Replenishment Pricing of Vegetable Commodities in a Supermarket as an Example in China

Jingzhe Lin, Chang Xu*, Ke Li

College of Science, Minzu University of China, Beijing, China, 100081

* Corresponding Author Email: 18052679075@163.com

Abstract. In this paper, for the automatic pricing and replenishment decision-making problem of vegetable commodities, it is found that there are certain seasonal characteristics in the sales of vegetables in each category, and wavelet analysis is carried out on the daily sales data of single products to get the sales cycle of each category. For the vegetable single product with large amount of data, all single products are divided into four categories of common household vegetables, packaged vegetables, seasonal special vegetables and special vegetables by systematic cluster analysis, and the time distribution characteristics of single products are described by the sales time series of typical members in each cluster. Based on the market demand theory, the cost-plus pricing and sales volume data of the categories were fitted, and the relationship was obtained to be negatively correlated. Taking the purchase quantity and pricing of each category as the decision variables, maximizing revenue as the objective function, and taking the historical maximum purchase quantity of each category and the historical maximum purchase quantity of the superstore as the constraints, a nonlinear programming model was established to formulate the pricing and replenishment strategies for vegetable products. By solving the model, the optimal strategy of replenishment volume and pricing for the future seven-day category is derived, corresponding to the maximum profit of ￥3124.75.

Keywords: Vegetable Commodity Replenishment, Wavelet Analysis, Cluster Analysis, Linear Programming Model.

1. Introduction

Vegetables are usually purchased from 3:00 a.m. to 4:00 a.m. Merchants are required to make replenishment decisions for the day based on historical sales and demand for each commodity without determining the specific individual product and purchase price [1]. From the perspective of market demand, there is often a correlation between the sales volume of vegetables and the time of day; from the perspective of commodity supply, the supply of vegetables is more abundant in April-October, at which time, due to the limitations of the sales space of the superstore, the development of a reasonable sales mix is the key to maximize the benefits [2-3].

Pricing and replenishment strategy of commodities is an important issue to enhance the profitability of retail enterprises, for which some scholars have already conducted relevant research. Yali Li studied the joint replenishment strategy and pricing of fresh commodities under the influence of preservation technology, and found that a joint replenishment strategy can achieve greater profits than individual replenishment [4]; Zhipeng Sun investigated the multi-channel supply chain pricing and coordination problems under centralized decision-making, decentralized decision-making, and coordinated decision-making, and analyzed the coordinated pricing problems of commodities from the perspective of both producers and retailers [5]; Jia Chen investigated the optimal pre-sale dynamic pricing and inventory duration in an e-commerce environment by establishing a pricing and inventory model under dynamic demand [6].

2. Vegetable merchandise sales analysis

The data were obtained from http://www.mcm.edu.cn/, which included the date of sale of vegetable items, time of code-sweeping sales, individual item code, individual sales volume (kg), type of individual sales, and discounting.

2.1. Seasonal fluctuation characteristics of sales volume of each vegetable category based on wavelet analysis

Since the two ends of the time series may produce boundary effects, in order to eliminate the influence of boundary effects on the analysis of the data at the beginning or end points of the series, we extend the length of the data by replicating and filling in the edge part of the data by signal extension. At the same time, the data are wavelet transformed using the complex Morlet wavelet function [7-8], where the mother wavelet function formula is:

$$\phi(z) = \exp\left(-\frac{t^2}{2}\right) * \exp(jw_0 t) \tag{1}$$

After telescopic translation of the complex Morlet wavelet function, we obtain:

$$\phi_{(a,b)}(z) = \frac{1}{\sqrt{a}} * \exp\left(-\frac{(t-b)^2}{2a^2}\right) * \exp\left[\frac{jw_0(t-b)}{a}\right] \tag{2}$$

where a is the scale factor and b is the translation distance. The wavelet can be scaled and transformed by changing the magnitude of a. As a result, the complex Morlet wavelet transform equation is obtained as:

$$WT(a,b) = \int_{-\infty}^{+\infty} f(z) * \phi^*_{(a,b)}\left(\frac{z-b}{a}\right) dz \tag{3}$$

