Optimization Decision Making of Vegetable Replenishment and Pricing Based on ARIMA Model and Greedy Strategy

Jiayun Xu[1, *], Xiangqi Lin[2], Jiahong Ye[2]

[1] School of Mathematical Sciences, Jiangsu University, Zhenjiang, China, 212013

[2] School of Mechanical Engineering, Jiangsu University, Zhenjiang, China, 212013

*Corresponding author: 3210101012@stmail.ujs.edu.cn

Abstract. Fresh products become the important part in people's daily life. At present, the market of fresh products in China reaches a scale of 5.3 trillion, and the number of weekly purchases of fresh products by ordinary residents reaches about 3 times. Due to their short freshness period, superstores face the challenge of daily replenishment pricing. In order to develop a reasonable replenishment and pricing plan, this paper determines the cost margin in the markup cost pricing method by formula conversion using four indicators: sales volume, wastage rate, cost unit price, and replenishment volume, and further derives a functional relationship between sales volume corresponding to the six categories and cost markup pricing. The ARIMA model is used to forecast the cost margin, cost unit price, and sales volume for the next seven days, and a greedy strategy is used to equate the value of the daily replenishment volume to the predicted value of the sales volume before the loss on that day in order to achieve the goal of reducing the cost and meeting the goal of maximizing the revenue per day with certainty of the sales volume to arrive at the optimal pricing decision and the daily replenishment volume for the next seven days. Finally, the relationship between sales volume and pricing was directly fitted using the least squares method to reformulate the pricing decision, while the model was tested and evaluated. The integrated predictive pricing decision model proposed in this model can be a better solution to the commodity pricing problems related to time series, like supply chain product management system and online vegetable commodity delivery platform. It can also be used in models that contain interactions between time-dimension and space-dimension variables, such as the prediction of traffic flow and people flow.

Keywords: ARIMA model, Greedy strategy, Pricing replenishment decisions.

1. Introduction

In recent years, most of the people are concerned about the date of production and shelf life when they are shopping for edible consumer goods, and cold chain logistics, mechanical refrigeration and other preservation technologies have been developed. [1] According to the data provided by the 2023 National College Students Mathematical Modeling Competition, we can analyze that vegetables are an indispensable class of commodities in superstores. However, due to their short shelf life and deterioration of traits over time, vegetable varieties that are not sold on the same day will not be available for sale the following day. Supermarkets usually restock daily based on historical sales and demand for various commodities. Because of the sale of many varieties of vegetables, different origins, and vegetables are generally early morning stocking, so the businessman does not know exactly the specific single product and commodity prices, to make the day's replenishment decision for each vegetable category, which has become the core of the vegetable sales process.

Replenishment decisions and pricing decisions rely heavily on reliable market demand analysis. From the demand side, there is often a correlation between the sales volume of vegetables and the time of year; from the supply side, the supply of vegetables is more plentiful from April to October, and the limitations of the superstore's sales space make a reasonable sales mix extremely important. When making decisions on vegetable replenishment and pricing optimization, supermarkets need to deeply analyze historical sales data and market demand to ensure that they can meet consumers' needs and maximize profits at the same time.

In this paper, we apply detailed data to explore the relationship between sales volume and cost-plus pricing, and use the ARIMA model to predict the cost margin, cost unit price, and sales volume for the next seven days by category [2][3], and adopt the greedy strategy [4] to equate the value of the daily replenishment volume to the predicted value of the sales volume before the depletion of that day, so as to achieve the goal of reducing costs and meeting the maximum daily revenue under the certainty of sales [5][6] and thus derive the optimal pricing and daily replenishment volume for the next seven days. Finally, the relationship between sales volume and pricing was directly fitted using the least squares method [7][8] to reformulate the pricing decision, while the model was tested and evaluated.

2. A study of vegetable replenishment and pricing optimization decisions based on ARIMA Model with greedy strategies

2.1. Data pre-analysis

Based on vegetable individual wholesale price data, Using the number as the horizontal coordinate and the price as the vertical coordinate, a scatterplot of wholesale prices is plotted as shown in Figure 1 below.

