ISSN 2522-3488 Vol. 9, Issue 1: 57-64, DOI: 10.25236/IJNDES.2025.090108

Pricing and replenishment decision-making of vegetable products based on time series analysis

Liwen Chen, Linghui Bu, Qiang Chen*

College of Mathematics and Physics, Yancheng Institute of Technology, Yancheng, Jiangsu, China
**Corresponding author: chq@ycit.cn*

Abstract: *Due to the short shelf life of vegetable products, their pricing and replenishment decisions are receiving increasing attention from supermarkets. This article first uses the Prophet model to remove time effects and provides the distribution pattern of vegetable product sales. Based on this, a double logarithmic model of vegetable product sales and cost markup pricing is established. LSTM is used to predict the cost price of vegetable products. Then, replenishment quantity and markup rate are used as decision variables, and pricing satisfying the cost markup formula is used as a constraint condition. The goal is to maximize supermarket profits, and a single objective optimization model is established. Finally, genetic algorithm is applied to solve the pricing and replenishment of vegetable products, providing optimal decisions for vegetable market management.*

Keywords: *Prophet model, Double logarithmic model, LSTM, Genetic algorithm*

1. Introduction

The fresh food supermarket industry is facing the challenges of short vegetable shelf life and rapid decline in product quality, which requires businesses to make accurate replenishment decisions in the uncertain situation of individual products and prices. The traditional cost plus pricing method and discount sales strategy are no longer able to meet market demand. Therefore, it is particularly important to develop a replenishment and pricing model [1-2] based on accurate market demand forecasting. This model can help supermarkets optimize replenishment quantities, reduce inventory backlog, improve capital turnover and operational efficiency. At the same time, it can also provide flexible pricing strategies to adapt to market demand and changes in product appearance, thereby increasing sales volume and profit margins. Using the data from Question C of the 2023 National College Student Mathematical Modeling Competition, this article aims to establish the relationship between sales volume, price, and cost, providing scientific basis for supermarkets to formulate vegetable sales combinations, pricing, and replenishment strategies for the next seven days, and enhancing their market competitiveness.

2. Analysis of Sales Patterns Based on Prophet

Due to the significant absolute differences in sales between different categories and individual products, the monthly sales ratio is calculated based on the sales ratio within a certain time range. For example, when calculating the monthly sales pattern, the monthly sales of each category or individual product are divided by the total sales to obtain the monthly sales ratio. The same applies to other time dimensions, facilitating comparisons between different categories or items.

From July 2020 to June 30, 2023, the daily sales volume of six types of vegetables showed clear seasonal fluctuations with peaks and valleys each year. The sales volume of aquatic rhizome vegetables is higher in August December and lower in other months, which may be related to the fact that the maturity period is mostly in August October, or the "Three Treasures of Autumn" such as water chestnut on the Mid-Autumn Festival; Most leafy vegetables have higher sales from July to October, but lower sales from February to June and November to December, possibly due to their fruiting period from July to September; Flower vegetables are relatively popular from July to November, while sales in other months are relatively sluggish, possibly due to their harvest period from July to November; Eggplant vegetables are in their peak sales season from May to August, while the rest of the months are in their off-season; The sales of chili vegetables are mostly concentrated from January to March and August to October, which may be affected by holidays such as Spring Festival and National Day; Edible mushrooms are in the peak sales season from January to February and October to December,

ISSN 2522-3488 Vol. 9, Issue 1: 57-64, DOI: 10.25236/IJNDES.2025.090108

possibly due to their maturity period in spring and autumn. Overall, people's demand for different types of vegetables varies with seasons and holidays, and this fluctuation has a certain impact on vegetable sales volume.

The Prophet model [3] is a process of automatically predicting time series and future trends based on additive models, suitable for non-stationary series with significant seasonal and holiday effects. This article uses the Prophet model with exogenous regression terms to fit the trend, periodicity, and holiday components. The model is as follows:

$$N_i(t)=g_i(t)+s_i(t)+h_i(t)+f\left(P_i(t)\right)+\varepsilon_i(t), \tag{1}$$

In the formula, $N_i(t)$ represents the sales volume of each vegetable category or item, $g_i(t)$ represents the trend term, $s_i(t)$ represents the season term, $h_i(t)$ represents the holiday term, $f\left(P_i(t)\right)$ represents the exogenous regression term, and $\varepsilon_i(t)$ represents the noise term.