Then there are wavelet coefficients calculated as:

$$|WT(a,b)| = \sqrt{WT_r(a,b)^2 + WT_i(a,b)^2} \tag{4}$$

Finally, the complex wavelet coefficients of the data after extension are obtained according to the extreme of the formula. Delete the complex wavelet coefficients corresponding to the expanded data to get the value of the original sequence corresponding to the complex wavelet coefficients. According to the calculated complex wavelet coefficients, calculate the square of the real part of the mode of the wavelet coefficients, and finally calculate the complex wavelet coefficients of the variance, according to the calculated value of the variance of the wavelet variance of the time series of the sales of the six categories plotted as shown in Figure 1.

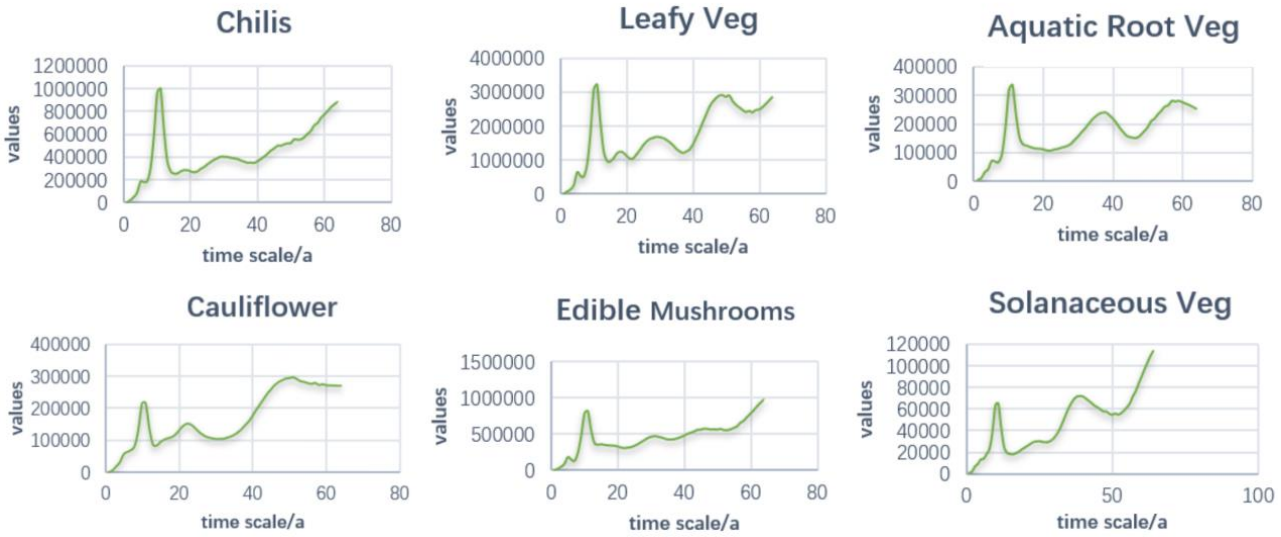


Figure 1. Wavelet variance diagram of each category

From the wavelet variance plot of each category, there are at least 2 obvious peaks in each category, i.e., there are at least two main cycles, and excluding cauliflower and Solanaceous veg, the first main cycle of the other categories are all 11. According to the obtained value of the first main cycle, the main cycle of each category is plotted as shown in Figure 2.