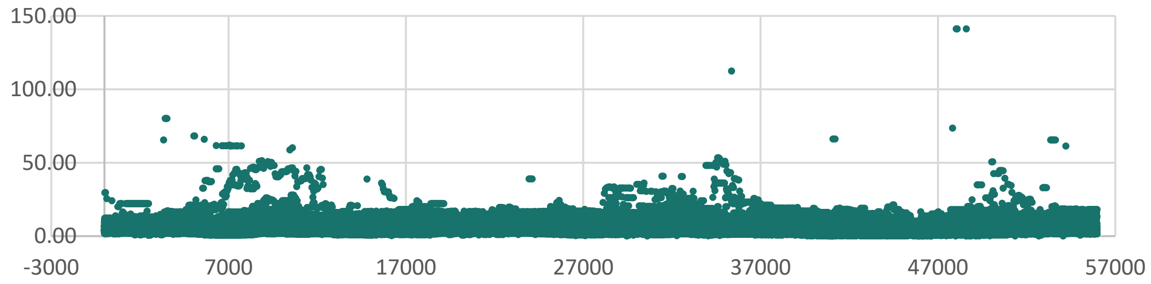

Figure 1. Wholesale price scatter plot

From Figure 1 above, it is concluded that the wholesale price of different vegetable items varies, with most of them located in the range of 0-20 RMB/kg , some data points located in the range of 20-80 RMB/kg , and very few of them with wholesale price up to 140 RMB/kg , with an average wholesale unit price of 5.96 RMB/kg .

A scatterplot was drawn to visualize the vegetable loss volume data, as shown in Figure 2 below.

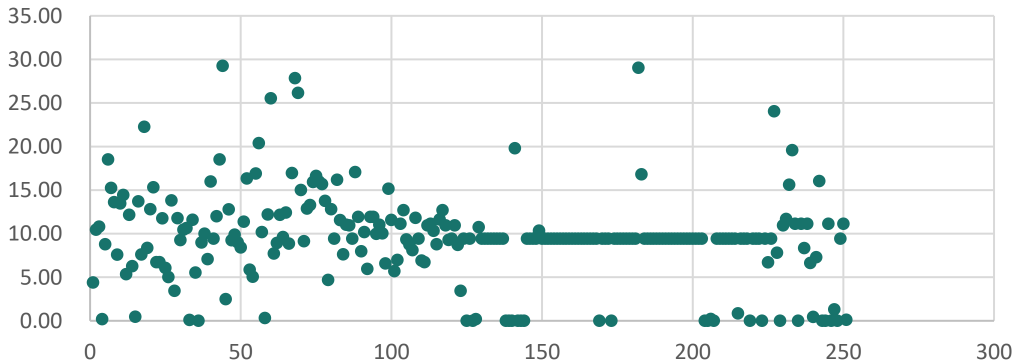

Figure 2. Scatterplot of recent vegetable wastage rates

From Figure 2 above, the wastage rate of most vegetables is concentrated around 10%, a small number of vegetables have wastage rates close to or equal to 0, and only a few have wastage rates higher than 20%.

After preliminary organization of the data, it is found that all the values of the attrition rate fall within [0,29.25%], and are unevenly distributed within this interval. Among them, there are 37 kinds of single products with a loss rate lower than 5%, accounting for 14.74% of the total number of single products; 131 kinds with a loss rate between 5% and 10%, accounting for 52.19% of the total number; 54 kinds with a loss rate between 10% and 15%, accounting for 21.51% of the total number; 21 kinds with a loss rate between 15% and 20%, accounting for 8.37% of the total number; 7 kinds with a loss rate higher than 20%, accounting for 2.79% of the total number.

After the initial analysis of the known data, the replenishment plan is done in terms of category for modeling, and the total flow chart is shown in Figure 3 below.

