Research on vegetable pricing and replenishment strategy based on genetic algorithm

Minjie Liang[1,*], Yufei Xie[1] and Shiyu Xie[1]

[1]School of Ocean University Of China, Qindao, China

*Corresponding author: 2930976998@qq.com

Abstract. In this paper, the average sales volume and cost plus pricing of 6 categories of vegetables are calculated. The regression analysis is carried out by using R language. Then, with profit maximization as the objective function and annual sales volume, replenishment volume, unit price of sales and loss rate as the constraint conditions, a quadratic programming model is established. The total replenishment of flowers and leaves, cauliflower, aquatic rhizome, capsicum, nightshade and edible fungi from July 1 to 7, 2023 was 111.27, 22.82, 21.24, 46.88, 14.23 and 40.40 kg, respectively. The expected pricing is 10.28, 12.92, 11.77, 20.56, 12.96 and 11.91 yuan/kg, with a total profit of 1843.8 yuan. On this basis, with the number of orders and sales pricing of merchantable items as decision variables, an integer programming model was established according to the constraints such as the number of items, minimum display quantity, price range and loss rate, etc. Genetic algorithm was used to solve the integer programming model, and the replenishment volume and pricing strategy of 35 items were obtained on July 1, with an estimated profit of 2398.2 yuan.

Keywords: Regression analysis; Quadratic programming; Integer programming; Genetic algorithm.

1. Introduction

The marketable production of vegetable commodities is an important part of the structural reform of the agricultural supply side, which plays a decisive role in ensuring farmers' "money bag" and supplying citizens with "vegetable basket"[1]. However, under the influence of factors such as the change of diversified demand in the consumer market, the vegetable industry faces problems such as poor connection between production and marketing and poor comprehensive benefits. In order to accelerate the effective docking of the market[2], promote the improvement of industrial quality and efficiency, improve the comprehensive competitiveness of the vegetable industry and the level of sustainable and healthy development, and promote the structural reform of the agricultural supply side, the real and effective market demand analysis is particularly important[3].

For fresh supermarket, vegetable products have short shelf life, short purchasing radius and low scale; High loss of goods transportation, unstable quality, high risk; There are many varieties of vegetables, scattered origin and difficult quality standardization[4]. At the same time, the production season is strong, the supply and demand relationship is unstable, and the price is volatile. The purchase transaction time of vegetables is usually before determining the item and purchase price. Business super can only be replenished according to the past sales of goods[5], and then adopt the method of "cost plus pricing", and discount the goods with poor transport damage and product phase change. Therefore, the reasonable pricing method and replenishment decision need the support of reliable market demand analysis to ensure the profit of the supermarket[6].

2. Methodology

2.1. Calculation of sales volume and cost plus pricing for each vegetable category

2.1.1 Average daily sales

According to the total sales volume of vegetable products in supermarket, we choose the daily average sales volume as the calculation index. Because the average total sales volume is based on historical data, it can effectively avoid seasonal and active sales fluctuations, provide more stable and reliable sales information, and enable the supermarket to better plan replenishment plans, inventory management, pricing strategies and supply chain strategies, thereby improving operational efficiency and profitability. Its calculation formula is as follows:

$$\mathrm{Q_i} = \frac{\sum_{j=1}^{n} \mathrm{Q_j}}{\mathrm{n}} \tag{1}$$

2.1.2 Monthly cost-plus pricing

Cost plus pricing is based on the unit cost of vegetable goods plus a certain percentage of profit to establish the sales unit price method, most of the supermarket will be based on the cost profit rate to determine the size of the added profit. Its calculation formula is as follows:

$$\mathrm{CP_i} = \mathrm{C_i}(1 + \mathrm{CW_i}) \tag{2}$$

$$\mathrm{CW_i} = \frac{\mathrm{W_{total}}}{\mathrm{C_{total}}} = \frac{\sum_{j=1}^{n} \mathrm{Q_j P_j} - \sum_{j=1}^{n} \mathrm{p_j}(1 + \mathrm{L_j})}{\sum_{j=1}^{n} \mathrm{q_j p_j}} \tag{3}$$

Where is the cost plus pricing, is the cost profit rate, is the commodity loss rate.

Based on the above calculation formula, the average sales volume and cost plus pricing are obtained to provide data support for subsequent regression analysis and planning model.

2.2. Regression analysis of total sales volume and cost plus pricing of each vegetable category

2.2.1 Unitary regression analysis

Let X be the cost plus pricing and Y the average sales volume. Based on the above data processing, a regression model is established:

$$\mathrm{Y_i} = \mathrm{a} + \mathrm{b} * \mathrm{X_i} + \varepsilon_\mathrm{i} \tag{4}$$

The expression of the least square solution is:

$$\hat{b} = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sum_i (x_i - \overline{x})^2} = r_{xy} \frac{S_y}{S_x} \tag{5}$$

$$\hat{a} = \overline{y} - \hat{b}\overline{x} \tag{6}$$

Where is the sample correlation coefficient of X and Y, and the sample standard deviation of X and Y, respectively.