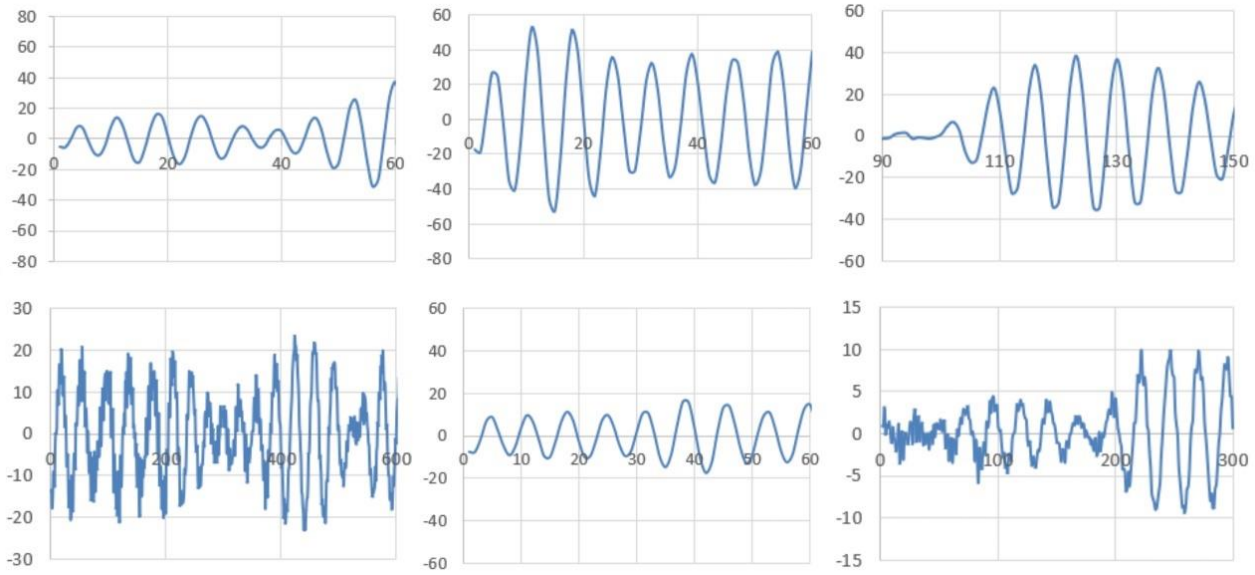


Figure 2. Main cycle chart of sales volume of each vegetable category

From Figure 2, it can be seen that the cycle of sales of chili, leafy veg, aquatic roots veg is about 7 days; the cycle of sales of cauliflower is about 30 days; the cycle of sales of edible mushrooms is about 3 days; the cycle of sales of solanaceous veg is about 25 days.

2.2. Potential patterns between the sales volume of each individual product based on cluster analysis

In order to explore the interrelationship between the sales volume of each single product, we carry out cluster analysis of the sales volume of all single products [9]. After analyzing the spectrogram of clustering results, we finally determined that the number of clusters selected for 5 clustering effect is the best, the specific clustering effect is shown in Table 1:

Table 1. Clustering effect table

Cluster	Members
Cluster 1	Niu Shou Lettuce, Sichuan Red Toon, Cabbage Moss, Young Pumpkin Stems. Radish Leaf, Hemerocallis citrina Baroni, etc.
Cluster 2	Local Chinese Cabbage, Black Rape, Baby Lettuce, Suizhou black rape, Yunnan lettuce (portion), etc.
Cluster 3	Absinthium, Malabar Spinach (portion), Water Chestnut, Honghu Oudai, Fruit Peppers, etc.
Cluster 4	Fresh Dumpling Leaves (bag)(1)
Cluster 5	Chinese cabbage seedling

Cluster 1 consisted of 128 individual items, mainly characterized by high overall sales, belonging to household vegetables, and having very obvious seasonal sales characteristics. The time-series graphs of sales volume of amaranth, water spinach, and Shanghai green in cluster 1 are shown in Figure 3.

