International Journal of New Developments in Engineering and Society
ISSN 2522-3488 Vol. 7, Issue 9: 57-64, DOI: 10.25236/IJNDES.2023.070910

Vegetable Replenishment and Pricing Strategy Based on the White Shark Optimization Algorithm

Weiming Li, Haobin Wu, Tao Kang, Kaican Huang, Huimin Ke, Bin Feng*

Guangzhou Maritime University, Guangzhou, 510725, China
**Corresponding author: fengbin@gzmtu.edu.cn*

Abstract: *With the increasing consumer demand in society, the variety of vegetables in fresh supermarkets has grown, leading to more complex vegetable pricing and replenishment strategies. To address these issues, this paper first employs the Pearson correlation coefficient to analyze the correlation between the sales volumes of different vegetable categories and individual items in the fresh supermarket. It also explores the temporal distribution patterns of vegetable sales volumes. Furthermore, the paper investigates the relationship between total sales volume and cost-based pricing. Based on this, using daily replenishment total quantity as the decision variable and the relationship between total sales volume and cost-based pricing as constraint conditions, and aiming to maximize supermarket revenue, it constructs a BiLSTM model based on the White Shark Optimization Algorithm. This model is used to formulate daily replenishment total quantities and pricing strategies for various categories in the coming week to maximize supermarket revenue.*

Keywords: *Vegetable Replenishment Strategy, Vegetable Pricing Strategy, Correlation Analysis, White Shark Optimization Algorithm, BiLSTM Model*

1. Introduction

As society's demand for consumer goods continues to rise, fresh supermarkets have gradually expanded their vegetable product offerings to meet changing consumer needs. This market trend has made pricing and replenishment strategies for vegetables increasingly complex, forcing supermarket managers to adopt more refined inventory management and pricing tools to better adapt to market demand. In past research, academics have conducted in-depth studies on the correlation between vegetable sales and pricing, and have proposed a variety of approaches to optimize supermarkets' inventory management and pricing strategies.

Prior academic studies have shown that there are complex correlations between vegetable sales volumes, which depend on multiple factors, including seasonal variations, market trends, and consumer preferences. For example, in past research, Haixia Deng[1], in her study of vegetable distribution patterns, clearly indicated that there is usually some correlation between sales of different vegetable categories and that vegetable sales have a temporal distribution pattern.These studies provide the basis of this paper for an in-depth discussion of vegetable sales and inventory management in fresh supermarkets.

However, there is a lack of in-depth research on the relationship between total supermarket sales volume and cost plus pricing. Understanding this relationship is essential for developing more accurate inventory management and pricing models. Therefore, this paper aims to fill this academic gap by analyzing the correlation between sales volumes by applying the Pearson correlation coefficient, investigating the temporal distribution pattern of vegetable sales, and exploring in depth the relationship between total sales volume and cost plus pricing in order to construct a comprehensive inventory management and pricing optimization model. The ultimate goal is to provide fresh supermarkets with more refined replenishment and pricing strategies to maximize their revenues.

2. The Temporal Distribution Patterns of Various Vegetable Categories and Their Correlations

2.1 Data source

This paper's data is sourced from https://cumcm.cnki.net. This paper has collected data from 6 vegetable categories, namely, Leafy Greens, Cauliflower, Aquatic Stems and Roots, Nightshades, Chili, and Edible Fungi.

2.2 The temporal distribution patterns of various vegetable categories

Due to variations in harvest seasons for different vegetables throughout the year, this paper will use monthly sales volume data to examine the distribution and changes in sales volumes of different vegetable categories.

