14] Ghelichi, Z., Saidi-Mehrabad, M., & Pishvaee, M. S. (2018). A stochastic programming approach toward optimal design and planning of an integrated green biodiesel supply chain network under uncertainty: A case study. Energy, 156, 661-687.

[15] Schiffer, M., & Walther, G. (2018). Strategic planning of electric logistics fleet networks: A robust location-routing approach. Omega, 80, 31-42.

[16] Tian, G., Lu, W., Zhang, X., Zhan, M., Dulebenets, M. A., Aleksandrov, A., ... & Ivanov, M. (2023). A survey of multi-criteria decision-making techniques for green logistics and low-carbon transportation systems. Environmental Science and Pollution Research, 30(20), 57279-57301.

[17] Nicoletti, B., & Appolloni, A. (2024). Green Logistics 5.0: a review of sustainability-oriented innovation with foundation models in logistics. European Journal of Innovation Management, 27(9), 542-561.

[18] Zhang, W., Zhang, M., Zhang, W., Zhou, Q., & Zhang, X. (2020). What influences the effectiveness of green logistics policies? A grounded theory analysis. Science of the Total Environment, 714, 136731.

[19] Guo, J., Wang, X., Fan, S., & Gen, M. (2017). Forward and reverse logistics network and route planning under the environment of low-carbon emissions: A case study of Shanghai fresh food E-commerce enterprises. Computers & Industrial Engineering, 106, 351-360.

[20] Y Halimi,Z Bouzid,S Halimi & N Ghellai. (2025). Exploring the impact of feature selection on load profile analysis and energy demand forecasting through enhanced support vector regression. Physica Scripta(1),016015-016015.

[21] Qing An,Xia Zhang,Lang Rao & Mengyan Zhang. (2025). Novel decomposed genetic algorithm for equivalent circuit model parameter optimization of lithium-ion battery. Journal of Energy Storage115116-115116.

[22] Hongzhi Wang,Haojie Ran & Shuhua Zhang. (2024). Location-routing optimization problem of country-township-village three-level green logistics network considering fuel-electric mixed fleets under carbon emission regulation. Computers & Industrial Engineering110343-110343.