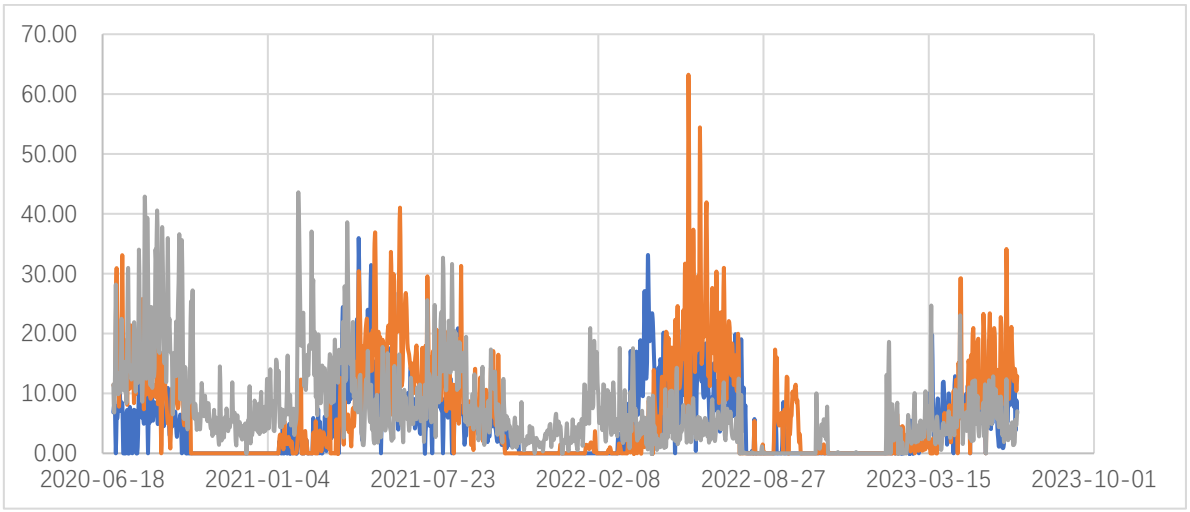


Figure 3. Time series diagram of representative sales volume of cluster 1

Cluster 2, with a total of 66 individual items, is mainly characterized by rapid sales growth after the small sales peak in 2022, and most of them are packaged sales items, which belong to packaged vegetables. The time series graphs of sales volume of Yunnan lettuce (portion), screw pepper (portion), and Yunnan romaine (portion) in cluster 2 are shown in Figure 4.

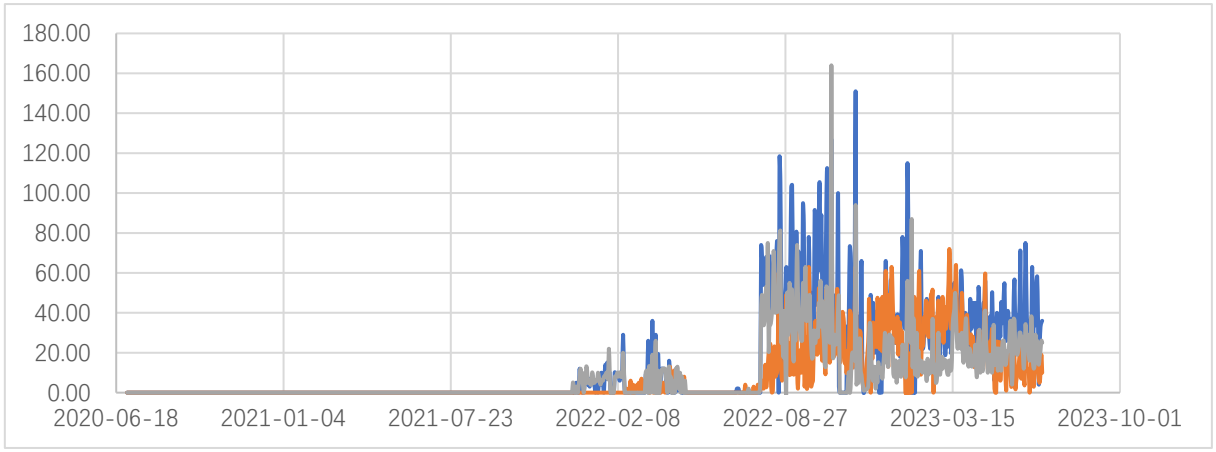


Figure 4. Time series graph of representative sales volume of cluster 2

The third cluster has 10 individual products, which are mainly characterized by an obvious seasonal effect, with a certain sales volume only at a specific time, belonging to seasonal special dishes. The representative time series diagram of the sales of red lotus root scallops, water caltrop, and fruit peppers is shown in Figure 5.

