

Vegetable Replenishment and Pricing Planning Based on ARIMA Forecasting

Xinchao Shao, Fuxi Xiao *, Fansheng Sun

School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, 710049

* Corresponding Author Email: xiao.fuxi@icloud.com

Abstract. In fresh food supermarkets, vegetables have a short shelf life and can deteriorate quickly, necessitating daily replenishment and pricing decisions based on historical sales to maximize profits. This article, based on the historical sales data of vegetable products from a major supermarket, constructs a replenishment and pricing planning model based on ARIMA forecasting for six types of vegetables, seeking the optimal replenishment and pricing strategy for the next seven days. The results indicate that leafy and flowering vegetables, as well as eggplant-type vegetables, require more replenishment on Fridays and weekends, whereas aquatic root vegetables need more on Thursdays and Fridays. The replenishment of chili vegetables and edible mushrooms fluctuates, and there is a constraint relationship between pricing and replenishment volume. This study aims to assist merchants in making reasonable replenishment and pricing decisions to reduce costs and maximize economic benefits.

Keywords: Unitary Linear Regression, ARIMA, Nonlinear Programming, Replenishment and Pricing.

1. Introduction

In fresh food supermarkets, a variety of fresh vegetables are offered, which generally have high water content, short shelf life, and are prone to spoilage and damage. The freshness of vegetables decreases over time, and those of inferior quality may end up being discounted or unsellable. For certain seasonal types of vegetables, the impact of the season also needs to be considered [1-2]. In response, merchants need to make daily, reasonable replenishment and pricing decisions based on historical prices and sales data of various vegetable categories [3-4] to achieve maximum economic benefits.

In recent years, scholars both domestically and internationally have extensively studied the pricing and replenishment strategies of perishable products. Tang [5] (2004) observed that when selling perishable, seasonal products, the season available for sale is short and the replenishment turnover time is long, preventing retailers from using actual sales data from the start of the season to update demand forecasts. Jia [6] (2011) analyzed the sales and pricing conditions of perishable goods in specific areas, suggesting that the optimal pricing strategy and order quantity for fresh products depend on the wholesale price, and there is a constant constraint relationship between the optimal pricing strategy and order quantity. These studies lack statistical analysis of extensive historical sales data, and research on the relationship between market demand and pricing is limited.

Given the issues in current research, this paper proposes a nonlinear programming model based on ARIMA forecasting [7-10], focusing on the optimal replenishment and pricing strategy for six vegetable categories. This study comprehensively considers factors such as time, the constraint relationship between pricing and sales volume, and the decisive role of market demand in replenishment strategy, hoping to help fresh food supermarkets resolve replenishment difficulties and enhance economic benefits as much as possible.

2. Vegetable replenishment and pricing planning model based on ARIMA forecasting

2.1. Regression analysis of daily vegetable sales volume and pricing

To explore the correlation between daily sales volume and pricing of various vegetable categories, this paper focuses on six common types of vegetables in supermarkets, using the weighted average price P (in yuan / kg) as the independent variable and daily sales volume S (in kg) as the dependent variable, considering the unitary linear regression analysis based on historical statistical data.

$$S_i = a_i P_i + b_i + \xi_i \quad (i = 1, 2, \dots, 6) \quad (1)$$

Using MATLAB for fitting, the linear relationship between daily sales volume and pricing for each vegetable category was obtained, as shown in Table 1:

Table 1. Results of unitary linear regression.

Vegetable category	Regression equation
Flower leaf class	$S_1 = -13.72 P_1 + 257.4 + \xi_1$
Florescent vegetables	$S_2 = -2.681 P_2 + 63.82 + \xi_2$
Aquatic rhizomes	$S_3 = -3.33 P_3 + 69.82 + \xi_3$
Solanaceae	$S_4 = -1.013 P_4 + 30.31 + \xi_4$
Chili peppers	$S_5 = -2.532 P_5 + 106.4 + \xi_5$
Edible fungi	$S_6 = -5.169 P_6 + 114.5 + \xi_6$

As Table 1 indicates, regardless of the vegetable category, there is always a negative correlation between daily sales volume and pricing, meaning that an increase in pricing leads to a decrease in sales volume. Therefore, in practical scenarios, vegetables should not be priced too high, as it may cause a significant decrease in daily sales volume, consequently reducing overall revenue.

2.2. Future daily sales forecast based on ARIMA(p, d, q) model

Based on the historical data of daily sales volume for the six vegetable categories, different Autoregressive Integrated Moving Average (ARIMA) models were constructed to predict the daily sales volume for different vegetable categories over the next seven days, with the actual daily sales volume expected to fluctuate around this predicted value.