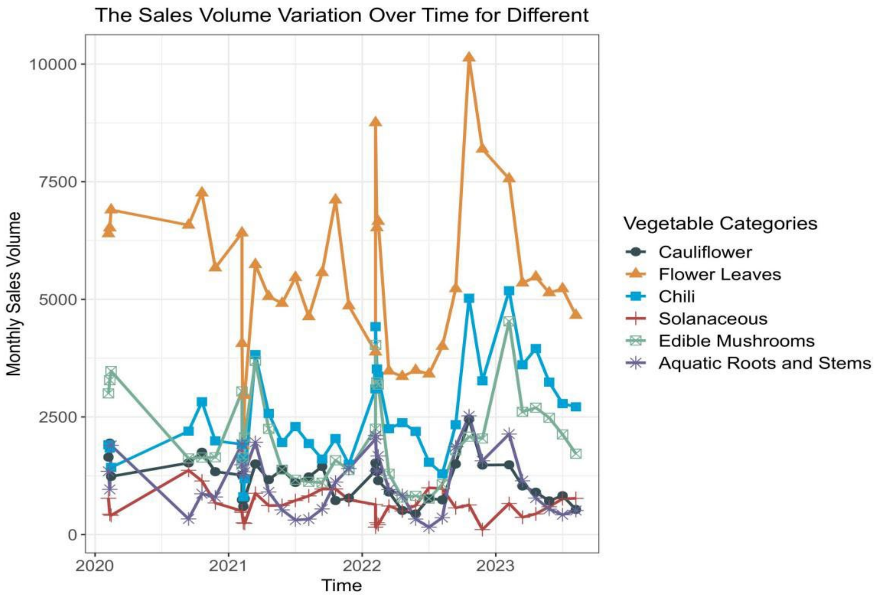

Figure 1: The Temporal Distribution of Vegetable Categories from 2020 to 2023

Analyzing the time distribution patterns of vegetable categories depicted in Figure 1, the following conclusions can be drawn:

1) Based on three years of sales data, the line graphs for the six vegetable categories exhibit a recurring trend. Sales volumes tend to rise from July to August each year, followed by a decline from August to December. However, there is an upward trend from December to January, and from January onwards until June of the following year, sales volumes gradually decrease.

2) When considering specific vegetable categories, it's observed that for Flower Leaves, the highest sales volumes are achieved in August, with the second-highest in January. In contrast, Cauliflower, Chili, Solanaceous, Edible Mushrooms, and Aquatic Roots and Stems hit their peak sales volumes in January, with the second-highest occurring in August.

3) Based on the monthly analysis of the line graphs over the three-year period, it can be further inferred that each of these six vegetable categories exhibits distinct seasonal variations, characterized by periods of low demand (off-season) and high demand (peak season). Both January and August represent the peak seasons for these categories, where sales volumes reach their highest levels.

2.3 Correlation Analysis among Vegetable Categories

Based on the monthly sales volume data for vegetable categories from July 2020 to June 2023, a Pearson correlation analysis was conducted to assess the relationships among the six categories.

(1) For calculating the sample covariance between sales data for each pair of categories, you can use the following formula.

(2)

$$cov(X,Y)=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) \tag{1}$$

In this formula, " X " and " Y " represent the sales data for two categories, " n " is the sample size, "

ISSN 2522-3488 Vol. 7, Issue 9: 57-64, DOI: 10.25236/IJNDES.2023.070910

"X_i" and "Y_i" are the "i" observed values in the sales data, "$\overline{X}$" denotes the sample mean, and "$E(X)$" represents the sample standard deviation.

(3) Utilizing Pearson correlation analysis to generate a correlation Heatmap, as depicted in Figure 2.