The trend term $g(t)$ is used to describe the non periodic changes of the time series g column, including saturation growth models and linear growth models based on logistic regression functions. The season term $s(t)$ uses Fourier order a numbers to simulate the periodicity of time series, $s(t)=\sum_{n=1}^{N}\left(a_n\cos\frac{2\pi nt}{T}+b_n\sin\frac{2\pi nt}{T}\right)$. The holiday term is used to describe the impact of special periods such as holidays and major events on time series, $h(t)=\sum_{i=1}^{K}k_i\bullet I\{t\in H_i\}$. The exogenous regression term $f\left(P_i(t)\right)$ represents the impact function of pricing determined based on the cost plus method on vegetable sales volume.

By training the data, the weekly effect, annual effect, holiday effect, and trend effect of sales time series can be extracted. Taking aquatic rhizomes as an example, their temporal decomposition diagram is as follows:

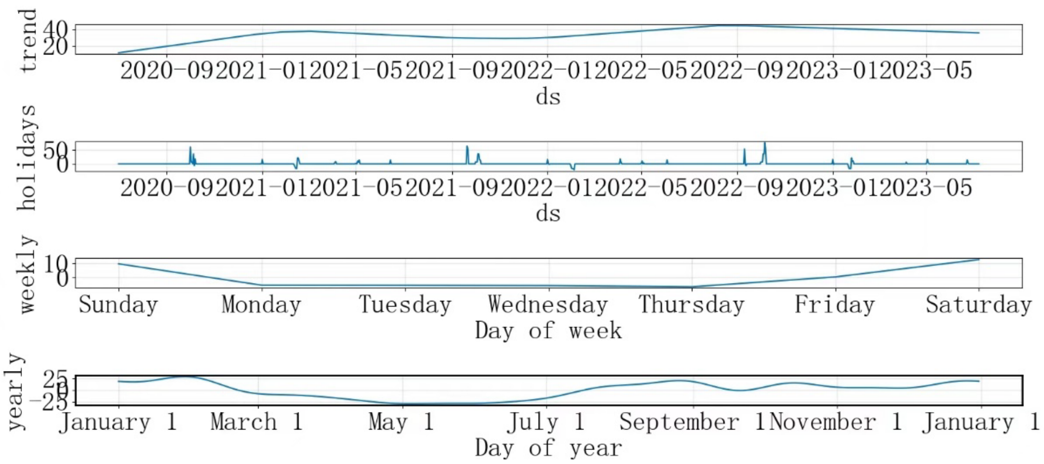

Figure 1. Time series decomposition diagram of aquatic rhizomes

According to Figure 1, there were two peaks in the sales of aquatic rhizomes in March 2021 and September 2022, but overall there was no significant growth trend. During the Mid-Autumn Festival and National Day holidays every year, the consumption of aquatic root vegetables will experience a significant increase, with sales significantly higher than during other holidays. During short holidays such as Qingming and Spring Festival, there may also be some small-scale consumption peaks. Vegetable consumption on weekends is usually significantly higher than on weekdays, which may be related to frequent family meals and weekend cooking activities. It is worth noting that January ,February, September and October are the peak sales periods for aquatic root vegetables, and the sales performance during this period is particularly outstanding, reflecting people's preference and demand for these vegetables in autumn and winter seasons. Overall, the consumption patterns during holidays and weekends, as well as seasonal fluctuations, profoundly affect the market volatility of aquatic root vegetables.

