Research on the Profit Maximization Model for Automatic Pricing and Restocking of Vegetable Commodities

Yanqi Zhang, Yue Wang, Yu Zhou*

School of Architecture and Civil Engineering, Xiamen University, Xiamen, China, 361005

*Corresponding author: zhouyu6660424@gmail.com

Abstract. Due to the perishable nature of fresh agricultural products and the uncertainty of the types and prices of individual items before daily procurement in supermarkets, these establishments typically need to determine daily replenishment and pricing strategies based on the historical sales and demand for various goods. This paper starts from the perspective of maximizing supermarket revenue, predicting the sales of various product categories through seasonal indices. Under the constraints of the sales space for goods, the paper further uses ARIMA time series forecasting for the sales volume and procurement prices of individual items during the corresponding time period. Cost markup equivalent coefficients, daily replenishment quantities, and discount coefficients are considered as decision variables. Based on the demand curve model in economics, the paper takes into account the counteraction of cost markup pricing coefficients on sales volume and the counteraction of discount coefficients on the discounted portion of sales. The objective functions include maximizing supermarket revenue and satisfying the market demand for various categories of vegetable products. The paper adopts a novel sales model of offering discounts on predicted perishable vegetables and establishes a multi-objective programming model, providing the optimal strategies for replenishment quantities and pricing for each item on July 1st.

Keywords: Profit Maximization Optimization Model, Time Series Forecasting Model, Multi-Objective Programming Model, Vegetable Commodities.

1. Introduction

Perishable goods are characterized by their perishable nature and short sales cycles, making their pricing and replenishment strategies significantly impact retail enterprises' profits. Therefore, the issues related to pricing and replenishment of perishable goods have long been widely discussed[1]. The replenishment and pricing strategies for vegetables have a substantial impact on supermarket revenue[2]. When retailers place orders on demand, market demand and order quantities are matched, and there is no inventory left for retailers at the end of each replenishment cycle[3]. In fact, perishable goods can satisfy market demand only within their shelf life; after that, the products lose their utility value due to spoilage[4]. Regarding pricing, the 'cost-plus pricing' method is generally used for vegetables, and discounts are applied to damaged or subpar items during transportation. Market demand. Therefore, making reasonable replenishment and pricing decisions has a significant impact on ensuring a favorable profit situation. Zhang Jinlong combined the characteristics of perishable products to establish a joint decision model for pricing and replenishment of perishable new products and designed heuristic algorithms to solve the model[5]. Peng Zuohe assumed an exponential distribution for the perishability rate of goods, considered the time value of capital, and allowed for shortages. They established a joint ordering and pricing model for perishable goods based on the complete unit quantity discount[6]. Tian Junfeng, Sun Xixiu, and Yang Mei took into account the quality level impact factor and the shape factor of the perishability rate and solved the optimal decision-making process[7].

In the above-mentioned study of the replenishment model for vegetables, most considerations have focused on the impact of price and quality on demand, but little attention has been given to a new sales method where supermarkets can directly offer discounts on the predicted quantity of perishable products when their quality decreases or they spoil. This provides a fresh perspective for maximizing supermarket revenue. To better simulate real-world scenarios, we treat the statistically obtained vegetable spoilage rate as the proportion of discounted items in supermarkets. Additionally, we

transition from a replenishment cycle of restocking after stockouts to daily replenishment. Using historical sales volume and wholesale price data, we provide replenishment pricing decisions for each individual item for the upcoming day.

2. Introduction of models

2.1. Seasonal Index Forecasting Model

The expression for the seasonal index model is:

$$X_t = T_t S_t C_t I_t \tag{1}$$

In which, $X_t$represents the total variation of the time series;T_t stands for the long-term trend; S_t represents seasonal variation; C_t accounts for cyclic variation; and I_t is the irregular variation.

Step1: Solving for seasonal variation S and irregular variation I

$$M_t = \frac{X_t + X_{t+1} + \ldots + X_{t+N-1}}{N} \tag{2}$$

N represents the quantity of data points within one cycle, and X_t is the observed value of data at time t.

$$SI_t = 100 X_t / M_t \tag{3}$$

Calculate the seasonal index r:

$$r_t = \frac{\sum_{j=1}^{T} SI_t^{(j)}}{T}, t = 1, 2 \ldots, N \tag{4}$$

Among which, T represents the number of cycles,$SI_t^{(j)}$corresponds to the i-th data point within the j-th cycle.