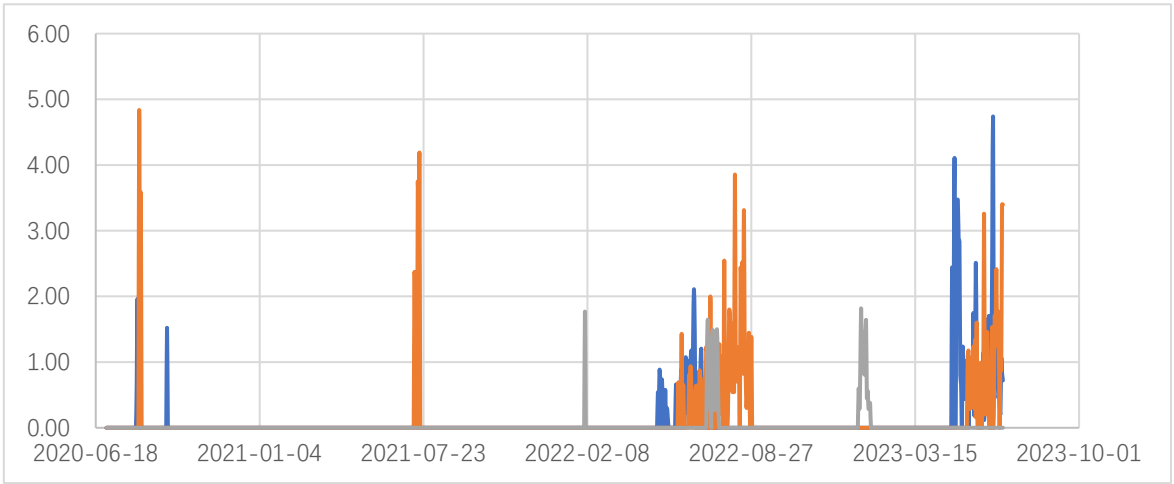


Figure 5. Time series graph of representative sales volume of cluster 3

The fourth and fifth clusters all have only one member and have only a few sales data over the years, they are special dishes and do not have any time sales characteristics, and we record them as a separate category in the classification.

3. Developing pricing and replenishment strategies for vegetable categories based on linear programming models

3.1. Fitting the relationship between total sales volume and cost-plus pricing of vegetable category

Based on the definition of market demand curve (market demand curve) of economic theory, we know that there is some kind of relationship between commodity pricing and market sales. However, the question only provides the daily pricing of each individual item, which requires us to calculate the daily pricing of each category.

We define:

$$Daily\ retail\ price\ of\ each\ vegetable\ category = \sum \frac{Daily\ purchase\ price\ of\ each\ vegetable\ category \times Supply}{Stock\ intake\ of\ this\ vegetable\ category\ on\ this\ day} \tag{5}$$

Therefore, we can get the daily pricing of each vegetable category and the total daily sales of each vegetable category.

We set q_i to denote the total amount of sales of the i-th category of vegetables, and the corner label i corresponds to the total amount of sales of chilis, leafy veg, aquatic root veg, cauliflower, edible mushrooms, and solanaceous veg, respectively.

We used the matlab curve fitter to perform a linear fit between total sales and pricing for each vegetable category. Obtained:

$$q_i = -11.25 * p_i + 75 \quad (i = 1, \dots, 6) \tag{6}$$

The linearly fitted R2 and Adjusted R2 are shown in Table 2.