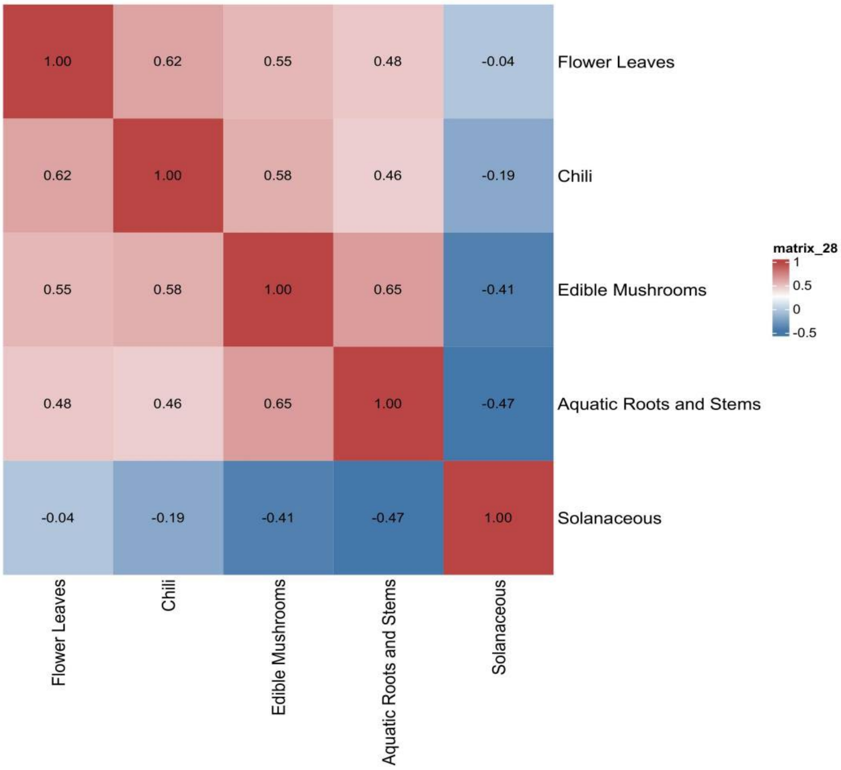

Figure 2: Correlation Heatmap of Sales Volumes among Various Vegetable Categories

In accordance with the correlation analysis depicted in Figure 2, a reasonable interpretation would be that the correlation between each category is not substantial. Therefore, it can be concluded that the sales relationships within each category exhibit weak correlations.

3. Vegetable Replenishment and Pricing Strategy

3.1 Visualization of the Relationship Between Each Vegetable Category and Cost-Plus Pricing

In order to study the relationship between each vegetable category and cost-plus pricing more accurately, this paper aggregates the data on a daily basis from the attached dataset, obtaining the daily total sales volume and daily average cost-plus pricing for each vegetable category from July 1, 2020, to June 30, 2023. Since vegetable pricing typically follows the 'cost-plus pricing' method, this analysis excludes sample data related to discounted sales and focuses only on data pertaining to regular pricing sales.

This paper conducts a visual analysis of the total sales volume and average cost-plus pricing for each vegetable category, represented in a scatter plot as depicted in Figure 3.

