

Research on Dynamic Pricing Strategy of Fresh Supermarket Based on ARIMA and Integer Programming Model

Zhuoxin Chen^{1,*}, Wenlu Liu², Zhongyan Lin²

¹ Department of Mathematics and Information Science Guangzhou University, Guangzhou, China, 510006

² School of Mathematics and Information Science, Guangzhou University, Guangzhou China, 510006

* Corresponding Author Email: 13268119468@163.com

Abstract. The study of automatic pricing and replenishment decision-making of vegetable products is a key measure for modern fresh food merchants to pursue high efficiency, accuracy and competitiveness. In this paper, the time series ARIMA(p,d,q) model and 0-1 integer programming model were used to realize the intelligent replenishment and pricing decision of fresh supermarkets, so as to ensure the reasonable pricing and sufficient inventory of vegetable products and meet the changing market demand. This paper obtains the replenishment strategy in the next week and the maximum profit of vegetable sales in Supermarket on July 1 is 3831.2 RMB.

Keywords: ARIMA, 0-1 integer programming, Dynamic pricing, Replenishment strategy, Violence search.

1. Introduction

In today's society, life supermarkets can be seen everywhere. The daily necessities of supermarkets greatly facilitate the daily needs of local residents [1], and also promote local economic development [2]. With the upgrading of the preservation system [3], the variety of goods in the living supermarket is gradually increasing, and fresh meat products and various vegetables can also be purchased in the supermarket. However, the shelf life of each commodity is different [4], and the freshness of the commodity is often an important consideration for customers to purchase [5].

Vegetables are an indispensable food in life. They are fresh food and have a short shelf life. The freshness of vegetables in supermarkets will decrease with the increase of time on the shelves. Many kinds of vegetables cannot be sold again the next day if they cannot be sold on the same day [6]. Therefore, it is very meaningful to study the sales of various commodities and the needs of local residents to determine the quantity of purchases and the price of goods to reduce the waste of goods. [7].

2. Replenishment strategy based on time series ARIMA (p, d, q) and 0-1 integer programming

According to the daily replenishment data of six major categories of vegetables, the time series diagram of the sales volume of six major vegetables with time is first drawn, and the ARIMA(p,d,q) time series model is established [8] to predict the daily replenishment volume of each vegetable in the next week. Then, the constraints are used to find the number of single items available for sale and establish a 0-1 integer programming model based on violent search [9]. Finally, the single product replenishment and pricing strategy that maximizes the profit on July 1 under the market demand are obtained [10].

2.1. Establishment of replenishment model based on time series and its algorithm steps

2.1.1 Construction of the model

(Data source of this paper:

http://www.mcm.edu.cn/html_cn/node/c74d72127066f510a5723a94b5323a26.html)

Since the sales volume of various vegetables changes with time, and the time interval is one day, the sales volume of various vegetables can be regarded as a random variable. The variable interval is one day, and the corresponding data is continuous. Therefore, the time series model can be used to predict the replenishment of various vegetables.

In this paper, the ARIMA model in time series was used to test the stationarity, determine the autoregressive coefficient and the partial regression coefficient before using the model, and test the white noise of the model residual.

2.1.2 Algorithm steps

Step 1: stationarity test, by making a time series diagram, observe whether there are abnormal values, if there are abnormal values, you need to transform the data, or use the difference method to process the data.