2.3. Secondary planning of sales volume and cost plus pricing of each vegetable category

2.3.1 Model building

(a) Decision variables: sales unit price and daily replenishment volume, where i represents different types.

(b) Objective function: to maximize the total profit of the quotient, that is:

$$\max\ W_i = \sum_{i=1}^{n} (P_i Q_i - p_i q_i) \tag{7}$$

Where, is the total profit, is the average sales volume, is the daily replenishment volume. Is the wholesale price, taking the weighted average of the unit cost over the years. Is the unit price of sales, which can be derived from the linear regression equation of total sales and cost plus pricing, namely:

$$\mathrm{p_i} = \frac{\mathrm{Q_i} - \mathrm{a_i}}{\mathrm{b_i}} \tag{8}$$

(c) Constraints:

Demand side constraint: historical minimum daily loss ≤ daily loss ≤ historical maximum daily loss, that is:

$$L_j = q_j - Q_j, \tag{9}$$

$$L_j \geq minL_j \geq 0 \tag{10}$$

$$minL_j \leq L_i \leq maxL_j \tag{11}$$

Among them, it is the historical daily loss, and the replenishment and sales volume of various vegetables over the years.

Supply side constraints: the quantity of replenishment is limited by the maximum value of replenishment in June, July and August of previous years, namely:

$$\min q_{678} \leq q_i \leq \max q_{678} \tag{12}$$

Price range constraint: the maximum and minimum values of the price, namely:

$$minP_j \leq P_i \leq maxP_j \tag{13}$$

Consumption constraint: Control vegetable loss rate based on historical data

2.4. Integer planning of order quantity and pricing of each vegetable item

2.4.1 Model building

a. Decision variables: pricing and order volume, where i represents different items.

b. Objective function: to maximize the total profit of the quotient, that is:

$$\max\ W_i = \sum_{i=1}^{n} (P_i Q_i - p_i q_i) \tag{14}$$

Where, is the total profit, is the sales pricing, is the average sales volume, is the order volume, is the wholesale price:

$$p_i = \frac{\sum_{j=1}^{t_i} p_{i,j}}{t_i} \tag{15}$$

Among them, the unit price of a single purchase item.

c. Constraints:

Demand-side constraint: The total number of salable items is controlled at 27-33, that is:

$$27 \leq i \leq 3 \tag{16}$$

Supply side constraint: the minimum display quantity (order quantity of each item) is greater than 2.5kg, that is:

$$2.5 \leq q_i \tag{17}$$

Selling price range constraint: maximum and minimum values of the 7-day selling price, i.e. :

$$minP_j \leq P_i \leq maxP_j \tag{18}$$

3. Results

3.1. Regression result

Based on the calculation data of average sales volume and cost plus pricing, regression model was established by using R language and regression analysis was conducted. The fitting results obtained were shown in the following table, in which R-squared, F-statistic and p-value values all met expectations with good fitting degree and high significance level.

Table 1. Regression analysis equations for each category

Vegetable category fitting equation	Fitting equation
Flowers and leaves	Y_1=258.12-15.54x
Cauliflower	Y_2=67.152-3.439x
Aquatic roots	Y_3=83.304-5.336x
Chili	Y_4=108.357-3.058x
Nightshade	Y_5=32.5063-1.4386x
Edible fungi	Y_6=125.346-7.183x

Using the above results, the regression line scatter plots of different categories are drawn, including regression lines fitted from the scatter data. We can see that the cost plus pricing and the average sales volume show a clear negative correlation.

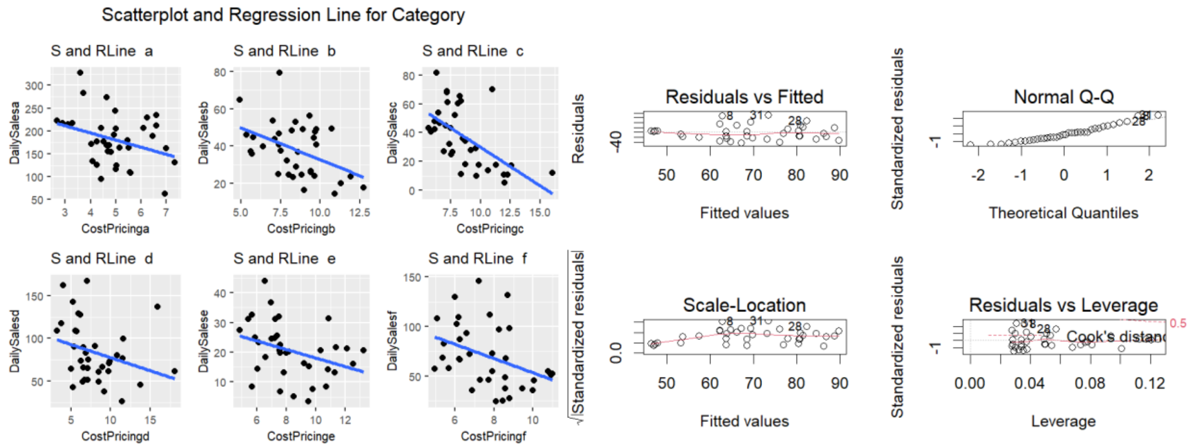

Figure 1. Scatter plot of regression analysis &Residual diagram of regression analysis