ISSN 2522-3488 Vol. 7, Issue 9: 57-64, DOI: 10.25236/IJNDES.2023.070910

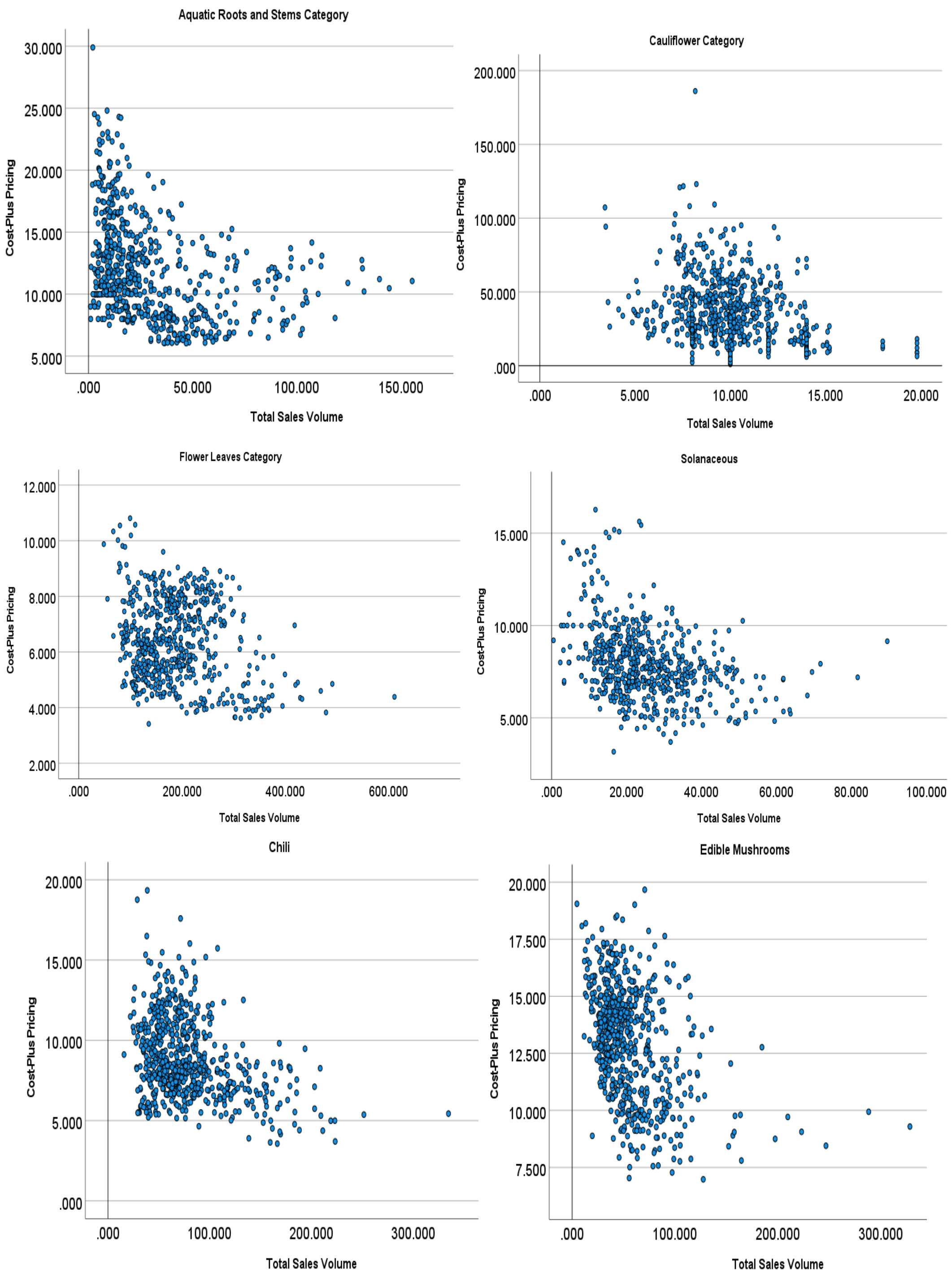

Figure 3: Scatter Plot of Sales Volume and Cost-Plus Pricing Relationship for Various Vegetable Categories

From the scatter plot distribution in Figure 3, it is evident that there are significant differences in the relationship between total sales volume and cost-plus pricing for various vegetable categories. To further analyze the relationships between each vegetable category and cost-plus pricing, individual regression analyses are required for each category.

3.2 Solving the Relationship Between Total Sales Volume and Cost-Plus Pricing for Each Vegetable Category

Using the fitting toolbox to derive the functional relationship expression between sales volume and cost-plus pricing, as presented in Table 1.

Table 1: The relationship between sales volume and cost-plus pricing

Vegetable Category	Function Expression	Fitting Error
Cauliflower Category	$y_1 = -15.667x_1 + 22.122$	0.233
Flower Leaves Category	$y_2 = -0.109x_2^2 + 1.452x_2 + 110.33$	0.269
Aquatic Roots and Stems Category	$y_3 = 0.0045x_3^2 - 0.905x_3 + 65.379$	0.252
Solanaceous	$y_4 = 0.0043x_4^3 + 0.296x_4^2 - 6.181x_4 + 54.7$	0.277
Chili Category	$y_5 = 0.0056x_5^2 - 0.597x_5 + 120$	0.213
Edible Mushrooms	$y_6 = 0.21x_6^2 - 13.12x_6 + 293.6$	0.278

Explanation: y_i represents the total sales volume for various vegetable categories, and x_i represents the cost-plus pricing for these vegetable categories.

3.3 Establishment and Solution of the Optimized Model with BiLSTM Improved by White Shark Optimization (WSO) Algorithm

3.3.1 Establishment of the Model for Maximizing Revenue in the Upcoming Week

(1) Determination of Decision Variables

Monthly replenishment quantity for each category (in kilograms) is represented as " x_{ij} ",Monthly average cost-plus pricing for each category (in yuan per kilogram) is represented as " y_{ij} ", Categories are denoted as " i " representing Cauliflower, Flower Leaves, Chili, Solanaceous, Edible Mushrooms, and Aquatic Roots and Stems, respectively," j " represents the time frame from July 1 to July 7, 2023. $(i = 1,2,3,4,5,6)(j = 1,2,3,4,5,6,7)$