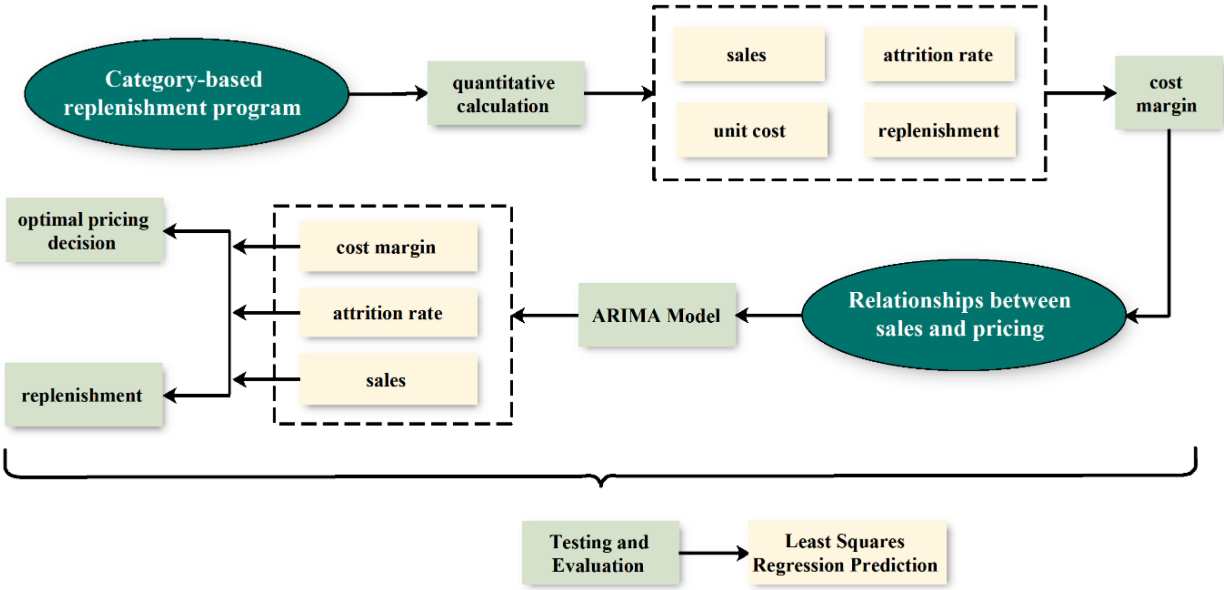

Figure 3. Flowchart for the study

2.2. Sales volume versus cost-plus pricing

In order to obtain commodity profits, it is also necessary to forecast the cost of the vegetable category, this paper uses the ARIMA model to forecast it in the next 7 days. The ARIMA forecasting model is a type of time series forecasting partner that is used in a variety of disciplines, including statistics and econometrics. Its essence is to predict the amount of deferral generated by the model's dependent variable, the amount of deferral generated after the uncertainty bias generated by the model, and the current amount of the model as the time series gradually shows stabilization.

According to the market survey, the pricing of vegetables is based on the "cost-plus pricing" method, which is a method of making pricing decisions for a product based on the unit cost price and a fixed percentage of profit, as shown in the following equation (1):

$$V = M + M \cdot \alpha = M \cdot (1 + \alpha) \tag{1}$$

where is pricing, is cost per unit, and is cost margin.

Since vegetables need to be replenished on a category-by-category basis, the definition of cost-per-unit is no longer universally appropriate, so the cost-per-unit in equation (1) above is corrected to cost, as in equation (2) below:

$$\overline{V_{N,k}} = \overline{C_{N,k}} + \overline{C_{N,k}} \cdot \alpha = \overline{C_{N,k}} \cdot (1 + \alpha) \tag{2}$$

where $\overline{V_{N,k}}$ is the average pricing of a category on a given day and $\overline{C_{N,k}}$ is the average cost.

For each category of vegetable sales per day, it is defined as the sum of the sales per day of the individual vegetable items in that category, i.e., it conforms to the following equation (3):

$$\overline{Q_{N,k}} = \sum_{j \epsilon N} Q_{j,k} \tag{3}$$

For each category of vegetables the cost per day is given in equation (4) below:

$$\overline{C_{N,k}} = \overline{M_{N,k}} \cdot \overline{I_N} \tag{4}$$

Where, $\overline{M_{N,k}}$ is the cost unit price of a category on a particular day, and $\overline{I_{N,k}}$ is the replenishment quantity.