Table 2. Linear fit R2 and Adjusted R2 table

i	1	2	3	4	5	6
R2	0.8871	0.8271	0.9689	0.8546	0.9631	0.8563
Adjusted R2	0.887	0.8269	0.9688	0.8544	0.963	0.8562

Therefore, we can get the total sales and cost-plus pricing relationship generalized as:

$$q = k * p + c \tag{7}$$

The two show a negative correlation, i.e., the higher the cost pricing, the lower the total sales.

3.2. Time series modeling to predict the price of vegetable inputs

We conducted a time series analysis of the wholesale prices of each category of vegetables for April-June 2023 (i.e., the feed prices of each category). Firstly, the original time series was differenced to the 1st order, and after the unit root test, the differenced series has reached the requirement of smoothness (P<0.01), which can be regarded as a smooth time series. Subsequently, we plot ACF and PACF plots for the smoothed time series to obtain the time series model ARIMA (0, 1, 1) [10].

We analyze the time series of wholesale prices of leafy veg, cauliflower, aquatic root veg, solanaceous veg, chilis, edible mushrooms in April-June 2023 to get the predicted wholesale prices in July 1-7, 2023, and the predicted results are shown in Table 3.

Table 3. Predicted values of wholesale prices of vegetables from July 1-7, 2023

	7/1	7/2	7/3	7/4	7/5	7/6	7/7
Leafy Veg	3.26	3.26	3.26	3.26	3.26	3.26	3.25
Cauliflower	7.93	7.93	7.92	7.92	7.91	7.90	7.90
Aquatic Root Veg	11.93	12.07	12.14	12.21	12.28	12.34	12.41
Solanaceous Veg	4.74	4.59	4.58	4.63	4.58	4.52	4.54
Chilis	3.59	3.51	3.46	3.42	3.40	3.38	3.36
Edible Mushrooms	3.67	3.66	3.66	3.65	3.64	3.64	3.63

3.3. Pricing and replenishment strategy model for vegetable categories based on linear programming

Let Q_{ij} denotes the incoming quantity of the j-th category of vegetables on day i, P_{ij} denotes the pricing of the j-th category of vegetables on day i, q_{ij} denote the total quantity of vegetables sold on day i for the j-th category of vegetables, m_{ij} denotes the purchase price of the j-th category of goods on day i, ($i = 1,\ldots,7, j = 1,\ldots,6$), while $j = 1,\ldots,6$ corresponds to the category of chilis, leafy veg, aquatic root veg, cauliflower, edible mushrooms and solanaceous veg, respectively.

Based on the above obtained linear relationship between total sales and pricing for each vegetable category, it is obtained:

$$q_{ij} = k * P_{ij} + c \tag{8}$$

We compare the quantity demanded with the quantity of sellable incoming goods and take the smaller value to calculate the daily profit, i.e., $min\{q_{ij}, (1 - r_j)Q_{ij}\} * (P_{ij} - m_{ij})$, and subtract the daily loss of unsold goods as $max\{Q_i - q_{ij}, 0\} * m_{ij}$, i.e., the maximum profit.

We set q* to denote the total sales of all categories in a day, z to denote the profit, m_{ij} to denote the predicted wholesale price of the jth vegetable item on day i of July, r_{ij} to denote the wastage rate of the jth vegetable item on day i of July, i=1,...,7, j=1,...,6, and j=1,. ...,6 when they correspond to chili, foliar, aquatic, cauliflower, edible mushrooms and eggplant, respectively.