ARIMA is a mathematical modeling method widely used in time series analysis. Its fundamental idea is to predict future data trends based on historical data patterns, combining autoregression (AR), differencing (I), and moving average (MA) components to capture non-stationarity and autocorrelation in data.

For time series data y_t , before applying the ARIMA model, d -th order differencing is required to reduce or eliminate trends and seasonal variations, resulting in a stationary series y'_t . The differencing process is as follows:

$$y'_t = \Delta^d y_t = (1 - L)^d y_t \quad (2)$$

Based on the differenced time series y'_t , the formula for the ARIMA model can be expressed as:

$$y'_t = \mu + \sum_{i=1}^p \phi_i y'_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (3)$$

Here, ϕ_i represents the fitting parameters of the AR model, describing the relationship between the current value and the past p time sequence values. θ_j Represents the fitting parameters of the MA model, describing the relationship between the current value and the errors of the past q time sequence values. ε_t And μ are the error term and constant term, respectively.

The steps to establish the ARIMA prediction model are as follows:

Step 1: Model parameter estimation

From formula (3), the ARIMA model can be represented as ARIMA (p, d, q), where p is the order of the AR model, q is the order of the MA model, and d is the order of differencing.

For the daily sales volume data of different vegetable categories, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots of the time series are made. Based on their tailing and truncating properties, the model parameters (p, d , and q) are preliminarily determined.

Then, the optimal model parameters are selected according to the Bayesian Information Criterion (BIC):

$$BIC = n \ln(T) - 2 \ln(\max) \quad (4)$$

Here, n represents the number of parameters in the model, reflecting its complexity, while $\max$ denotes the maximum likelihood function value of the model, indicating the degree of fit to the data. A smaller BIC value implies a better model.

Step 2: Model verification

Conduct a white noise test (such as the Ljung-Box test) on the stationary series. If the data is not white noise, it indicates that there is still a correlation between the series values, necessitating a re-ordering and fitting of the ARIMA model.

2.3. Replenishment and pricing planning model based on Monte Carlo simulation

2.3.1. Determination of the objective function

Based on the forecast of daily sales volume for each vegetable category over the next seven days, daily revenue is set as the objective function to establish a single-objective nonlinear programming model. For vegetable category i , the sales profit on the j -th future day is $Q_i(j)$, thus the total revenue of the supermarket on the j -th future day is Q_j :

$$Q_j = \sum_{i=1}^6 Q_i(j) = \sum_{i=1}^6 (P_{ij} - C_{ij}) \times V_{ij} \quad (j = 1, 2, \dots, 7) \quad (5)$$

Here, P_{ij} , C_{ij} and V_{ij} represent the average pricing, average cost, and daily sales volume, respectively, of vegetable category i on the j -th future day. Considering that the average costs of different vegetable categories do not fluctuate significantly in the short term, this paper takes the average cost of the past week as the cost for the next seven days.

Based on the correlation between daily sales volumes and pricing in section 2.1, the optimization objective can be expressed as:

$$\max \sum_{i=1}^6 (P_{ij} - C_{ij}) \times (a_i P_{ij} + b_i + \xi_i) \quad (j = 1, 2, \dots, 7) \quad (6)$$

2.3.2. Determination of constraint conditions

Using the daily sales volume forecast data for different vegetable categories from section 2.2, combined with the correlation between daily sales volumes and pricing in section 2.1, the expected pricing of each vegetable category over the next seven days can be estimated:

$$P_{ij \text{ predict}} = (V_{ij \text{ predict}} - b_i) / a_i + \xi_i' \quad (7)$$

The actual pricing should fluctuate around this value. Based on historical fluctuations in vegetable pricing, the fluctuation range can be set within a certain interval, thus establishing the constraint conditions for vegetable pricing:

$$0.85 P_{ij \text{ predict}} \leq P_{ij} \leq 1.15 P_{ij \text{ predict}} \quad (8)$$

2.3.3. Construction of the revenue maximization model

By consolidating the objective function and constraint conditions, the supermarket revenue maximization model can be summarized as follows:

$$\begin{aligned} \max \sum_{i=1}^6 (P_{ij} - C_{ij}) \times (a_i P_{ij} + b_i + \xi_i) \\ \text{s.t. } 0.85 P_{ij \text{ predict}} \leq P_{ij} \leq 1.15 P_{ij \text{ predict}} \\ (i = 1, 2, \dots, 6; j = 1, 2, \dots, 7) \end{aligned} \quad (9)$$

2.3.4. Strategy solving based on Monte Carlo simulation

Since MATLAB's solution to nonlinear programming problems tends to fall into local optima, and considering that the independent variable of the objective function is only vegetable pricing, the model can be optimized based on Monte Carlo simulation.