After that, we make corrections to r. Assuming the corrected data is R, which represents seasonal variation S and irregular variation I:

$$R_t = r_t / \left(\frac{\sum_{t=1}^{N} r_t}{N} \right) \tag{5}$$

Step2: Solving for the long-term trend T

Based on the overall trend, fit X_t with a linear or quadratic function and represent the fitting relationship as T_t:

$$T_t = a + bt \tag{6}$$

$$T_t = a + bt + ct^2 \tag{7}$$

Using leaf type as an example, we will perform fitting with a linear function.

Step3: Solving for cyclic variation C:

$$C1_t = \frac{M_t}{T_t} \tag{8}$$

Where M_t represents the average value of periodic data, and T_t is the data fitted for the long-term trend. Taking the average of $C1_t$ yields C:

$$C = \frac{\sum_{t=1}^{n} C1_t}{n} \tag{9}$$

2.2. ARIMA model

The ARIMA model is a form of the Auto Regressive Moving Average (ARMA) model that incorporates differencing. This model is widely used for forecasting and modeling time series data[8].

Let y_t represent the actual values of a time series at different time points t. The autoregressive (AR) component of the ARMA model can be expressed as:

$$y_t = \sum_{i=1}^{p} \gamma_i y_{t-i} \tag{10}$$

In the equation: γ_i represents the autocorrelation coefficient of the AR component, and p is the order of the autoregressive terms.

Let ε_t be the white noise error term at time t, then the moving average (MA) component of the ARMA model can be represented by equation:

$$y_t = \mu + \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{11}$$

In the equation, μ represents the expected value of y_t, q is the order of the moving average terms, and θ_j(j=1, 2, …, q) represents the model parameters.

Equation 11 indicates that the value at time point t linearly depends on the current and past q white noise error terms. Combining equations 10 and 11, the ARMA model can be represented as:

$$y_t = \sum_{i=1}^{p} \gamma_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{12}$$

In practical applications, to capture the stationary characteristics of a time series, the differencing method is included in the ARMA model, resulting in the ARIMA model.

Assuming Δdy_t represents the d-th order difference of y_t,ARIMA (p, d, q) can be expressed as:

$$\Delta^d y_t = \sum_{i=1}^{p} \gamma_i \Delta^d y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{13}$$

The accuracy of ARIMA model predictions for a time series depends on the values of p and q. The parameters p and q are determined based on the AIC (Akaike Information Criterion) criterion:

$$AIC = 2k - 2ln(L) \tag{14}$$

Note: k is the number of parameters, and n is the likelihood function.

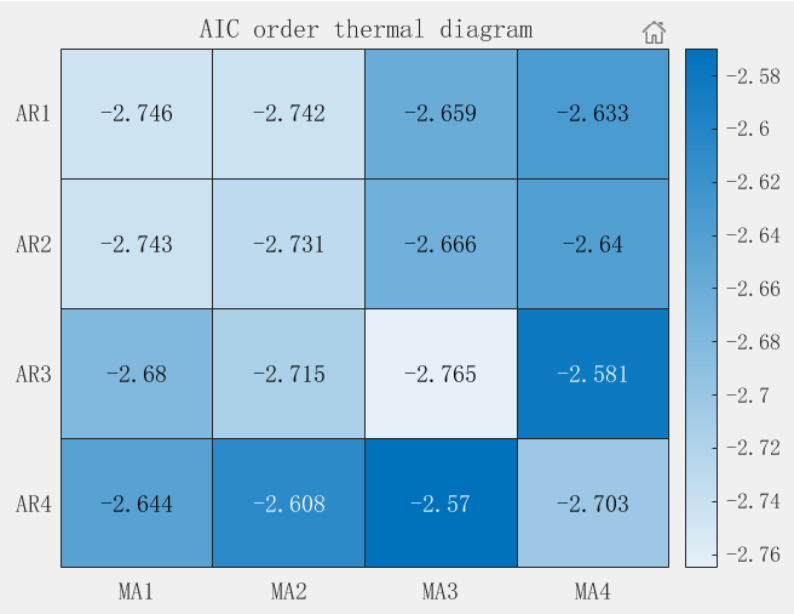

Figure 1. AIC Order Selection Matrix Plot

As shown in Figure 1, it is the AIC order selection matrix plot. For the corresponding model in Figure 1, the best order is obtained with p=3 and q=3.