We constrain that the pricing will always be no less than the purchase price, and that the sales of vegetables in each category in the next seven days will not be greater than the maximum value of the previous sales of that category, nor less than the minimum value of the previous sales of that category, and at the same time, we constrain that the total sales of vegetables in all categories on each of the next seven days will not be greater than the maximum value of the total sales of all categories on each of the previous days, nor less than the minimum value of the total sales of all categories on each of the previous days, and so we build the model as follows:

$$\begin{aligned}\max \sum z = \max \sum_{i=1}^{6} \sum_{j=1}^{7} \{\min\{q_{ij}, (1 - r_j)Q_{ij}\} * (P_{ij} - m_{ij}) \\ - \max\{Q_{ij} - q_{ij}, 0\} * m_{ij}\}\end{aligned} \tag{9}$$

$$\begin{cases} P_{ij} \geq m_{ij} \\ \min q \leq q_{ij} \leq \max q \\ \min q^* \leq \sum_{j=1}^{6} q_{ij} \leq \max q^*, i = 1, \ldots, 7 \end{cases} \tag{10}$$

Solving this linear programming model yields Table 4:

Table 4. Pricing and replenishment policy by vegetable category

category	pricing	replenishment
Chilis	6.54	85.05
Leafy Veg	7.63	401.1
Aquatic Root Veg	3.24	41.25
Cauliflower	7.91	57.88
Edible Mushrooms	5.29	73.09
Solanaceous Veg	4.57	105.86

Therefore, we obtain the pricing and replenishment strategy for each vegetable category for July 1-7, 2023 as shown in Table 4, with a maximum profit of ￥3,124.75 for all seven days when goods are purchased at the replenishment levels as above as well as sold at the pricing as above for each vegetable category.

4. Conclusion

In this paper, the sales cycle and time distribution characteristics of each vegetable are obtained through wavelet analysis and systematic cluster analysis, and then the pricing and replenishment problems of vegetable commodities are solved through a linear planning model. The planning model considers comprehensively, including the situation that the demand is greater than the incoming quantity and the incoming quantity can't be sold in full, and cleverly avoids the calculation of the discount rate, which is included in the attrition rate; however, the planning model does not take into account the increase in the cost of labor, water, and electricity due to the increase in the incoming quantity, so that the program has some errors.

This paper includes a system clustering model, a demand pricing model, a stocking planning model, and a pricing planning model. This is a relatively complete set of models to solve the super planning problems, but also can be further promoted, he applies to other classes of goods planning problems, the need to consider more on the basis of the original model of the daily inventory; and can be solved for fresh supermarkets, such as stocking, pricing, and other issues.

References

[1] Fan Fengyan. Research on pricing model of supermarket fresh goods [D]. Fujian Normal University,2020.

[2] Mao Lisha. Research on Pricing Strategy and Production and Marketing Mode of Vegetable Wholesale Market under the Perspective of Supply Chain[D]. Central South Forestry University of Science and Technology,2023

[3] HU Hanli, CAO Yu, WU Kan. Research on sales model selection and pricing of fresh supply chain based on pre-sale[J]. Operations Research and Management,2022,31(11):128-134.

[4] Yali Li. Joint replenishment and pricing collaborative decision making considering investment in preservation technology [D]. Chongqing Jiaotong University,2023.

[5] Sun ZP. Research on multi-channel supply chain pricing and coordination under O2O model[D]. Beijing Jiaotong University, 2021.

[6] Chen Jia. Joint decision making of pricing, ordering and preservation of fresh agricultural products under pre-sale model[D]. Southwest Jiaotong University, 2022.

[7] He Junqi, Guo Xinjia, Chen Yunfei et al. A multi-timescale study of rainfall series based on Hilbert-Yellow transform and wavelet analysis[J/OL]. Soil and water conservation research:1-12[2023-11-18].

[8] YINHAN LIAO, HONGWEI WANG, TONGSHOU LU et al. Periodic variation characteristics of energetic electrons in near-Earth space studied by wavelet analysis[J]. Astronomical Journal,2021,62(03):69-86.

[9] ZHOU Yu,DUAN Yongrui. Retail merchandise sales prediction based on clustering and machine learning[J]. Computer System Applications,2021,30(11):188-194.

[10] TIAN Ge,WU Kuang,TIAN Hongrui. Matplotlib-based K-means-ARIMA model for sales prediction of retail goods during traditional holidays[J]. Computer Knowledge and Technology,2022,18(18):59-60+63.