A study of vegetable replenishment and pricing based on nonlinear programming

Jinpeng Ye[1, *], Weipeng Ye[2], Gongwen Liu[3], Xuejun Chen[4], Yinglin Huang[5]

[1]Department of college of mathematics and computer, Shantou University, Shantou, China, 515821

[2]College of Engineering, Shantou University, Shantou, China, 515063

[3]College of Science, Shantou University, Shantou, China, 515063

[4]College of Mathematics and computer, Shantou University, Shantou, China, 515821

[5]College of Business, Shantou University, Shantou, China, 515821

*Corresponding author: 22jpye1@stu.edu.cn

Abstract. Vegetable replenishment and pricing are the key concerns of supermarket, and accurately predicting the future demand and pricing of vegetables in supermarket is of great significance for supermarket to improve their revenue. This paper predicts the mean and variance of the demand and pricing of each vegetable category in supermarket in the future based on the ARIMA time series and constructs a nonlinear planning model to maximize the revenue. The model yields the optimal demand and pricing for the supermarket for the next seven days, and the results show that the sales of all vegetable categories except edible fungi continue to rise. The results fit the historical situation and the predictions are accurate, which provides a reference basis for the replenishment and pricing of the supermarket.

Keywords: Vegetable Replenishment and Pricing, ARIMA, Time Series, Nonlinear Programming.

1. Introduction

Today's consumers are increasingly concerned about the freshness of vegetables due to their perishable nature, which makes it difficult for them to be stored for long periods of time[1]. Therefore, the quality of vegetables generally decreases with time.

The shelf life of fresh vegetables is measured in days, and after the shelf life, the value of vegetables significantly decreases.[2] Therefore, supermarkets need to order and sell vegetables according to their daily demand[3]. If the supermarket stocked too much, it would result in wastage of vegetables and economic loss; if it stocked too little, the supply short of demand and potential revenue would be missed; similarly, both too high and too low selling price of vegetables are not conducive to the sale of vegetables. Therefore, appropriate vegetable replenishment and pricing is of great significance for supermarkets to maximize revenue. In this process, it is crucial to regulate the demand and pricing of the vegetable category, as it directly determines the subsequent sales volume and the supermarket's daily revenue.

For the prediction of sales volume, there are methods of manual prediction based on experience, but there are problems of obvious subjectivity and low accuracy [4]; there are methods of linear prediction [5] and nonlinear prediction [6], but ignoring the characteristics of historical data changes over time.

For pricing strategy, Minghua,W et al[7]. used back propagation neural network to predict the price of agricultural products in 2012; Guo Qiang applied genetic algorithm to neural network for price prediction in 2011[8]. However, the theoretical significance of the above machine learning methods is lacking, and for time series dominated by linear components, the fitting effect will be unstable and the accuracy needs to be improved[9].

And the autoregressive integrated moving average model (ARIMA) pioneered by Box- Jenkins would be a suitable choice for modeling vegetable price and sales volume forecasting[10]. In this paper, the sales data, wholesale prices, and recent depletion data of a Chinese supermarket from July

1, 2020 to June 30, 2023 are selected and fitted to the historical data using ARIMA time-series analysis to predict future demand and pricing. Subsequently, the objective function based on revenue is constructed, along with the upper and lower bound constraints utilizing the mean and variance derived from ARIMA. This establishes a nonlinear programming model. Utilizing the particle swarm optimization algorithm, the nonlinear programming model with upper and lower bound constraints is resolved. Finally, the replenishment and pricing of goods are determined under the condition of maximum benefit.

2. Modeling component

2.1. ARIMA time series analysis

Traditional regression forecasting methods for historical data that change over time do not take into account new trends in development and the impact of other factors, so the accuracy is poor. While ARIMA model is a time series analysis method, which consists of autoregressive model (AR), algorithmic moving average model (MA) and difference algorithm together, by analyzing and fitting the historical data, it can well predict the future data changes and trends [11].

The modeling steps are as follows:

(1) Linear expansion of historical data by AR modeling

$$Y_t = \kappa + \sum_{i=1}^{q} \phi_i \epsilon_{t-i} + \epsilon_t \tag{1}$$

Where α_i denotes the autocorrelation coefficient; t denotes the error at moment t, v denotes the constant term, and P denotes the number of historical data.

(2) Calculation of error accumulation by MA modeling

$$\begin{gathered} Y_t = \kappa + \sum_{i=1}^{q} \phi_i \epsilon_{t-i} + \epsilon_t \\ D^{(d)} Y_t = \mu + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \phi_i \epsilon_{t-i} \end{gathered} \tag{2}$$

In equation (2), ϕ_i is the coefficient of the bias value after adding weights; q is a linear combination of order q of the current data Y_t being expanded into the error generated by the autoregressive process; and κ is a constant.