2.3. Genetic algorithm

Genetic algorithm is a multi-parameter, multi-combination optimization method that simulates the principle of "natural selection, survival of the fittest" in the process of natural evolution. Its main features include population-based search methods and the exchange of individual information within the population [9]. Genetic algorithms have a slow convergence speed, and fixed crossover and mutation probabilities can disrupt individuals with higher fitness. During the iterative process, there may be frequent occurrences of stagnation or degradation[10]. The specific operational steps are as shown in Table 1:

Table 1. Genetic Algorithm Solution Steps

Genetic Algorithm Solution Steps
Step1: Set the initial state, parameters, encoding, etc., and import the data.
Step2: Initialize and start the iterative evolution.
Step3: Select, recombine, mutate, and merge feasible solutions.
Step4: Allocate fitness values based on the objective function values.
Step5: Calculate the index of the current best individual.
Step6: Repeat steps 3, 4, and 5 until evolution meets the optimum or reaches the maximum number of generations.
Step7: Output the results.

3. Results

3.1. Establishment of models

We first extract the non-discounted data for each product category and forecast the sales. For this question, the forecast is based on the valid data from June 20, 2021, to July 7, 2023 (with only June 20 to June 30, 2023, data used for 2023), with a focus on predicting the sales for July 1st for each category. Due to the seasonality of sales, we use a seasonal index forecast.

First, we conduct data preprocessing. We extract the single products that appeared between June 24 and June 30, 2023, and exclude products with too few days of appearance (<12 days) and products with sales lower than the minimum display quantity. This screening results in 37 single products and their corresponding data. We forecast the sales and wholesale prices for these selected single products on July 1, 2023, using an ARIMA forecasting model with a differencing order of i=1. As this is a multi-objective optimization problem, we use a priority solution method to solve the first-level single-objective optimization: maximizing revenue. To facilitate solving this with a genetic algorithm (GA), we consider the effect of the cost increment coefficient on sales as a linear model:

$$\xi = a + b \cdot r \tag{15}$$

To prevent the cost increment coefficient from having a significant impact on sales, we let ξ=2 when r=0.2; Let ξ=0.2 when r=2. The computed linear function is as follows:

$$\xi = 2.2 - r \tag{16}$$

To solve the problem by introducing 0-1 variant y_{j1}(j=1,2....37):

$$y_{j1} = \begin{cases} 1, \text{When item j is selected} \\ 0, \text{When item j is not selected} \end{cases} \tag{17}$$

Based on the above analysis, establish a planning model to solve for the maximum supermarket revenue.

Step1: Model establishment

1)Decision variables: Cost markup pricing coefficients for each individual item r_j, Replenishment quantities for each individual item Q_{j1}, discount factor η_{j1}, 0-1 variables y_{j1}

2)Objective function: Maximize the revenue:

$$\max R = \sum_{j=1}^{37} \left[\left(1 + r_j\right) \cdot c_j \cdot D_j \cdot \xi - Q_{j1} \cdot c_j + \left(x \cdot Q_{j1} \cdot \varepsilon\right) \cdot c_j \left(1 + r_j\right) \cdot \eta_{j1} \right] \cdot y_{j1} \tag{18}$$

3)Constraint conditions:

First, the profit margin should not be too large or too small. Here, it is set as $0.1 \leq r_j \leq 2$. Secondly, the replenishment quantity for selected items should be greater than their minimum display quantity, and the replenishment quantity for selected items should be greater than the sales quantity plus the loss coefficient multiplied by the replenishment quantity, as seen in equations 20 and 21. Additionally, the number of selected items should be greater than 27 and less than 33, as seen in equation 22. Finally, the sum of replenishment quantities for each category of products should be less than the forecasted sales quantity for that category, as seen in equation 23.

$$0.1 \leq r_j \leq 2 \tag{19}$$

$$Q_{j1} \geq 2.5 \tag{20}$$

$$Q_{j1} \geq \left(c_j + x \cdot Q_{j1}\right) \cdot y_{j1} \left(0 \leq \eta_{j1} \leq 0.8\right) \tag{21}$$

$$27 \leq \sum_{j=1}^{37} y_{j1} \leq 33 \tag{22}$$

$$\sum_{j=1}^{37} Q_{j1} \leq D_j \tag{23}$$

$$Q_{j1} > 0 \tag{24}$$

Among which, c_j is wholesale prices for each individual item in problem three, D_j is forecasted sales volume for each individual item.