Since profit is the difference between sales and cost, in order to maximize the profit of the superstore, with reference to "Exploration of employee scheduling problem of greedy strategy based on predicting warehouse receiving and dispatching volume", a greedy strategy [4] is adopted so that the replenishment volume is minimized. After calculating the attrition rate, it happens to satisfy the customer's demand, i.e. the replenishment volume is defined as the following equation (5):

$$\overline{I_{N,k}} = \frac{\overline{Q_{N,k}}}{1 - \overline{L_N}} \tag{5}$$

Combining equations (8)-(11) yields the relationship between pricing and sales volume in equation (6) below:

$$\overline{V_{N,k}} = \frac{\overline{M_{N,k}} \cdot (1 + \alpha_{N,k})}{1 - \overline{L_N}} \cdot \overline{Q_{N,k}} \tag{6}$$

Where $\overline{Q_{N,k}}$ is the sales volume of a category on a particular day, and $\overline{L_N}$ is the average wastage rate. From the above formula, it is initially determined that the pricing is directly proportional to the sales volume.

The final expression for the cost margin using the cost-plus pricing method is given in equation (7) below:

$$\alpha_{N,k} = \frac{\overline{V_{N,k}} \cdot \left(1 - \overline{L_N}\right)}{\overline{M_{N,k}} \cdot \overline{Q_{N,k}}} - 1 \tag{7}$$

Then further analyze the relationship between sales volume and cost-plus pricing, i.e., to find the relationship between sales volume and cost margin. The formula for the specific parameter in equation (7) is as follows.

(1) The total pricing of each category on a given day is:

$$V_{N,k} = \sum_{j \epsilon N} V_{j,k} \tag{8}$$

(2) The average attrition rate for each category is:

$$\overline{L_N} = \frac{1}{|N|} \sum_{j \epsilon N} L_j \tag{9}$$

where $|N|$ denotes the number of item number sets in the N category.

(3) The average cost per unit for each category on a given day is:

$$\overline{M_{N,k}} = \frac{1}{|N|} \sum_{j \epsilon N} M_{j,k} \tag{10}$$

(4) The average sales volume for each category on a given day is:

$$\overline{Q_{N,k}} = \frac{1}{|N|} \sum_{j \epsilon N} Q_{j,k} \tag{11}$$

In addition, the expression for profit is as follows:

$$P_{N,k} = S_{N,k} - C_{N,k} \tag{12}$$

From eqs. (8)-(12), the calculations are carried out using C++, and finally the cost margins of six different categories are solved by Eq. (7), as shown in Table 1 below.

Table 1. Cost margins for each category for the next 7 days

Time	philodendron	cauliflower	Aquatic rhizomes	eggplant	capsicum	edible mushroom
2023/7/1	0.3809	0.5592	0.3153	0.7238	0.0405	0.1761
2023/7/2	0.3596	0.4634	0.3283	0.6902	0.0473	0.1520
2023/7/3	0.3540	0.5453	0.3299	0.6809	0.0538	0.1313
2023/7/4	0.3525	0.5097	0.3302	0.6720	0.0600	0.1249
2023/7/5	0.3521	0.5018	0.3303	0.6637	0.0658	0.1218
2023/7/6	0.3519	0.5222	0.3304	0.6557	0.0713	0.1204
2023/7/7	0.3518	0.5067	0.3306	0.6483	0.0765	0.1195

From the table above, we can get that the cost margin of each category fluctuates in the first few days, but in the end, they all tend to stabilize. It is assumed that the supermarkets do not have a clear understanding of the market demand at the beginning, and after a period of business, they have enough understanding of the stable price ecosystem of the market. The pricing will stabilize in a longer term, and the cost margin will fluctuate only when new factors flood in.

2.3. Daily replenishment and pricing decisions

In order to derive the replenishment volume and pricing strategy for the next 7 days (2023-7-1 to 2023-7-7), prediction is made based on the existing data, after which the relevant covariates in Eq. (6) can be solved to obtain the pricing decision.