Monte Carlo simulation generates a large number of random samples and uses them as independent variables to obtain the corresponding objective function values, selecting the solution with the highest function value as the initial solution for the model problem. Due to the large number of samples and their random generation, the final solution is generally the global optimum, which can also speed up convergence and reduce computation time for this problem.

3. Results

3.1. Forecast of daily sales volume for six types of vegetables over the next seven days

For different vegetable categories, the optimal ARIMA model parameters were selected. Based on historical statistical data, the daily sales volume (in kg) for the next seven days was forecasted, with results shown in Figure 1:

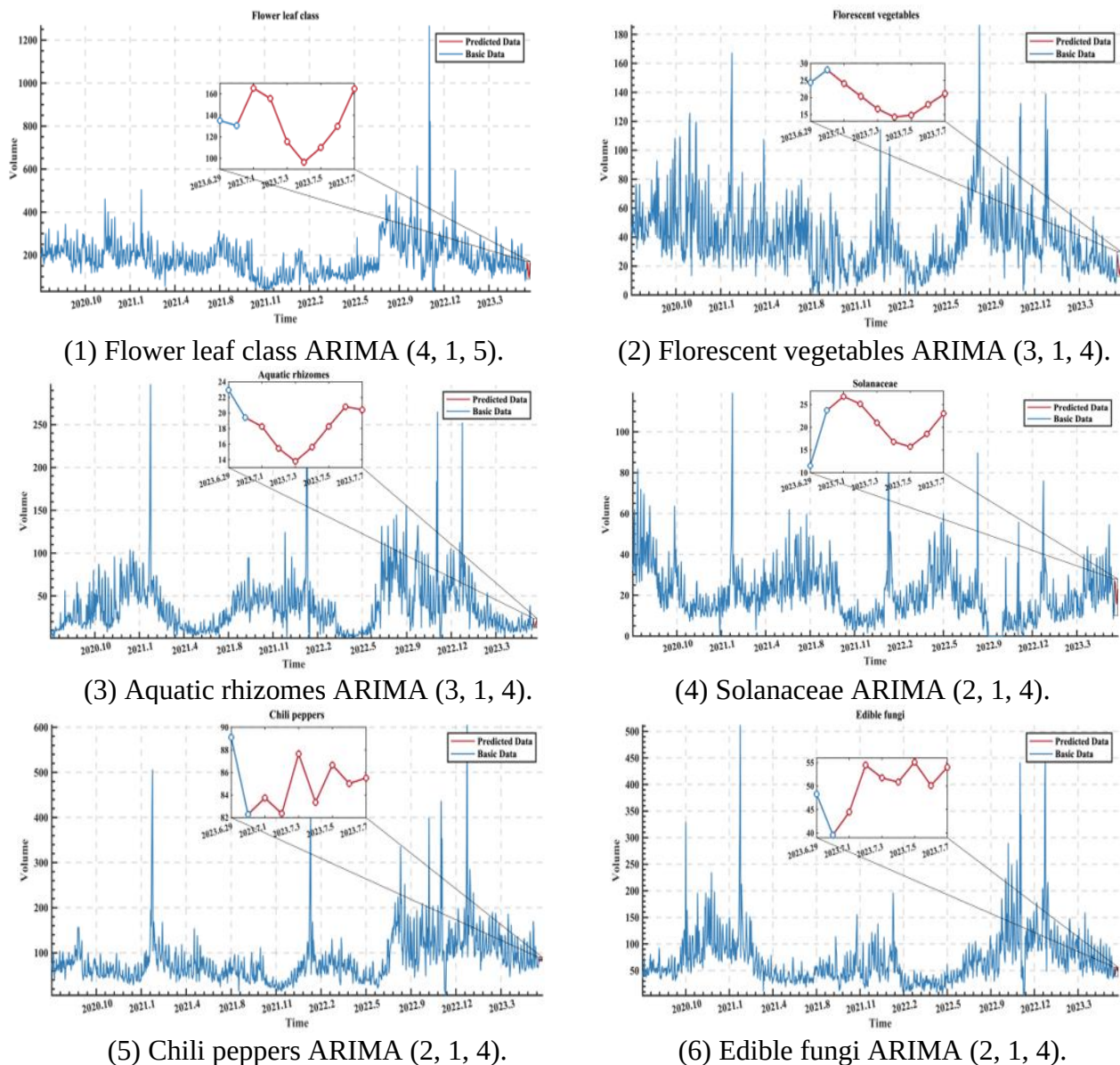


Figure 1. Forecast of daily sales volume for the next seven days.