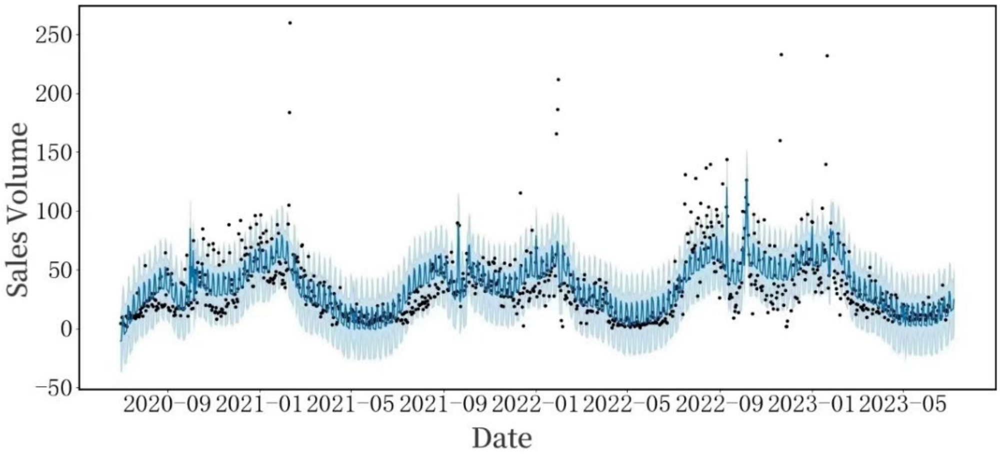

Figure 2. Prophet regression plot of aquatic rhizomes

From Figure 2, it can be seen that the Prophet model eliminates the influence of outliers and missing values. The regression results show that aquatic rhizomes and edible mushroom vegetables have higher sales volume in January February and October December each year, and lower sales volume in May July. The sales volume of cauliflower and leafy vegetables is relatively high in September each year, but the overall sales volume in 2022 is lower than other years. Chili vegetables experience a sales peak in September October and January February each year, indicating a high demand from customers during holidays such as National Day and Spring Festival. Eggplant vegetables have the highest sales volume from June to August each year, with significant seasonal characteristics.

3. Pricing and replenishment decisions for the vegetable categories

Firstly, it is necessary to solve the exogenous regression term in the Prophet model. By establishing a double logarithmic model, the impact function $f\left(P_i\left(t\right)\right)$ of cost markup pricing on vegetable sales volume can be obtained.

3.1 Double logarithmic model of category sales

Supermarket sales use the cost plus method as a pricing strategy, which sets the price of vegetable products by adding a certain percentage of profit to the unit cost. The cost plus pricing can be calculated using the following formula:

$$P = C \bullet \left(1+q\right), \tag{2}$$

Among them, P represents cost markup pricing, C represents unit cost, and q represents markup rate.

(1) Accounting for vegetable cost prices

For each individual vegetable, its cost is:

$$C_{i_j}\left(t\right) = \frac{U_{i_j}\left(t\right)}{1-r_{i_j}}, \tag{3}$$

Among them, i_j represents the j-th vegetable item under the i-th category, $C_{i_j}\left(t\right)$ represents the cost price of the i_j-th vegetable item on the t-th day, $U_{i_j}\left(t\right)$ represents the wholesale price of the i_j-th vegetable item on the t-th day, and r_{i_j} represents the loss rate of the i_j-th vegetable item.

The cost $C_i(t)$ of each vegetable category can be expressed as the average cost per item.

(2) Calculation of vegetable markup rate

The formula for calculating the markup rate of each vegetable item is as follows:

$$q_{i_j}(t)=\frac{V_{i_j}(t)-U_{i_j}(t)}{U_{i_j}(t)}\times 100\%, \tag{4}$$

Among them, $q_{i_j}(t)$ represents the markup rate of the i_j-th vegetable item on the t-th day, $V_{i_j}(t)$ represents the sales price of the i_j-th vegetable item on the t-th day, and $U_{i_j}(t)$ represents the wholesale price of the i_j-th vegetable item on the t-th day.

The markup rate $q_i(t)$ of each vegetable category can be expressed as the average markup rate of a single product.

To analyze the relationship between the total sales volume of various vegetable categories and cost markup pricing, this article takes the total sales volume of each vegetable category as the dependent variable and the daily average cost markup pricing of each category as the explanatory variable, and combines the Prophet model to establish a double logarithmic model [4]:

$$\ln f\left(P_i(t)\right)=\alpha_i+\beta_i\ln P_i(t)+\gamma_i\ln\left(\sum_j\sum_t f\left(P_i(t)\right)\bullet P_{i_j}(t)\right)+\varepsilon_i, \tag{5}$$

Among them, t represents the number of days of sales, $N_i(t)$ represents the total sales of the vegetable category on the t-th day, $P_i(t)$ represents the cost markup pricing of the i-th vegetable category on the t-th day, α_i represents the constant parameter of the i-th vegetable category, β_i represents the demand price elasticity of the i-th vegetable category, γ_i represents the expenditure elasticity of the i-th vegetable category, $P_{i_j}(t)$ represents the total sales of the j-th vegetable item under the i-th category on the t-th day, ε_i represents the error term of the i-th vegetable category, $f\left(P_i(t)\right)$ represents the total sales volume of the i-th vegetable category after excluding time effects.