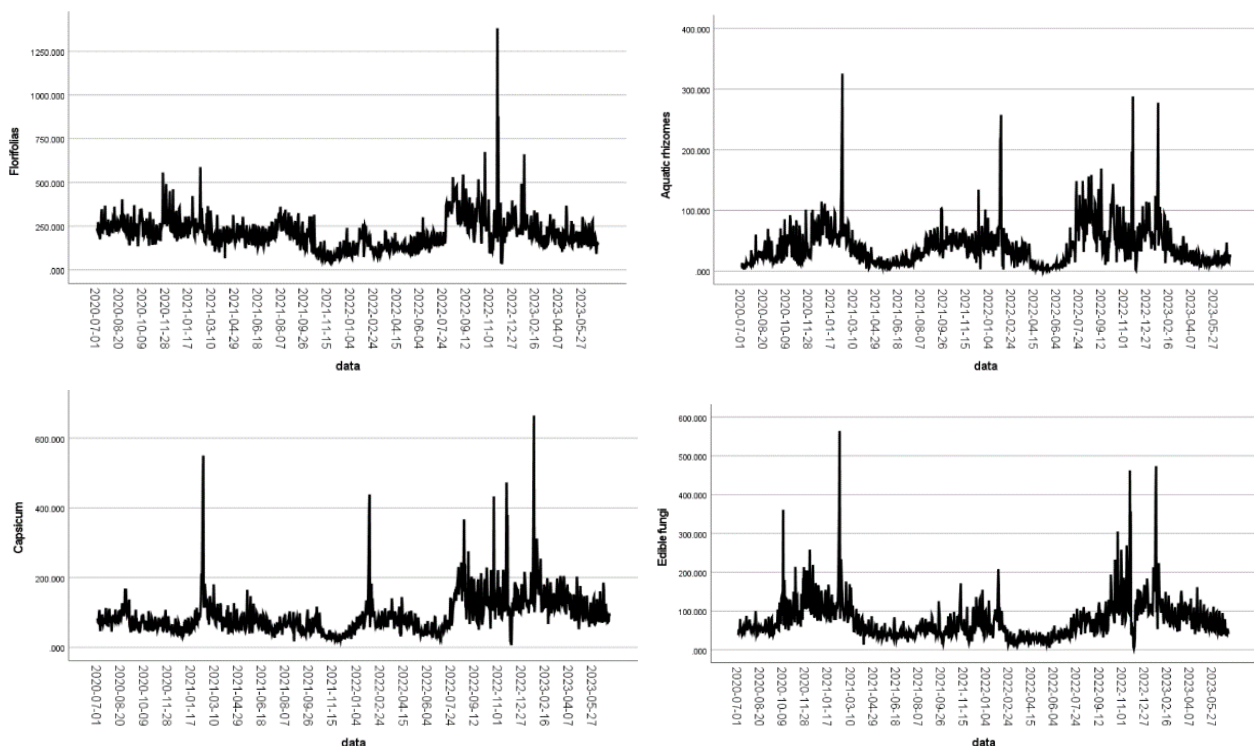


Figure. 1 Time sequence diagram of four types of vegetables

(Note: Considering the large number of vegetables, only the time series diagrams of cauliflower, eggplant, pepper and edible fungi are shown here.)

It can be seen from Figure.1 that from left to right and from top to bottom, the total daily sales of cauliflower, eggplant, pepper and edible fungi changed with time, and the trend of change was relatively stable, fluctuating up and down around a mean line, with no upward or downward trend, so there was no need for differential treatment. (Floral and rhizome vegetables that are not shown in the picture also show a roughly stationary distribution)

Step 2: Determine the autoregressive correlation coefficient p and partial regression correlation coefficient q . The corresponding ACF and PACF diagrams are drawn as shown in Figure 2:

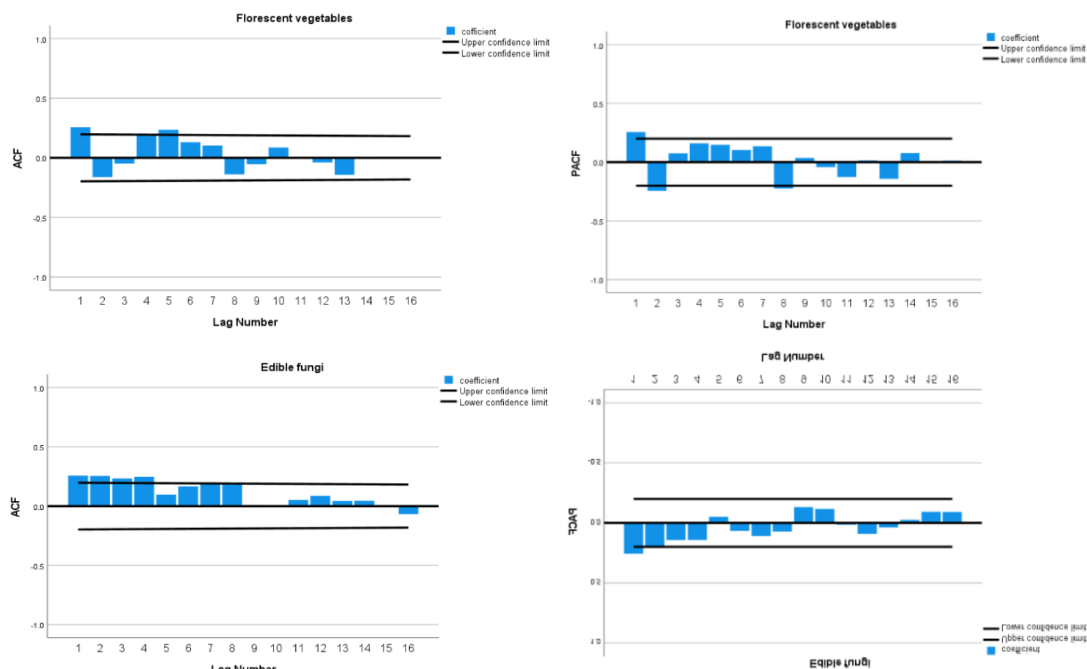


Figure. 2 ACF (left) and PACF (right) of the sales of cauliflowers and edible fungi vegetables.

(Note: Considering that there are many categories of vegetables, here only ACF and PACF charts showing the sales of cauliflower and edible fungi vegetables are shown.)

Figure.2 depicted the ACF and PACF of cauliflowers and edible fungi, from which we can see the number of autoregressive terms and moving average terms. Among them, the number of autoregressive terms of cauliflower is $p=1$, that is, 1-order tailing, and the number of moving average terms is $q=9$, that is, 9-order tailing ; the number of autoregressive terms of edible fungi is $p=4$, that is, 4-order tailing, and the number of moving average terms is $q=6$, that is, 6-order tailing. (The autoregressive coefficients and moving average terms of the other four types of vegetables not shown in the Figure are divided into: the number of autoregressive terms of flower and leaves and peppers $p=1$, and the number of moving average terms $q=9$; the number of autoregressive terms $p=1$ and the number of moving average terms $q=3$ for aquatic rhizomes; the number of autoregressive terms $p=1$, the moving average number of terms $q=7$).

Step 3: Perform residual white noise test. The ACF diagram and PACF diagram of the residuals are made in Figure 3.

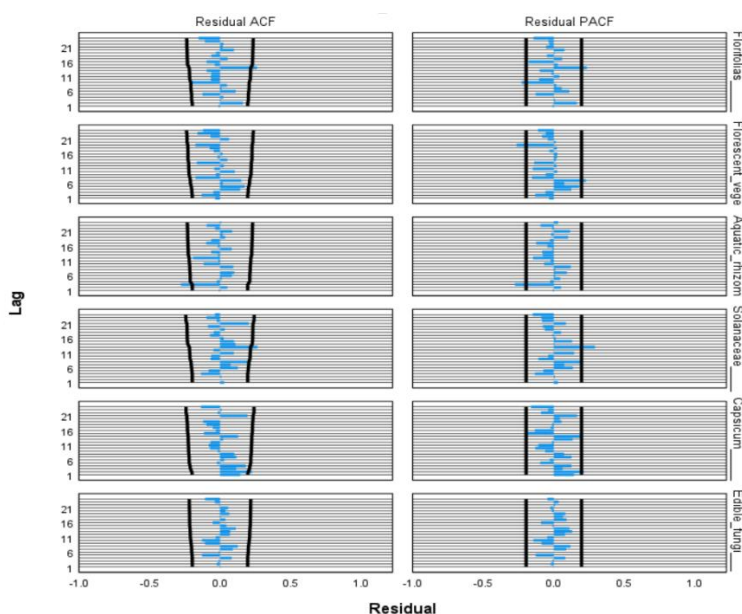


Figure. 3 ACF diagram (left) and PACF diagram (right) of residuals of Figure 3)