As the sales volume of vegetable commodities and time often have a certain correlation, so the establishment of ARIMA model for prediction, although the sales volume of vegetable commodities into a certain seasonal pattern of change can be predicted using the SARIMA, but due to the need to give the last 7 days with the measurements, it is still the use of ARIMA model for prediction, the specific principles are as follows:

Based on the idea that the time series of the sales volume of vegetable commodities is generated by a certain stochastic process and that this stochastic process is not affected by time changes, it is believed that it is possible to accurately speculate the structure of this future stochastic process through the ARIMA model, using the observations of the existing time series. As a result, the predicted value of this time series can be expressed as a linear function of the current and lagged periods of the lag and random disturbance terms, i.e., its general expression is shown below:

$$Y_t = c + \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \ldots + \beta_q \epsilon_{t-q} \tag{13}$$

Based on the above concepts, the modeling process of the ARIMA model for this problem can be subdivided into the following three steps:

Step 1. Construct the model for the smooth time series:

Make full use of the PP or ADF test [10] to carry out the unit root test on the original time series of vegetable commodities, with reference to the journal in 2023 *Research on the optimal model for short-term CPI forecasting in China--Selection and optimization based on several time series models*. If the series does not meet the smoothness conditions, the non-smooth time series is transformed into a smooth one by means of differential transformation or logarithmic differential transformation, and then formally construct an ARIMA model for the smooth time series.

Step 2. Determine the model order:

Through the autocorrelation (AC) coefficient and partial autocorrelation (PAC) coefficient and other statistics that can describe the characteristics of the series, initially determine the possible forms of the model, and then refer to the AIC and other fixed-order criterion, choose the best to use from the possible forms of the model.

Step 3. Parameter Estimation and Diagnostic Tests:

Test the validity of the model itself, the significance of the model parameters, and whether the residual series is a white noise series, respectively. If the model passes the test, the model setting is considered to be basically correct; if it fails, go back to the second step to re-determine the form of the model and diagnostic test again until we get the form of the model with the correct setting, and finally get the general expression of this model.

The cost unit price, sales volume, and replenishment volume were predicted respectively, and the relevant data were imported into SPSSPRO, which finally resulted in the predicted volumes for the next 7 days for the three indicators as shown in Tables 2-Table 4 below.

Table 2. Unit cost for the next 7 days

Time	philodendron	cauliflower	Aquatic rhizomes	eggplant	capsicum	edible mushroom
2023/7/1	3.6869	7.8023	11.6559	4.4403	5.9467	4.6861
2023/7/2	3.6758	7.8769	11.5613	4.4355	5.9456	4.6923
2023/7/3	3.6784	7.8199	11.4400	4.4265	5.9446	4.6909
2023/7/4	3.6755	7.8642	11.3229	4.4271	5.9435	4.6884
2023/7/5	3.6748	7.8305	11.2100	4.4277	5.9425	4.6857
2023/7/6	3.6731	7.8569	11.1011	4.4283	5.9414	4.6830
2023/7/7	3.6719	7.8371	10.9959	4.4289	5.9404	4.6804

As can be seen from the above graph, the costs for the next 7 days are all leveling off, with little fluctuation in costs, with aquatic roots and tubers having the highest costs and leafy cauliflower having the lowest costs.

Table 3. Sales volume for the next 7 days

Time	philodendron	cauliflower	Aquatic rhizomes	eggplant	capsicum	edible mushroom
2023/7/1	125.3660	22.8639	20.1310	22.6240	83.8857	52.9877
2023/7/2	116.1004	20.4237	21.0748	20.8351	85.2910	42.0045
2023/7/3	118.8319	19.2753	21.4582	17.6597	85.7888	32.2762
2023/7/4	121.3844	18.72648	21.8317	17.3255	85.8894	43.5609
2023/7/5	125.1678	18.45604	22.1955	16.6678	85.8960	51.5693
2023/7/6	126.5064	18.31481	22.5499	18.4071	85.8979	44.4784
2023/7/7	125.4935	18.23356	22.8951	19.3969	85.9069	45.4050

As can be seen from the chart above, in addition to the eggplant and aquatic roots, the value is located near 20, the remaining four categories of vegetable sales volume is not the same, there will be a certain degree of fluctuation, but ultimately there is a tendency to flatten out. philodendron have the largest sales volume. It is presumed to be related to the number of sub-category items and people's daily demand for the amount of.