(3) Calculate the number of differences d to transform the data into a smooth series

The use of AR model needs to have smoothness, so it is necessary to perform the difference calculation to obtain the required number of differences d .

Combine the AR and MA models to obtain the expression for the ARIMA (p ,d ,q) model

$$D^{(d)} Y_t = \mu + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \phi_i \epsilon_{t-i} \tag{3}$$

In Eq. (3),$D^{(d)}Y_t$ is the d-order difference sequence of the time series data Y_t; μ is a constant.

2.2. The establishment of nonlinear planning models

2.2.1 Establishment of the objective function

Supermarket to achieve maximum net income for the purpose (i.e., for the profit), so set z for the supermarket profit; supermarket purchase quantity for y; ordering transportation to the supermarket process will exist in the loss, the final successful transportation to the supermarket amount of vegetables is the supermarket can achieve the maximum sales volume on the day. Assuming that the sales volume forecast is accurate, the vegetables transported to the supermarket is the sales volume of the day.

Which can be written in the form:

$$order\ quantity \times (1 - loss\ rate(\%)) = sales\ volume \tag{4}$$

With this formula (4), the historical three-year order quantity data can be calculated.

And the profit for one day is as follows:

$$sales\ volume \times (sales\ price - trade\ price) - loss\ rate \times \frac{sales\ volume}{1 - loss\ rate} \times trade\ price \tag{5}$$

Where "$sales\ volume \times (sales\ price - trade\ price)$" is the revenue received by the supermarket through the sale of vegetables, "$\frac{sales\ volume}{1 - loss\ rate}$" is the initial order quantity of the supermarket, and "$loss\ rate \times \left[\frac{sales\ volume}{1 - loss\ rate}\right] \times trade\ price$" is the money lost by the supermarket from the loss of the vegetable portion.

2.2.2 Determine the constraints

For the objective function, selling price, sales volume and wholesale price are three independent unknown variables.

Subsequently, the mean and variance of these three variables are predicted by fitting the ARIMA time series, the mean±variance of the three variables is used as the constraint range.

2.2.3 Particle swarm algorithm solution

Particle swarm algorithm is a kind of method to realize the optimal value search through the cooperation and competition between individuals of the population, the essence is that the particles first fly in the optimization space, and exchange information with the group, and finally find the optimal position. In the process of optimization search, each particle needs to update its velocity and position, and its iterative equation is as follows[12]:

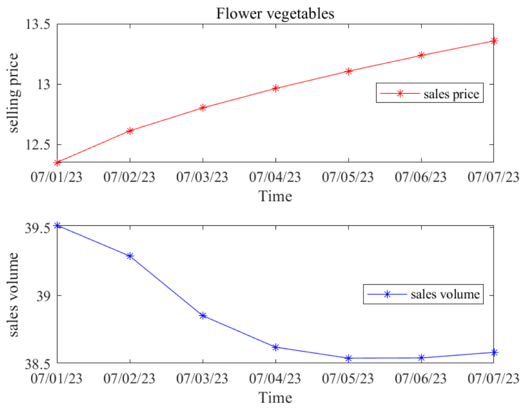
$$v_{i+1} = \omega v_i + c_1 r_1 (pbest_i - x_i) + c_2 r_2 (gbest - x_i) \tag{6}$$

$$x_{i+1} = x_i + v_{i+1} \tag{7}$$

Where $v(i)$ denotes the velocity of the ith particle; *x(i)* denotes the position of the ith particle; $pbest$ is the current optimal position of the individual ith particle; $gbest$ is the current optimal position of the particle swarm as a whole; ω is the weight coefficient, which denotes the degree of retention for the last velocity; c_1, c_2 is the learning factor of the particles, which denotes the degree of adjustment of the particles towards the optimal position; and r_1, r_2 are random numbers in the range of [0,1].

3. Results

By solving through nonlinear programming, the data of the wholesale price and revenue of each category sales of the supermarket for the next 7 days are obtained. The line graph is shown as Figure 1、2、3、4、5、6.



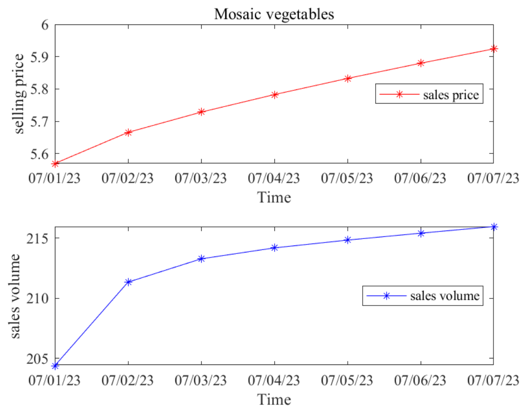
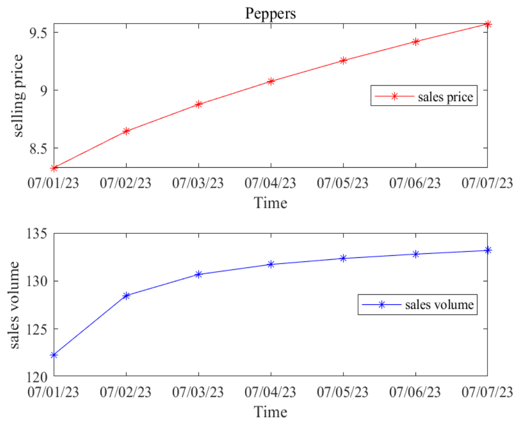