3.2 Cost prediction of vegetable categories based on LSTM

In order to facilitate the establishment of replenishment and pricing decision-making models for various vegetable categories in the future, the cost prices of various vegetable categories from July 1-7, 2023 were first predicted using time series, and these data were used as known parameters for the model.

LSTM [5] is a variant of recursive neural network used for processing and predicting sequential data. Using the LSTM model to predict costs, the cost price predictions for six vegetable categories for the next week are shown in Table 1.

Table 1. Cost Forecast Data

Categories	07-01	07-02	07-03	07-04	07-05	07-06	07-07
Aquatic rhizomes	13.89	13.91	13.69	13.48	12.93	12.71	24.64
Flowers and leaves	3.81	3.75	3.73	3.68	3.67	3.68	3.66
Florescent vegetables	8.26	8.21	8.15	8.06	7.98	7.91	7.86
Eggplant type	4.99	5.15	5.14	5.08	5.13	5.14	5.16
Chili peppers	3.81	3.94	4.02	4.06	4.08	4.08	4.12
Edible fungi	4.38	4.37	4.40	4.46	4.56	4.71	4.83

Taking eggplant vegetables as an example, the actual cost values of eggplant vegetables are generally consistent with the predicted values, with a high goodness of fit of 0.87 and average absolute error and mean square error of only 0.46 and 0.35, respectively. It is considered that the fitting effect is good. The goodness of fit of LSTM models for other vegetable categories is greater than 0.6, indicating good performance. Among them, the goodness of fit of cauliflower is as high as 0.86, which is comparable to that of eggplant; The goodness of fit of edible mushrooms is the lowest, at 0.63, but the fitting effect is relatively good. The cost curve predicted by LSTM is shown in Figure 3.

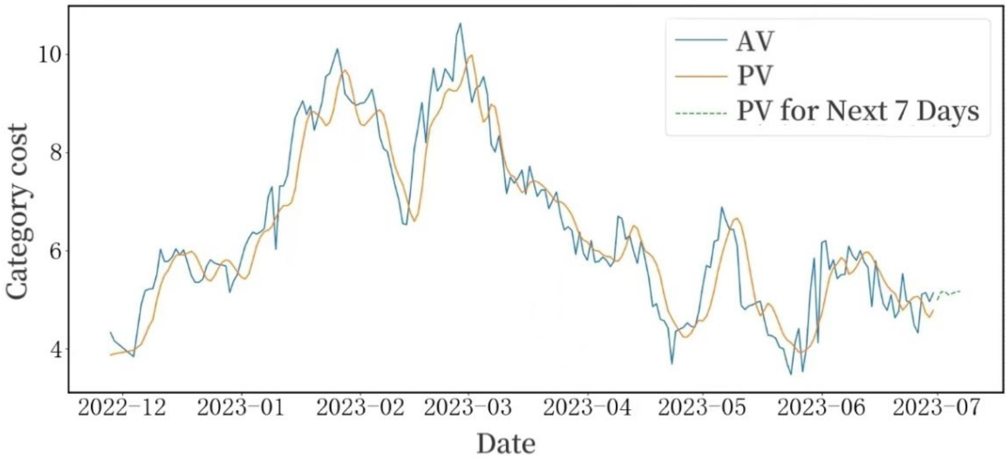

Figure 3. LSTM prediction results for eggplant species

3.3 Category Pricing and Replenishment Model

The vegetable sales on that day can be divided into two parts: one is sold directly based on pricing, and the other is sold at a discount due to uncontrollable factors such as transportation and spoilage. The discount rate for each vegetable category can be obtained by weighted averaging the ratio of the selling prices of similar items before and after discounts throughout the day.