(2)Determination of the Objective Function

In this paper, the replenishment plan is conducted on a per-category basis. The goal is to maximize the total revenue over 7 days in the fresh supermarket, providing the daily replenishment quantity and pricing strategy for each vegetable category for the upcoming week (July 1-7, 2023).

$$x_{ij} = \frac{X_{ij}}{1 - k_{ij}} \tag{2}$$

$$c_{ij} = \frac{C_{ij}}{x_{ij}} \tag{3}$$

$$w_{ij} = \frac{y_{ij} - c_{ij}}{c_{ij}} \tag{4}$$

$$\max M_1 = \sum_{i=1,j=1}^{i=6,j=7} w_{ij} * X_{ij} \tag{5}$$

" X_{ij} " represents the daily replenishment quantity for each vegetable category." k_{ij} " stands for the

ISSN 2522-3488 Vol. 7, Issue 9: 57-64, DOI: 10.25236/IJNDES.2023.070910

average spoilage rate for each vegetable category." c_{ij} " denotes the unit cost for each vegetable category." w_{ij} " represents the interest rate for each vegetable category." M_1 " corresponds to the total revenue of the fresh supermarket over 7 days.

(4) In summary, the model for maximizing revenue over the upcoming week is as follows.

$$\max M_1 = \sum_{i=1,j=1}^{i=6,j=7} w_{ij} * X_{ij} \tag{6}$$

$$s.t.\begin{cases} y_{ij} = \left(1 + w_{ij}\right)c_{ij} \\ W_{ij} = w_{ij} * X_{ij} \\ X_{ij} = x_{ij}\left(1 - k_{ij}\right) \\ c_{ij} = \dfrac{C_{ij}}{x_{ij}} \\ x_{ij}, y_{ij}, w_{ij} \geq 0 \\ i = 1,2,\cdots,6 \\ j = 1,2,\cdots,7 \end{cases} \tag{7}$$

$$\begin{cases} y_{1j} = -15.667x_{1j} + 22.122 \\ y_{2j} = -0.109x_{2j}^2 + 1.452x_{2j} + 110.33 \\ y_{3j} = 0.0045x_{3j}^2 - 0.905x_{3j} + 65.379 \\ y_{4j} = 0.0043x_{4j}^3 + 0.296x_{4j}^2 - 6.181x_{4j} + 54.7 \\ y_{5j} = 0.0056x_{5j}^2 - 0.597x_{5j} + 120 \\ y_{6j} = 0.21x_{6j}^2 - 13.12x_{6j} + 293.6 \end{cases} \tag{8}$$

3.3.2 Establishment of a Time Series Model with BiLSTM Based on White Shark Optimization Algorithm

To obtain the replenishment quantities and pricing strategies for the upcoming 7 days that maximize the supermarket's revenue, this paper has established a BiLSTM time series model to forecast the supermarket's replenishment quantities for the next 7 days. It has also employed the White Shark Optimization Algorithm[2] to enhance the BiLSTM time series model, aligning the replenishment quantities and pricing strategies with the goal of maximizing the supermarket's revenue. This approach ultimately yields the replenishment quantities and pricing strategies for the upcoming 7 days.

3.3.3 BiLSTM Time Series Model Principles

The LSTM[3] network addresses the issue of long-term memory and gradient vanishing in RNNs. Bidirectional Long Short-Term Memory (BiLSTM)[4] is a variant of LSTM, which comprises forward and backward LSTM layers, further optimizing the long-term memory and gradient vanishing issues in RNNs. By understanding the concept of interdependence between past and future data in time series, BiLSTM can simultaneously consider both past and future data information[5], resulting in more

ISSN 2522-3488 Vol. 7, Issue 9: 57-64, DOI: 10.25236/IJNDES.2023.070910

accurate and realistic predictions.[6]

3.3.4 The Model's Performance

Table 2: Training and Testing Set Performance Table

	Training Set	Testing Set
MSE	580.4609	340.6550
RMSE	20.4425	30.4007
MAE	18.6245	26.9811
MAPE	1.8885	2.3141
R^2	0.73766	0.76566

A comprehensive analysis of Table 2 shows that the model exhibits a good fit and the error is within acceptable limits.

3.3.5 Replenishment Quantity and Pricing Strategy for the Next Seven Days

The replenishment and pricing strategies from July 1 to July 7 are obtained as shown in Table 3 and Table 4.