Table 4. Restocking volume for the next 7 days

Time	philodendron	cauliflower	aquatic rhizomes	egg plant	capsicum	edible mushroom	total
2023/7/1	138.25	26.10	22.54	24.23	91.03	57.30	359.45
2023/7/2	128.04	23.31	23.60	22.32	92.56	45.42	335.24
2023/7/3	131.05	22.00	24.03	18.92	93.10	34.90	323.99
2023/7/4	133.86	21.37	24.44	18.56	93.21	47.10	338.55
2023/7/5	138.04	21.07	24.85	17.85	93.21	55.76	350.78
2023/7/6	139.51	20.90	25.25	19.72	93.22	48.09	346.69
2023/7/7	138.39	20.81	25.64	20.78	93.23	49.10	347.94

As can be seen from the above figure, the trend of replenishment is consistent with the trend of sales volume, which can better match the amount of people's demand, minimize the cost and increase the revenue of the superstore, taking into account the wear and tear.

As a result, the calculation is carried out through equation (6) and finally the pricing for the next 7 days is obtained as shown in Table 5 below.

Table 5. Pricing for the next 7 days

Time	philodendron	cauliflower	aquatic rhizomes	eggplant	capsicum	edible mushroom
2023/7/1	5.09	12.17	15.33	7.65	6.19	5.51
2023/7/2	5.00	11.53	15.36	7.50	6.23	5.41
2023/7/3	4.98	12.08	15.21	7.44	6.27	5.31
2023/7/4	4.97	11.87	15.06	7.40	6.30	5.27
2023/7/5	4.97	11.76	14.91	7.37	6.33	5.26
2023/7/6	4.97	11.96	14.77	7.33	6.37	5.25
2023/7/7	4.96	11.81	14.63	7.30	6.40	5.24

As can be seen from the above graph, the pricing for the next 7 days tends to flatten out and fluctuate to a small degree for each category, with the highest pricing for the aquatic roots and tubers and the lowest pricing for the flowers and foliage category.

Calculations using equation (14) resulted in the final profit as shown in Table 6 below.

Table 6. Profit for the next 7 days

Time	Philodendron	Cauliflower	aquatic rhizomes	eggplant	capsicum	edible mushroom	total profit
2023/7/1	128.52	74.52	45.90	65.55	-22.34	23.52	315.67
2023/7/2	109.63	51.80	50.83	57.21	-19.21	13.95	264.21
2023/7/3	109.85	60.88	51.60	47.65	-15.96	7.58	261.60
2023/7/4	111.40	54.23	52.02	46.08	-12.88	8.91	259.76
2023/7/5	114.71	52.09	52.41	43.72	-9.95	9.76	262.75
2023/7/6	115.79	54.78	52.75	47.65	-7.10	8.15	272.01
2023/7/7	114.78	52.20	53.09	49.57	-4.42	8.13	273.36

As can be seen from the above table, profits will fluctuate at the beginning of the ups and downs of change, but ultimately tend to level off, the profits of the pepper category into an upward trend. Among them, the flower and leaf category has the highest profit, and the negative profit in the pepper category is thought to be due to its high degree of wear and tear, resulting in poor quality, the business discount processing.

3. Results

3.1. The establishment of model

When adopting cost-plus pricing to determine the relationship between sales volume and pricing, the key is to determine the cost margin, so this paper through the formulas converted to the four indicators of sales volume, depletion rate, cost, replenishment volume to determine the cost margin, and then determine the cost-plus pricing and the sales volume of the function of the equation. The value of replenishment volume is equated to the value of sales volume before depletion to satisfy the optimal pricing decision in the maximum profit, in order to give the pricing decision for the next 7 days, ARIMA is used to predict the cost margin, cost per unit, and sales volume, which is substituted into the functional relationship between them, and ultimately gives the optimal pricing decision and daily replenishment volume. The model was then tested and evaluated by directly fitting the relationship between them using least squares to re-measure the pricing decision.