It can be seen from Figure.3 that the autocorrelation coefficients and partial autocorrelation coefficients of all lag orders of various vegetables are basically distributed around 0, that is, the residuals are basically within the two lines, that is, there is no significant difference with 0.

However, it is not accurate enough to observe only from the Figure. Therefore, the Q test is performed on the residuals. The test results are shown in the following table :

Table. 1 Q test results

model	Model fitting degree statistics	Young-Box Q (18)		
	Smooth R ²	statistics	DF	significance
Flora Model1	0.336	23.336	16	0.105
Cauliflower Model2	0.102	19.664	17	0.292
Pepper model3	0.412	20.084	17	0.270
Aquatic rhizomes model4	0.573	23.718	14	0.051
Eggplant model5	0.463	22.173	15	0.103
Edible Fungi Model6	0.543	11.829	16	0.756

It can be seen from Table 1 that the p value obtained by Q test on the residuals of flower and leaves, cauliflowers, peppers, aquatic rhizomes, eggplants and edible fungi, that is, the significance level is greater than 0.05, indicating that the null hypothesis cannot be rejected (null hypothesis: H0: residuals are white noise sequences), indicating that the residuals of flower and leaves, cauliflowers, peppers, aquatic rhizomes, eggplants and edible fungi are white noise sequences. Therefore, ARIMA (p,d,q) model can be truly established.

Step 4: Establish ARIMA (p,d,q) model.

The number of autoregressive terms p and the number of moving average terms q obtained in Figure 1 and the results obtained with the help of SPSS expert modeler are as follows:

Table. 2 ARIMA model results

statistical model			
	Type	model	model types
Model ID	Flower and leaves	Model1	ARIMA(1,0,9)
	Cauliflower	Model2	ARIMA(1,0,9)
	Pepper	Model3	ARIMA(5,0,7)
	Aquatic rhizome	Model4	ARIMA(0,0,1)
	Eggplant	Model5	ARIMA(0,0,9)
	Edible fungus	Model6	ARIMA(0,0,9)

It can be seen from Table 2 that in the ARIMA model corresponding to the six major vegetables of flower and leaves, cauliflowers, aquatic rhizomes, eggplants, peppers and edible fungi, d = 0, there is no need for differential treatment.

Therefore, the model established is:

$$\begin{cases} h'_t = \alpha_0 + \sum_{i=1}^p \alpha_i h'_{t-i} + \varepsilon_t + \sum_{i=1}^q \beta_i \varepsilon_{t-i} \\ h'_t = \Delta^d y_t = (1 - L)^d h_t \end{cases} \quad (1)$$

Among them, y'_t is the sales volume of various vegetables, t is the time node, ε_t is the interference term, β_i is the moving average coefficient, L is the lag operator, α_i is the autoregressive coefficient.

2.2. The establishment of violence search model based on 0-1 integer programming and its algorithm steps

2.2.1 Construction of the model

In this paper, according to the constraint conditions, establishes a 0-1 integer programming model, obtains the number of items that meet the ordering conditions and meet the demand, and then uses the violent search, and then combines the time series of the six major vegetables on July 1. The replenishment model, and then use the final search for the optimal solution to obtain the maximum income of the merchant.

2.2.2 Algorithm steps

Firstly, using the data from June 24 to June 30, we can get 49 kinds of available sales orders, calculate the profit coefficient and sort them according to the profit level.

$$f_i = (1 - l_i) k_i - m_i \quad (2)$$

f_i is the profit coefficient, l_i is the loss of type I vegetables, k_i is the pricing of type I vegetables, and m_i is the cost of type I vegetables.