Table 3: Pricing Strategy for the Next 7 Days

Time	Flower Leaves Category	Cauliflower Category	Aquatic Roots and Stems Category	Solanaceous	Chili	Edible Mushrooms
1	8.3698	9.8554	9.8354	10.4454	8.2185	9.8405
2	8.2686	9.8554	9.8354	10.4512	8.2112	9.8335
3	8.3641	9.8305	9.8354	10.4380	8.2185	9.8324
4	8.4615	9.8359	9.8354	10.4321	8.2212	9.8334
5	8.3589	9.9474	9.8354	10.4280	8.2185	9.8343
6	8.2564	9.8354	9.8354	10.4321	8.2185	9.8356
7	8.3538	9.8397	9.8354	10.4284	8.2185	9.8362

Table 4: Replenishment Quantity Forecast for the Next 7 Days

Time	Flower Leaves Category	Cauliflower Category	Aquatic Roots and Stems Category	Solanaceous	Chili	Edible Mushrooms
1	83.9096	48.5599	38.7724	37.9893	108.2921	49.6733
2	83.9094	48.5167	38.7429	37.9654	108.3105	49.6586
3	83.9192	48.4635	38.7133	37.9416	108.3289	49.6039
4	83.9293	48.4303	38.6837	37.9176	108.3474	49.5491
5	83.9388	48.3872	38.6582	37.8938	108.3659	49.4944
6	83.9386	48.3480	38.6246	37.8679	108.3842	49.4366
7	83.9485	48.3007	38.5981	37.8460	108.4097	49.3848

4. Conclusions

The findings of this study lead to the conclusion that, while the BiLSTM model based on the White Shark Optimization Algorithm has successfully determined the replenishment quantity and pricing strategy that maximize revenue, further improvement in accuracy necessitates a more substantial dataset. The LSTM model exhibits proficiency in handling time-series data; however, to better capture sales trends and fluctuations, a larger and more diverse dataset is required. This could be achieved by expanding the scope of data collection, increasing data sampling frequency, or introducing additional relevant features.

Moreover, integrating deeper domain expertise in sales data could potentially enhance the predictive performance of the model. Understanding factors such as market dynamics, seasonal influences, and consumer behavior provides the model with a more comprehensive context, thereby strengthening its applicability in real-world business environments.

In summary, this research provides decision-makers with an effective tool based on deep learning and optimization algorithms for formulating optimal replenishment and pricing strategies. Nevertheless, for further enhancement of the model's accuracy and robustness, future studies may focus on improving

ISSN 2522-3488 Vol. 7, Issue 9: 57-64, DOI: 10.25236/IJNDES.2023.070910

data quality and optimizing the model structure to meet the evolving demands of the market.

References

[1] Zhu K , Wang H , Kong Y ,et al. Optimization of the Replenishment Strategy for the Supplier Based on Genetic Algorithm[J].International Journal of Business & Management, 2011, 6(1):218-222. DOI:10.5539/ijbm.v6n1p218.
[2] Shi M , Liao X , Yang H ,et al. Research on Vegetable Greenhouse Strategy Based on Multi-objective Distributed Constraint Optimization[C]//IEEE International Conference on Artificial Intelligence and Computer Applications. IEEE, 2021.DOI:10.1109/ICAICA52286.2021.9498123.
[3] Xinbo Huang, Weitao Jiang,Yongcan Zhu,Yi Tian. Transformer fault prediction based on time series and support vector machine[J]. High Voltage Technology, 2020, 46(07):2530-2538.
[4] Yanyan Dong. Research on the construction of ARIMA-BiLSTM model in vegetable price prediction under different combination methods Hong Qihan. [D]. Northeast University of Finance and Economics, 2022.
[5] Yan-Qiu L , Ge C , Ying Z ,et al.Optimization model for replenishment and distribution problems based on grouping strategy[J].Shenyang Gongye Daxue Xuebao/Journal of Shenyang University of Technology, 2017, 39(2):165-169.DOI:10.7688/j.issn.1000-1646.2017.02.09.
[6] Gu Sihong. A study on dynamic pricing of fresh products in H-retailers considering freshness variation [D]. Donghua University, 2023.