After obtaining the maximum profit for the single-objective optimal value, further observation reveals, $\sum y_{i1}<33$ and the difference between the sum of the demand forecast values for various product categories and the sum of the replenishment quantities for individual items is 77.45866. Further selection is carried out for the remaining unselected single products to maximize the number of items that can be sold. This can lead to increased profits and better satisfaction of market demand.

1)Decision variables: Cost markup pricing coefficients for individual items r_j,Replenishment quantities for various product categories Q_{j2},Discount coefficients η_{j2}、0-1variables y_{j2}

2)Objective function: Minimize the difference between replenishment quantity and market demand.

$$\min \sum_{i=1}^{6} D_i - \sum_{j=1}^{37}\left(Q_{j1}\cdot y_{j1}\right) - \sum_{j=1}^{37-\sum_{j=1}^{37} y_{j1}} \left(Q_{j2}\cdot y_{j2}\right) \tag{25}$$

$$Q_{j2} \geq 2.5 \tag{26}$$

$$0 \leq \eta_{j2} \leq 0.8 \tag{27}$$

$$\sum_{j=1}^{10} y_{j2} = 33 - \sum_{j=1}^{37} y_{j1} \tag{28}$$

$$\sum_{j=1}^{10} Q_{j2} \leq D_j - \sum_{j=1}^{37} Q_{j1} \tag{29}$$

$$Q_{j2} > 2.5 \tag{30}$$

$$\max R > 0 \tag{31}$$

3.2. Results analysis

First, sales volume for each product category on July 1, 2023, is forecasted using a seasonal index prediction model. It can be seen from the forecasted values and error values that the forecast errors for each product category are relatively small, indicating good forecasting results, as shown in Table 2.

Table2 . Sales forecast for various product categories on July 1, 2023

Type / Statistical items	Foliage plants	Cauliflower	Aquatic tubers and rhizomes	Eggplant	Chillies	Edible mushrooms
Forecasted sales/kg	128.41	26.39	22.30	17.24	48.95	33.83
Inaccuracies	0.0126	0.0018	0.0319	0.2799	0.1124	0.0657

Next, the ARIMA model is used to forecast the sales and wholesale prices for known individual item data on July 1, 2023. It can be observed that overall, both price and sales forecast errors are relatively small, indicating a good forecasting performance. However, there is some error, particularly when sales fluctuate significantly, as shown in Table 3.

Table 3. Sales and wholesale price forecast for 37 individual items on July 1, 2023

Type / Item Number	Forecasted sales	Inaccuracies	Forecasting wholesale prices	Inaccuracies
5	3.035476840	0.302665432	2.155078146	0.000496425
7	12.986144040	0.032727053	1.953225774	0.003805858
10	1.395564248	0.242977695	4.095802376	2.114849909
16	6.045118935	0.208221451	2.704809748	0.002304174
~	~	~	~	~
248	4.776228447	0.269895870	1.938343656	2.029919700

Subsequently, genetic algorithms are used to calculate the maximum revenue for the supermarket. To prevent genetic algorithms from getting trapped in local optima, a condition R > 850 is set to obtain a better global optimum solution, as shown in the forecasting results in Table 4.

Table 4. Replenishment quantity and pricing strategy data for individual items on July 1, 2023 (Partial)

Type / Item Number	Cost Increment Pricing Factor r_j	0—1 varianty_{j1}	Replenishment Quantity Q_{j1}	Discount Factor η_{j1}
5	1.350420928	1	3.724667526	0.412258512
7	0.621182464	1	15.03267106	0.532674494
10	0.622387087	1	2.49991469	0.540095636
16	0.593760111	0	3.990290696	0.490918947
~	~	~	~	~
248	0.600192787	1	4.775432132	0.258077854

At this point, the supermarket maximizes its revenue, with R_{max}=1007.980 yuan.