Then the 0-1 integer programming is carried out to obtain the number of items that meet the constraints by listing the constraints by the title:

Constraint conditions:

$$\begin{cases} 27 \leq x_i \leq 33 \\ x_i = 0 \text{ or } 1 \\ \sum_{i=1}^{q_a} y_i \geq M_a \\ y_i \geq 2.5 \end{cases} \quad (3)$$

x_i is the item in the i , y_i is the order quantity corresponding to the i item, q_a is the total number of items of the first a vegetable in the 49 categories divided into six categories of vegetables, and M_a is the replenishment quantity (demand) of the first a vegetable obtained.

Finally, the goods obtained above are violently searched by using the replenishment model and pricing decision model based on time series to find the optimal pricing scheme, so as to maximize the benefits of the supermarket.

According to the loss rate of each item, the above selected 27 categories of items, the average loss rate of each category of vegetables is calculated from the previous data, and the average loss rate of the j Category in the 27 categories of items is recorded as l_j ; let the total sales volume of the j th kind of vegetables in 2020.7.1 be c_j^* , and arrange c_j^* from small to large. In order to eliminate the interference of outliers and avoid the contingency of the experiment, the cost pricing is taken as a 95% comparison interval. The total sales volume of the j th kind of vegetables on the i th day is recorded as c_{ij} , where $i = 1, 2, \dots, 1030$, $j = 1, 2, \dots, 27$; s_j of class j in July 1, 2023, where $a = 1, 2, \dots, 7$; the average price of all kinds of vegetables is recorded as k_j , and the average wholesale price is recorded as m_j ; the profit of category j vegetables is set to w_j .

When $c_{ij} \geq s_j$, ($i = 1, 2, \dots, 1030$)

$$w_{aj} = s_j \cdot (1 - l_j) \cdot ((1 - p_i) * k_j + p_i * k_j) - s_j \cdot m_j \quad (4)$$

When $c_{ij} < s_j$, ($i = 1, 2, \dots, 1030$)

$$w_{aj} = \begin{cases} c_j \cdot ((1 - p_i) * k_j + p_i * k_j) - s_j \cdot m_j & , c_j \leq s_j \cdot (1 - l_j) \\ s_{aj} \cdot (1 - l_j) ((1 - p_i) * k_j + p_i * k_j) - s_j \cdot m_j & , c_j > s_j \cdot (1 - l_j) \end{cases} \quad (5)$$

Violence traverses all data from the above equation to obtain a profit matrix for each type of vegetable on July 1, which is recorded as A , and the corresponding maximum profit is recorded as w_j^* . The maximum profit of each vegetable is extracted and summed every day, and the sum of the

profits of various vegetables can be obtained (as shown in Equation (5)), that is, the maximum profit W is obtained.

$$W = \sum_{j=1}^6 w_j^* \quad (6)$$

3. Experimental results and analysis

3.1. Solution of replenishment model based on time series

By using SPSS to solve (2), the replenishment quantity from July 1 to 7 is obtained as follows:

Table. 3 Replenishment of various vegetables

Days	flower and leaves	cauliflower	aquatic rhizome	solanaceous	pepper	edible fungi
1	144.727	25.531	24.082	20.889	104.369	47.444
2	139.689	23.386	24.248	19.432	103.516	49.905
3	138.698	22.391	24.414	19.925	86.166	53.373
4	153.232	23.698	24.582	19.035	75.98	52.28
5	159.25	24.664	24.75	22.527	78.472	50.907
6	152.318	25.906	24.92	23.204	89.111	48.857
7	155.75	27.684	25.091	24.743	102.299	51.379

Note: The above food replenishment units are kg.

It can be seen from Table 3 that the total amount of daily replenishment of six kinds of vegetables from July 1 to July 7, such as flower and leaves, cauliflowers, aquatic rhizomes, eggplants, peppers and edible fungi on the first day, was 144.727 kg, 25.531 kg, 24.082 kg, 20.889 kg, 104.369 kg and 47.444 kg.