3.2. Model Analysis and Evaluation

The profitability is given by analyzing the link between the relational equations, and in this section the regression prediction is done directly using the least squares method, given the relationship, and the principle is as follows:

Linear regression is expressed in the form:

$$y = w'x + e \tag{14}$$

where e is the error obeys a normal distribution with mean zero.

In general, there is often more than one factor affecting y. Assuming that there are a total of k factors $x_1, x_2, \ldots, x_k$, a linear relationship can usually be considered as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \tag{15}$$

Then make n independent observations of y and $x_1, x_2, \ldots, x_k$ at the same time to obtain n sets of observations $(x_{t1}, x_{t2}, \ldots, x_{tk})$, $t = 1, 2, \ldots, n\,(n > k+1)$, which n sets of observations satisfy the following equation:

$$y = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \ldots + \beta_k x_{tk} + \varepsilon_t \tag{16}$$

where $\varepsilon_1 \ldots \varepsilon_n$ are uncorrelated and ε are random variables with the same distribution. To represent the above equation in terms of a matrix, let:

$$Y = X\beta + \varepsilon \tag{17}$$

Use the least squares method to obtain the solution of β.

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{18}$$

where $(X^T X)^{-1} X^T$ is called the pseudo-inverse of X.

Then the slope b is calculated using the following equation:

$$u(b) = \frac{\sigma}{\sqrt{\sum (x - \bar{x})^2}} \tag{19}$$

Then calculate the intercept a:

$$u(a) = u(b) \cdot \sqrt{\frac{\sum x_i^2}{n}} \tag{20}$$

The resulting fitted relationship between sales volume and cost pricing for each category is shown in Table 7 below.

Table 7. Fitting the relationship between sales volume and pricing based on the least squares method

	Sales volume and cost pricing equation
philodendron	$V = 6.553 - 0.004Q$
cauliflower	$V = 10.623 - 0.031Q$
aquatic rhizomes	$V = 11.294 - 0.042Q$
eggplant	$V = 8.557 - 0.001Q$
capsicum	$V = 9.647 - 0.012Q$
edible mushroom	$V = 9.482 - 0.013Q$

This will be calculated by following the steps in Section 7.3, resulting in the final pricing for the next 7 days under the least squares fit as shown in Table 8 below.

Table 8. Pricing for the next 7 days by category

Time	philodendron	cauliflower	aquatic rhizomes	eggplant	capsicum	edible mushroom
2023/7/1	5.85	9.91	10.45	8.53	8.67	8.79
2023/7/2	5.89	9.99	10.41	8.54	8.65	8.94
2023/7/3	5.88	10.03	10.39	8.54	8.65	9.06
2023/7/4	5.87	10.04	10.38	8.54	8.64	8.92
2023/7/5	5.85	10.05	10.36	8.54	8.64	8.81
2023/7/6	5.85	10.06	10.35	8.54	8.64	8.90
2023/7/7	5.85	10.06	10.33	8.54	8.64	8.89

As can be seen from the above figure, the pricing fitted by the least squares method is almost constant without any change, with the lowest pricing for the floral and foliage category, and the highest pricing for the aquatic roots and tubers category.

Calculations using equation (20) resulted in the final profit as shown in Table 9 below.

Table 9. Profit for the next 7 days by category

Time	Philodendron	Cauliflower	aquatic rhizomes	eggplant	capsicum	edible mushroom	total profit
2023/7/1	223.86	23.06	-52.39	85.47	185.72	197.43	663.16
2023/7/2	213.03	20.41	-53.45	78.86	187.50	162.23	608.56
2023/7/3	216.40	21.20	-51.86	67.07	188.18	128.79	569.78
2023/7/4	220.21	19.97	-50.24	65.79	188.39	167.54	611.67
2023/7/5	225.28	20.54	-48.61	63.29	188.50	193.12	642.13
2023/7/6	227.24	19.92	-46.97	69.86	188.60	170.80	629.43
2023/7/7	226.10	20.29	-45.33	73.58	188.71	173.94	637.28

From Table 9 above, it is found that the daily profits of its aquatic root vegetables are all negative, which is not in line with the practical significance; comparing it with the final pricing decision in section 2.3, it is found that there is a large difference in each of the predicted values. Comparing it with the existing 36-months sales data, it is found that its total daily profits are much higher than the average daily profits of the past three years. Therefore, it was initially determined that the regression prediction method utilizing the least squares method was not reliable.