According to the reality, the ever-changing market demand faced by supermarkets can be represented by sales volume $N_i(t)$. The profit of the supermarket on the t-th day can be expressed as:

$$W(t)=\sum_{i=1}^{6}\left[P_i(t)-C_i(t)\right]\bullet\min\left\{M_i(t)\bullet(1-r_i),N_i(t)\right\}+\left[d_i\bullet P_i(t)-C_i(t)\right]\bullet N_i(t)\bullet r_i, \quad (6)$$

This article takes the replenishment quantity $M_i(t)$ and price markup rate $q_i(t)$ of various vegetable categories as decision variables to achieve the maximization of profits for supermarkets.

The objective function is to maximize the benefits of supermarkets, which can be expressed as:

$$\max\sum_{i=1}^{6}\left\{\left[P_i(t)-C_i(t)\right]\bullet L_i(t)+\left[d_i\bullet P_i(t)-C_i(t)\right]\bullet\left[M_i(t)-L_i(t)\right]\right\}, \quad (7)$$

Among them, $L_i(t)=\min\left\{M_i(t)\bullet(1-r_i),N_i(t)\right\}$ represents the sales volume of the i-th item without discount on the t-th day.

The constraint conditions are as follows:

(1)The daily replenishment volume of each vegetable category shall not be less than the daily sales volume, but shall not exceed the maximum daily sales volume of the previous three years.

$$N_i(t)\le M_i(t)\le N_{i\max}, i=1,2,\cdots,6, \quad (8)$$

(2)To ensure the revenue of supermarkets, the pricing of each vegetable category must not be lower than 110% of the cost, and to prevent excessive pricing from causing unsold vegetables, the pricing must not exceed 150% of the cost.

$$110\%\times C_i(t)\le P_i(t)\le 150\%\times C_i(t), i=1,2,\cdots,6, \quad (9)$$

ISSN 2522-3488 Vol. 9, Issue 1: 57-64, DOI: 10.25236/IJNDES.2025.090108

(3)The pricing and sales volume of various vegetable categories conform to the double logarithmic model mentioned above.

$$\ln f\left(P_i(t)\right)=\alpha_i+\beta_i \ln P_i(t)+\gamma_i \ln\left(\sum_j \sum_t f\left(P_i(t)\right) \bullet P_{i_j}(t)\right)+\varepsilon_i, i=1,2,\cdots,6, \tag{10}$$

The discount rate, loss rate, cost price, and sales volume of each vegetable category must satisfy the above equation.

In summary, a single objective optimization model can be established as follows:

$$\max \sum_{i=1}^{6}\left\{\left[P_i(t)-C_i(t)\right] \bullet L_i(t)+\left[d_i \bullet P_i(t)-C_i(t)\right] \bullet\left[M_i(t)-L_i(t)\right]\right\},$$

$$s.t.\begin{cases} N_i(t) \leq M_i(t) \leq N_{i\max}, \\ 110\% \times C_i(t) \leq P_i(t) \leq 150\% \times C_i(t), \\ L_i(t)=\min\left\{M_i(t) \bullet\left(1-r_i\right), N_i(t)\right\}, \\ \ln f\left(P_i(t)\right)=\alpha_i+\beta_i \ln P_i(t)+\gamma_i \ln\left(\sum_j \sum_t f\left(P_i(t)\right) \bullet P_{i_j}(t)\right)+\varepsilon_i, \\ N_i(t)=g_i(t)+s_i(t)+h_i(t)+f\left(P_i(t)\right)+\varepsilon_i(t), \\ P_i(t)=C_i(t) \bullet\left(1+q_i(t)\right), \\ q_{i_j}(t)=\dfrac{V_{i_j}(t)-U_{i_j}(t)}{U_{i_j}(t)} \times 100\%, \\ q_i(t)=\dfrac{q_{i_j}(t)}{n_i}, i=1,2,\cdots,6. \end{cases} \tag{11}$$

The optimization model can be solved using genetic algorithm. Genetic algorithm utilizes computer simulation operations to transform the problem-solving process into the crossover, mutation, and other processes of chromosome genes in biological evolution. Setting the population size to 100, maximum algebra to 500, crossover rate to 0.7, and mutation rate to 0.2, the daily replenishment total and pricing strategy results are shown in Table 2 and Table 3.