Figure 1 shows that the sales of aquatic root vegetables, eggplant-type vegetables, and edible mushrooms exhibit noticeable seasonality. For aquatic root vegetables, the peak sales season is from August to February of the following year, and the off-season is from April to June. For eggplant-type vegetables, the peak sales season is from May to August, and the off-season is from November to December. The peak sales season for edible mushrooms is from October to February of the following year, and the off-season is from April to July. The forecasted daily sales volume for the next seven days of different vegetable categories indicates a significant increase in sales of leafy vegetables, flowering vegetables, and eggplant-type vegetables on July 1, July 2, and July 7. For aquatic root vegetables, a notable increase in sales is expected on July 6 and July 7, while the daily sales volume of chili vegetables and edible mushrooms fluctuates.

3.2. Optimal replenishment strategy for six types of vegetables over the next seven days

An algorithm was written in MATLAB to solve the model, determining the optimal replenishment and pricing strategy for six types of vegetables for the next seven days, as shown in Figure 2: (where the blue solid line represents the daily replenishment volume (in kg) for each vegetable category, and the red solid line represents the average daily pricing (in yuan / kg))

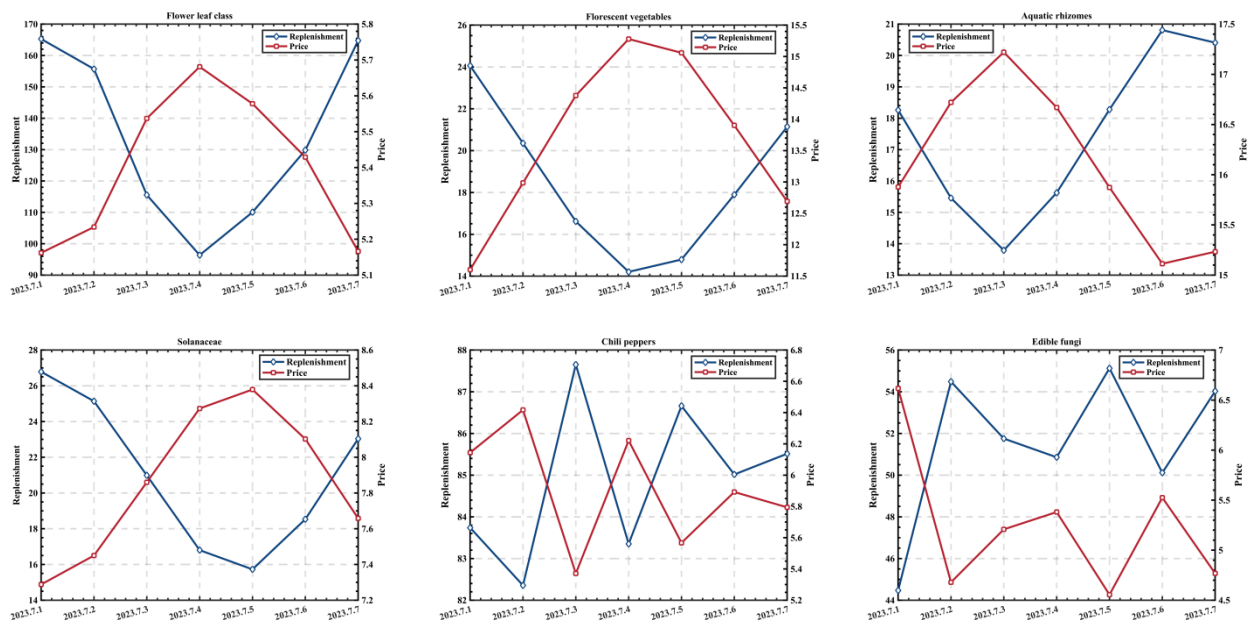


Figure 2. Optimal replenishment volume and average pricing for the next seven days.

Figures 2 illustrate that the replenishment patterns for leafy vegetables, flowering vegetables, and eggplant-type vegetables are quite similar. Increased replenishment is required on July 1 (Saturday), July 2 (Sunday), and July 7 (Friday), as customer numbers on holidays are higher than on weekdays, and a slight reduction in pricing can stimulate consumption. For aquatic root vegetables, increased replenishment is also required on July 6 (Thursday) and July 7, with a slight decrease in pricing. For chili vegetables and edible mushrooms, daily replenishment volumes fluctuate, necessitating timely adjustments in pricing.