At the same time, except for edible fungi, the points on the scatter plot are mostly clustered below the straight line, that is, with the increase of cost pricing, the average sales volume gradually decreases, which is consistent with the negative slope of the linear regression model. In addition, there are points on each graph that are far away from the straight line, and the existence of these points may affect the fitting effect of the regression model, requiring further outlier analysis or data processing.

Taking the Mosaic class as an example, we can also draw regression analysis residual plot, Q-Q plot, etc., to verify its linearity, normality, heteroscedasticity and independence, and observe the degree of model fitting. In the figure 1, the residuals are randomly distributed around the zero line, and there is no obvious pattern, which conforms to the linear hypothesis. The distribution of points in the normal graph is near the straight line, which is approximately normal distribution and meets the expectation.

3.2. Results of quadratic programming model

According to the above model, the corresponding quadratic programming model is established in MATLAB and solved to obtain the maximum profit. Then, the regression analysis expression is brought in to obtain the total replenishment and pricing strategy of each vegetable category from July 1 to 7, 2023. The specific results are shown in the following table.

Table 2. Secondary planning results of each category

Vegetable category	Replenishment volume	Estimated sales volume	k	b	Estimated price	Attrition rate (%)
Flowers and leaves	111.27	99.75	-15.40	258.00	10.28	10.35
Cauliflower	22.82	22.71	-3.44	67.15	12.92	0.47
Aquatic roots	21.24	20.52	-5.34	83.30	11.77	3.40
Chili	46.88	45.47	-3.06	108.36	20.56	3.00
Nightshade	14.23	13.90	-1.44	32.51	12.96	2.34
Edible fungi	40.40	39.83	-7.18	125.35	11.91	1.42
Profit: 1,843.8 yuan						

That is, the total replenishment of flowers and leaves, cauliflower, aquatic roots, capsicum, nightshade and edible fungi from July 1 to 7, 2023 was 111.27, 22.82, 21.24, 46.88, 14.23 and 40.40 kg, respectively. The expected pricing is 10.28, 12.92, 11.77, 20.56, 12.96 and 11.91 yuan/kg, with a total profit of 1843.8 yuan.

3.3. The result of integer programming model based on genetic algorithm

The integer programming model is composed of 32 multivariable programming models, and its planning model is affected by the replenishment strategy of the previous day. Therefore, when solving the automatic pricing and replenishment schemes, the model needs to be solved day by day, and the replenishment scheme of the t-1 day should be obtained first as the basis for the solution of the T-day. We choose genetic algorithm to solve the above model and find the replenishment quantity and pricing strategy. Genetic algorithms iteratively optimize these solutions, constantly improving and finding a near-optimal solution. This method is suitable for complex oversupply and pricing problems which need to consider many factors and constraints.

The corresponding integer programming model is established in MATLAB, and the genetic algorithm is used to solve it, and the maximum profit, the corresponding total quantity of replenishment and the pricing strategy are obtained. The replenishment quantity and pricing strategy of 35 products on July 1 are obtained, and the expected profit is 2398.2 yuan.

4. Conclusion

In the establishment of automatic pricing and replenishment decision-making system, we have carried out sufficient calculation and demonstration, selected reasonable and effective available indicators, measured many indicators, involving a wide range of aspects, can be comprehensively and effectively forecast. Using reasonable and effective planning model, the model has sufficient theoretical basis and can solve practical problems. The problem solutions are interlinked and hierarchical. Based on the previous problems, the index system reflecting the market demand is obtained through optimization and improvement.

References

[1] Xu Zhao Zhou. Research on product pricing/quality decision of supply chain considering consumer reviews from the perspective of BOPS [D]. Guangdong industry, 2022.

[2] Li Yali. Joint replenishment and pricing coordination decision considering fresh-keeping technology investment [D]. Chongqing Jiaotong University Science, 2022.

[3] Zhang Jinlong, WU Xiang, XU Haoxuan. Perishable products pricing and replenishment joint decision model [J]. Journal of systems engineering, 2018 (01) : 79-89.

[4] Wang Yuanlan. Demand forecast of automatic rice vending machine considering replenishment and shortage [D]. Nanjing University Science, 2021.

[5] Zhao X.Research on pricing and preservation decision of fresh agricultural products dual-channel supply chain [D]. Southeast University Science, 2021.

[6] Wan Guangyu, Li Qingsong, Cao Yu. Based on the service level of price discount and retailers perishable product pricing research [J]. Journal of financial theory and practice, 2019, 40 (5) : 142-147.