Table 2. Daily replenishment total (unit: kg)

Date	Flowers and leaves	Florescent vegetables	Aquatic rhizomes	Eggplant type	Chili peppers	Edible fungi
July 1st	220.121	22.864	32.276	106.954	94.156	94.248
July 2nd	212.255	20.424	30.231	102.354	89.521	90.615
July 3rd	156.373	19.375	14.517	70.997	63.417	62.435
July 4th	147.494	18.727	16.592	71.542	60.524	61.728
July 5th	159.148	19.456	17.194	71.136	67.431	67.492
July 6th	148.641	18.325	16.877	72.659	67.374	65.419
July 7th	167.734	18.234	25.691	85.452	77.982	78.347

Table 3. Pricing Strategy (Unit: Yuan)

Date	Flowers and leaves	Florescent vegetables	Aquatic rhizomes	Eggplant type	Chili peppers	Edible fungi
July 1st	5.421	9.364	14.571	5.631	4.167	5.487
July 2nd	4.672	9.127	14.952	6.354	4.642	5.156
July 3rd	4.247	9.246	14.243	6.817	5.542	5.517
July 4th	4.175	8.957	14.305	6.467	5.105	5.357
July 5th	4.364	8.994	13.954	6.175	5.253	5.753
July 6th	4.439	9.132	14.341	6.854	5.815	5.254
July 7th	3.899	9.439	15.612	6.546	5.945	5.198

ISSN 2522-3488 Vol. 9, Issue 1: 57-64, DOI: 10.25236/IJNDES.2025.090108

When using the daily replenishment volume and pricing strategy of the above categories, the maximum 7-day total revenue of the supermarket is 27316.814 yuan.

4. Pricing and replenishment decisions for individual vegetable products

The scale of supermarket operations is often limited by the sales location. In the case of limited sales space, supermarkets need to replenish their inventory in a timely manner in order to sustain their operations. The quantity of sales items should not be too large, controlled between 27-33 categories, and the display quantity of each item should not be too large, with a minimum requirement of 2.5 kilograms.

As in the previous chapter, the double logarithmic model is first used to obtain the relationship between the total sales volume of individual vegetable products and cost markup pricing, and LSTM is used to predict the cost price of individual vegetable products. Record k_j as whether to replenish the j-th item and introduce a 0-1 variable:

$$k_j=\begin{cases}1,\text{replenish the } j\text{-th item},\\ 0,\text{do not replenish the } j\text{-th item},\end{cases} \quad j=1,2,\cdots,49, \tag{12}$$

In order to maximize the profits of supermarkets while meeting the market demand for various vegetable categories as much as possible, a single objective optimization model can be established with the daily replenishment volume of a single product as the decision variable:

$$\max\sum_{j=1}^{49}\left\{\left[P_j(t)-C_j(t)\right]\bullet L_j(t)+\left[d_j\bullet P_j(t)-C_j(t)\right]\bullet\left[M_j(t)-L_j(t)\right]\right\}k_j$$

$$s.t.\begin{cases}27\le\sum_{i=1}^{49}k_j\le 33,\\ N_j\ge 2.5,\\ L_j(t)=\min\left\{M_j(t)\cdot(1-r_j),N_j(t)\right\},\\ \ln f\left(P_j(t)\right)=\alpha_j+\beta_j\ln P_j(t)+\gamma_j\ln\left(\sum_j\sum_t f\left(P_j(t)\right)\cdot P_j(t)\right)+\varepsilon_j,\\ N_j(t)=g_j(t)+s_j(t)+h_j(t)+f\left(P_j(t)\right)+\varepsilon_j(t),\\ P_j(t)=C_j(t)\cdot\left(1+q_j(t)\right),\\ q_j(t)=\dfrac{V_j(t)-U_j(t)}{U_j(t)}\times 100\%,\\ k_j=\text{0 or 1},\ j=1,2,\cdots,49.\end{cases} \tag{13}$$