To further verify the rationality of the replenishment strategy, the leafy vegetables with the highest sales volume were selected. The daily sales volume data for one week in the same period from 2020-2022 was compared with future replenishment strategies, as shown in Figure 3:

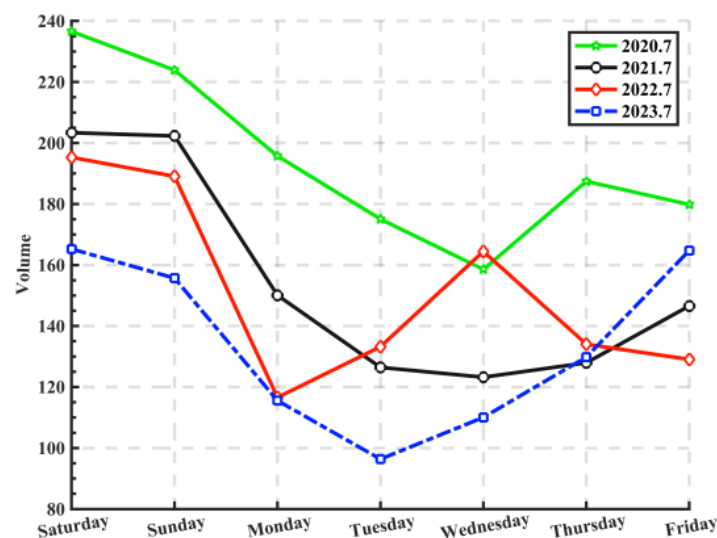


Figure 3. Comparison of daily sales in the same period (Flower leaf class).

According to Figure 3, during the same period from 2020 to 2022, the sales of leafy vegetables were highest on weekends. In 2020 and 2022, the peak sales for leafy vegetables occurred on Thursdays and Wednesdays of the workweek, respectively (though these values were less than weekend sales). The replenishment strategy for the next seven days is very similar to the daily sales

pattern in the same period of 2021, with increased replenishment on weekends due to higher vegetable demand, the lowest replenishment demand on Tuesday, and a gradual increase in daily replenishment thereafter. This analysis demonstrates the rationality of our replenishment strategy, aligning with the sales patterns of the vegetable market during the same period.

4. Conclusions

This article, based on historical pricing and sales data of six types of vegetables, has constructed a replenishment and pricing planning model based on ARIMA forecasting, studying the replenishment and pricing issues of fresh supermarket vegetable products. The results show that leafy vegetables, flowering vegetables, and eggplant-type vegetables require increased stock on Fridays and weekends. Aquatic root vegetables need more stock on Thursdays and Fridays, while the daily replenishment of chili vegetables and edible mushrooms fluctuates. Vegetable pricing should be timely adjusted according to the constraint relationship with replenishment volume, ensuring that vegetables are sold out on the same day as much as possible. In practical scenarios, when merchants face short-term vegetable replenishment and pricing issues, they can make effective decisions based on the methods presented in this article.

References

- [1] Dharavath R, Khosla E. Seasonal ARIMA to Forecast Fruits and Vegetable Agricultural Prices[C]//2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS). IEEE, 2019.
- [2] Biao Z, Lingxian Z, Zetian F U, et al. Analysis and Forecast of Vegetable Price Trend Based on Seasonal Index[J]. Northern Horticulture, 2017.
- [3] Fengyan F. Study on the pricing model of fresh commodities in supermarket [D]. Fujian Normal University, 2020.
- [4] Wujun, Cao, Xiali, et al. Pricing Policy of Supermarket's Fresh Agricultural Products Based on Demand-Price Elasticity[C]//The. 2012.
- [5] Tang C S, Rajaram K, Alptekino. Lu A, et al. The Benefits of Advance Booking Discount Programs: Model and Analysis [J]. Management Science, 2004, 50(4):465-478.
- [6] Jia J, Hu Q. Dynamic ordering and pricing for a perishable goods supply chain [J]. Computers & Industrial Engineering, 2011, 60(2):302-309.
- [7] Shukla M, Jharkharia S. Applicability of ARIMA Models in Wholesale Vegetable Market: An Investigation [J]. International Journal of Information Systems and Supply Chain Management (IJISSCM), 2013, 6(3):105-119.
- [8] Mingming Z, Yipeng H, Junqin W, et al. Forecast of Vegetable Prices in China Based on ARMA Model—A Case Study of Cabbage, Cucumber and Tomatoes[J]. Journal of Changjiang Vegetables, 2015.
- [9] Minmin Z. Research on dynamic pricing strategy of fresh community supermarket based on time context A [D]. Southwest University of Finance and Economics, 2022.
- [10] Mi1randa Suci, Helia Vembri Noor. Quantitative analysis on the fluctuation of vegetable price in supermarket [J]. MATEC Web of Conferences, 2018, 154.