Lastly, further selection is carried out for the remaining unselected single products to maximize the number of items that can be sold, effectively improving the degree of satisfaction with market demand and increasing profits. The forecast results are shown in Table 5.

Table 5. Replenishment Quantity and Pricing Strategy Data for New Single Products on July 1, 2023 (Partial0—1 variant)

Type / Item Number	Cost Increment Pricing Factor r_j	0—1 $varianty_{j2}$	Replenishment Quantity Q_{j2}	Discount Factor η_{j2}
6	1.365045447	1	20.03535686	0.623738379
40	0.646715223	1	11.05623133	0.144943138
59	0.285793431	1	9.930047719	0.334915928
123	0.84956263	1	7.377346405	0.643905587
~	~	~	~	~
242	0.392197562	1	9.263508063	0.495467867

The newly calculated unit profit is 7.264 yuan, with an overall deviation from market category demand of 13.274 kg. The total profit is 1015.244 yuan, with a deviation from demand of 13.274 kg.

From the forecast results, the multi-objective planning model demonstrates good predictive performance, striving to meet market demand for various vegetable products while ensuring maximum supermarket revenue.

4. Conclusions

Historical sales and wholesale prices provide a basis for vegetable replenishment pricing. However, previous research has not analyzed the relationship between discounted vegetable sales and the level of discount and how it impacts sales. This gap has limited the ability to provide valuable recommendations for supermarket decisions. In this paper, a linear function is used to fit the impact of prices on predicted sales and the relationship between discounts and sales of discounted items, better simulating real-world scenarios. Finally, a genetic algorithm is employed to solve the decision results, resulting in outcomes that closely match market demand. The difference from market demand is only 13.27 kg, demonstrating good predictive performance and the ability to provide effective recommendations for supermarket decisions.

References

[1] Chen Hong, Zhou Zongfang, Chen Jun. Inventory Replenishment Strategy during the Value-Added Period of Fresh Agricultural Products. Systems Engineering, 30(1):91–96,2012.

[2] Zhang Rui, Lin Feng, Jia Tao. Optimal Ordering and Pricing Decisions for Non-Perishable Goods Considering Shelf Life[J]. Operations Research and Management,2019,28(05):2634.

[3] Lei Lijuan. Replenishment and Pricing Decisions in Perishable Food Supply Chain Based on Virus Marketing Demand Forecast[D]. Beijing Jiaotong University, 2021.

[4] Yang Tianshan, Yuan Gonglin. Research on Dynamic Pricing and Replenishment Strategies for Perishable Goods Considering Green Delivery Technology Investment[J/OL]. Chinese Journal of Management Science:1-13[2023-09-27].

[5] Zhang Jinlong, Wu Xiang, Xu Haoxuan. Joint Decision Model for Pricing and Replenishment of Perishable New Products[J]. Journal of Systems Engineering,2018,33(01):7989.

[6] Peng Zuohe, Tian Peng. A Pricing and Inventory Model for Perishable Goods Based on Quantity Discounts[J]. Journal of University of Shanghai for Science and Technology,2004(06):565568+574.

[7] Tian Junfeng, Sun Xixiu, Yang Mei. A Generalized Model for Joint Replenishment and Pricing Decision Considering Demand and Price-Quality Relationships[J]. Logistics Engineering and Management,2014,36(05):188192.

[8] Li An, Gao Mengmeng, Chen Xi, Zhang Xiaohua, Li Jingbo, Li Liangtao. Predicting the Suitable Distribution Area of Taihang Mountain Genus Using the MaxEnt Model and Future Climate Conditions[J]. Henan Agricultural Science,2021,50(04):137146.

[9] Wang Lei, Wang Yixuan, Li Dongdong, et al. Research on Mobile Robot Path Planning Based on Improved Genetic Algorithm[J/OL]. Journal of Huazhong University of Science and Technology (Natural Science Edition):113[20230927].

[10] Li Yan, Yuan Hongyu, Yu Jiaqiao, et al. A Review of the Application of Genetic Algorithms in Optimization Problems[J]. Shandong Industrial Technology,2019(12):242243+180.