3.2. Violence search model based on 0-1 integer programming to solve the maximum profit

There are 27 available items obtained from step3 in 1.2.2. The specific daily replenishment quantity and pricing strategy of each item on July 1 are calculated by using relevant data and Matlab. The following Table.4 only shows the results of the daily replenishment quantity and pricing strategy of some items:

Table. 4 Replenishment quantity and pricing on July 1 (part)

Kinds	Pricing (kg/RMB)	Purchase quantity (kg)	Profit (RMB)	Discount probability
Yunnan lettuce	12.000	4.028	49.592	0.037
Bamboo leafy vegetables	4.000	51.426	97.356	0.022
Pumpkin tip	6.000	13.387	280.500	0.040
flowering cabbage	6.900	39.541	333.866	0.132
Muercai	4.500	23.827	234.249	0.222
Chinese cabbage	4.900	18.141	153.389	0.212
Yellow cabbage (1)	14.000	28.204	385.389	0.038
Spinach	12.000	27.192	353.898	0.073
Tian Qi	9.000	14.169	131.188	0.028
Chinese cabbage at Dongmenkou	6.000	4.381	36.491	0.003

In order to achieve the goal of maximum profit, the sales volume of each single product should be equal to the supply volume. From Table 4, we can see that the specific pricing, purchase volume, profit and discount probability of each item (only some items are listed in the table) on July 1. For example, for Yunnan lettuce, the purchase volume of Supermarket on July 1 should be 4.028 kg, the pricing should be marked as 12.000 kg / RMB, the discount probability is 3.7 %, and the profit is 49.592 RMB. At this time, the maximum profit of all items on July 1 was 3831.2 RMB.

4. Conclusion

This paper obtains the daily replenishment volume of the six major vegetables in Supermarket in the next week and the specific pricing, purchase volume, discount probability and profit of each single product on July 1, and the final maximum income is 3831.2 RMB. In this paper, the effective single product, single product loss rate, pricing, cost and profit coefficient are mainly used. Under the constraint conditions, the specific replenishment plan of each single product on July 1 is obtained through the violent search based on 0-1 integer programming. The purpose is to maximize profits. The experimental results show that this replenishment strategy can make the optimal profit planning for the next week 's excess replenishment volume and the replenishment volume and pricing of the single product on July 1.

References

- [1] Wu Chunlin. A quarter of an hour is a happy life in the convenient life circle. [N] Yinchuan Daily, 2022-01-21 (002).
- [2] Zhao Dan. Y Fresh Supermarket Marketing Strategy Improvement Research [D]. Lanzhou Jiaotong University, 2021.
- [3] Zhang Peng, Hao Congcong, Xue Youlin et al. Research Progress of Vegetable Storage Technology [J]. Packaging Engineering, 2023,44 (05): 111-120.
- [4] Vegetables preservation tips [J]. Changjiang vegetables, 2019 (09): 63.
- [5] YAO Ziwei. Study on the price decision of fresh products considering price and price reduction time point [D]. Southeast University, 2022.
- [6] Huang Jianxing. Vegetable price fluctuation and prediction research [D]. South China Agricultural University, 2019.
- [7] Hu Deyu. Discussion on the main food market demand of urban residents in China and the dietary characteristics of Chinese residents [J].Quality and market, 2021 (05) : 118-120.
- [8] Yang Haimin, Pan Zhisong, Bai Wei. A review of time series prediction methods [J].Computer Science, 2019,46 (01) : 21-28.
- [9] Guo Jidong, Zhang Kaibin, Jiang Junhao, etc. Balance improvement design of No.1 workshop production line based on program analysis and 0-1 integer programming model [J].Electromechanical engineering technology, 2020,49 (02) : 12-16 + 64.
- [10] Sabir B L, Farooque A J. Dynamic pricing and profit maximisation: a simulation-based approach for agri-fresh products retailing in India [J]. Int. J. of Inventory Research,2017,4(4).