Subsequently, in order to prove the unreliability of the regression prediction of the least squares method, the perturbation of the relationship between sales volume and cost-plus pricing for each category was measured as shown in Table 10 below.

Table 10. Least Squares Prediction of Perturbation of Fit

	philodendron	cauliflower	aquatic rhizomes	eggplant	capsicum	edible mushroom
R^2	0.081	0.067	0.031	0.054	0.139	0

As can be seen from Table 10 above, the perturbation of the relationship between sales volume and cost-plus pricing for each category is extremely low, so it is considered that there is irrationality in the fitting using the least squares method, and the pricing decision proposed by this model better meets the requirements of the context of the question and has a strong reliability.

4. Conclusions

In order to maximize profits, vegetable superstores often adopt ways such as restocking in response to changes in market demand and lowering pricing at the right time, while the daily restocking volume and pricing of each vegetable category have a key impact on superstore revenues. In this paper, based on the data of all past trading days, we build a forecasting and decision-making model based on the ARIMA model and greedy strategy, and use the least squares method to fit the relationship between sales volume and pricing to provide vegetable superstores with optimal suggestions for replenishment and pricing decisions. The experimental results show that the model has good predictability and robustness, and has certain practical application value. For example, the model can better solve the pricing problem of commodities related to time series, which can be used in supply chain product management system, online vegetable commodity distribution platform and so on. Of course, it can also be used in models that contain the interaction between variables in both time and space dimensions, such as the prediction of traffic flow and pedestrian flow.

It should not be overlooked that there is also room for optimization in this model. We can apply time series forecasting. The distribution pattern of different categories or different single products is given in the form of time series, and the interrelationship of different categories or different single products is given in the form of spatial series. Meanwhile, cluster analysis is also a good medium. Based on the standard deviation, coefficient of variation and other statistical indicators, 251 vegetable items were clustered, and the feature centers were selected to represent the relevant numerical characteristics of the items similar to them.

References

[1] Wang Shuhua, Li Binghai, Xue Zhongyuan, et al. Cold chain logistics + mechanical cold storage to meet the fresh supply of fruits and vegetables in the market [J]. New Agriculture, 2023, (24): 54-56.

[2] Zhang Shuyan, Chen Qibing, Cai Yijie. Research on river water quality prediction by ARIMA model based on exponential smoothing [J]. Guangdong Chemical Industry, 2024, 51(06): 95-98.

[3] Zhang He. Vegetable sales prediction based on integrated learning[D]. Yunnan Normal University,2021.

[4] Li Xiaoli, Gao Tianhao, Peng Jiali. Exploration of employee scheduling problem of greedy strategy based on predicting warehouse receiving and dispatching volume [J]. China Storage and Transportation, 2022, (12): 92-94.

[5] Gu Sihong. A study on dynamic pricing of fresh products in H-retailers considering freshness variation[D]. Donghua University, 2023.

[6] Liu Hongyu. Research on Dynamic Inventory Replenishment and Delivery Modeling and Solving for Enterprises Based on PI [D]. Northeastern University, 2020.

[7] Zhang Xiaoyong. Research on the correction algorithm of a certain type of programmable resistance box based on the least squares method [J]. Heilongjiang Science, 2023, 14(24): 22-25.

[8] Liu Li, Hao Xianfeng, Wang Yudong. Stock market return prediction based on time-varying robust weighted least squares [J]. Journal of Management Science, 2024, 27(01): 141-158.

[9] Zhou Xiangyu, Li Si. Short-term forecast of express business volume in Jiangsu Province based on TOPSIS criterion and SARIMA model [J]. Science and Industry, 2023, 23(17): 136-142.

[10] Chen Lin, Xie Jiachun, Ma Liangli. Research on the optimal model for short-term CPI forecasting in China--Selection and optimization based on several time series models [J]. Contemporary Economy, 2022, 39(05): 9-16.