Among them, $P_j(t)$ represents the cost markup pricing of the j-th vegetable item on the t-th day, $M_j(t)$ represents the replenishment quantity of the j-th vegetable item on the t-th day, $L_j(t)=\min\left\{M_j(t)\bullet(1-r_j),N_j(t)\right\}$ represents the sales volume of the j-th vegetable item without discount on the t-th day, $C_j(t)$ represents the cost price of the j-th vegetable item on the t-th day, r_j represents the loss rate of the j-th vegetable item, $N_j(t)$ represents the sales volume of the j-th vegetable item on the t-th day, d_j represents the discount rate of the j-th vegetable item, $f\left(P_j(t)\right)$ represents the total sales volume of the j-th vegetable item after excluding time effects, and k_j represents whether to replenish the j-th vegetable item.

ISSN 2522-3488 Vol. 9, Issue 1: 57-64, DOI: 10.25236/IJNDES.2025.090108

The results obtained by using genetic algorithm to solve the optimization model are as follows:

Table 4. Vegetable single item replenishment and pricing decisions

Item	Restocking quantity	Price	Item	Restocking quantity	Price
Baby Bok Choy	12.04	5.54	Seven Colored Peppers (2)	4.57	6.22
Purple Eggplant (2)	21.07	4.75	Honghu Lotus Root Belt	9.30	12.22
Out of Town Chrysanthemum	15.35	4.81	Bok choy (1)	19.54	2.02
High Melon (1)	8.99	9.52	Small Wrinkled Skin (portion)	4.58	1.98
Round Eggplant (2)	18.74	5.69	Screw Pepper	3.15	2.92
Zhijiang Qinggeng Scattered Flowers	19.69	9.66	Milk Cabbage	12.31	4.18
Spinach (portion)	26.54	5.11	Shanghai Green	4.17	5.65
Heart of Cabbage	25.07	5.20	Bamboo Leaf Vegetable	8.54	3.59
Amaranth	15.57	4.07	Flammulina velutipes (box)	19.25	4.15
Green Eggplant (1)	25.87	5.02	Net Lotus Root	5.91	12.89
Mushroom Double Mix	13.64	3.55	Cordyceps Flower	12.42	3.65
Fresh Fungus	22.71	1.98	Purple eggplant	6.41	3.22
Wild Lotus Root	12.74	12.17	Mugwort	20.54	16.40
Red Lotus Root Strap	52.88	8.63	Xixia Mushroom	30.60	12.54
Water Chestnut	5.74	12.32			

According to Table 4, 29 types of vegetable products can be sold during this period, and the maximum profit on July 1st can reach 1236.52 yuan.

5. Conclusions

We conducted an in-depth analysis of the distribution pattern of vegetable sales and the relationship between sales volume and cost markup pricing by establishing Prophet models and double logarithmic models. These models are not only applicable to current data, but also have broad potential for generalization. To improve the prediction accuracy and practicality of the model, multiple sources of data such as market research and customer behavior can be combined, and real-time data streams can be introduced to make the model more dynamic and adjust pricing strategies and replenishment decisions in real time.

References

[1] Zhan C. Optimization Study on Pricing and Replenishment Decision of Vegetable Products Based on LSTM and ARIMA Models [C]//2024 International Conference on Electronics and Devices, Computational Science (ICEDCS). 0[2025-02-26]. DOI: 10. 1109/ICEDCS64328. 2024. 00148.

[2] Ma M. Automatic Pricing and Replenishment Strategies for Vegetable Products Based on Data Analysis and Nonlinear Programming [J]. Highlights in Science Engineering and Technology, 2024.

[3] Gu A, Yan Z, Zhang X, et al. Research on the Modeling of Automatic Pricing and Replenishment Strategies for Perishable Goods with Time-Varying Deterioration Rates[J]. Axioms (2075-1680), 2024, 13(1). DOI: 10. 3390/axioms13010062.

[4] Xia L, Zhang J, Wen B. Optimization decision model of vegetable stock and pricing based on TCN-Attention and genetic algorithm [J]. Journal of Agricultural Mechanization Research, [2025-02-26].

[5] Yadav R, Sreenath S, Uchoi R. Vegetable Processing Wastes as Dietary Ingredients for the Striped Catfish Pangasianodon hypophthalmus: A Case Study[J]. Aquaculture Research, 2016.