Figure 1. Flower Vegetables Sales Trend Figure 2. Foliage Sales Trend



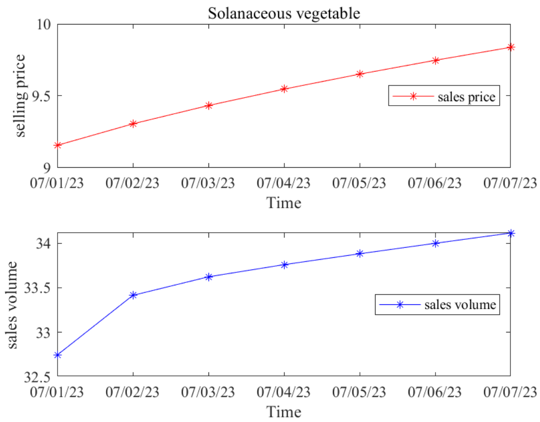
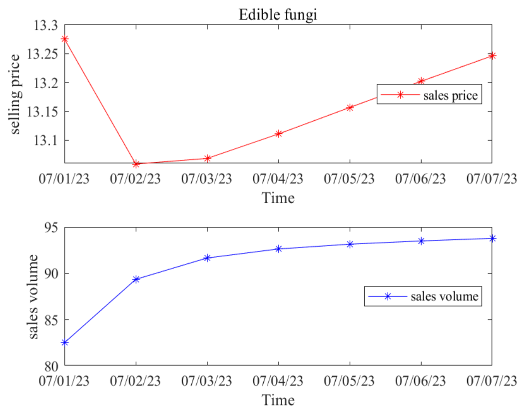


Figure 3. Chili Sales Trend Figure 4. Solanaceous Sales Trend



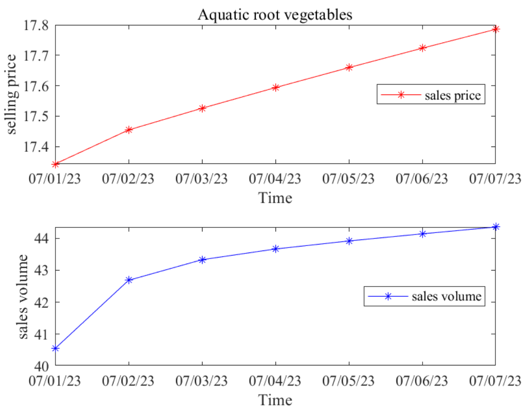


Figure 5. Edible Fungic Sales Trend Figure 6. Aquatic Root Sales Trend

For the results of this data, it can be found that in the next seven days, the unit price and order quantity of each vegetable category are basically rising gradually, showing a small increase in the trend of the small cycle. In general, it can be considered that the demand increases, so the order quantity increases; at the same time, the demand relationship affects the price, and the unit price of sales also rises gradually.

And flower vegetables sales volume of seven days gradually reduced but the unit price of sales increased. In China, flower vegetables are usually sown in July, a time when the supply of cauliflower is slowly decreasing. As a result, the rarity of cauliflower during this period leads to a higher demand, which in turn drives up its price.

For chili peppers, the sales unit price decreased by 0.15 on July 2, 2023, which is not a big fluctuation and belongs to the price change in a small cycle.

Comparing the sales demand and sales unit price in previous years(Table 1、2):

Table 1: Historical Daily Sales Volume Statistics of Vegetables by Category

Vegetable Category	Mosaic vegetables	Aquatic Root	Flower vegetables	Solanaceous vegetables	Pepp er	Edible fungi
Avg Daily Sales	181.30	37.06	38.14	20.49	83.6 4	69.49
Std Dev	87.55	31.42	22.89	13.57	53.7 9	48.73
Next 7-days Sales Average	212.77	43.23	38.85	33.65	130. 21	90.95

Table 2: Historical Unit Price Statistics of Vegetables Sold by Category

Vegetable Category	Mosaic vegetables	Aquatic Root	Flower vegetables	Solanaceous vegetables	Pepper	Edible fungi
Avg Daily unit sales price	6.24	10.37	9.47	8.90	10.25	12.15
Std Dev	1.52	3.77	2.48	2.49	4.45	2.65
Next 7 days' avg selling price	5.77	17.58	12.91	9.53	9.03	13.16

Plotting a line graph demonstrates the following:

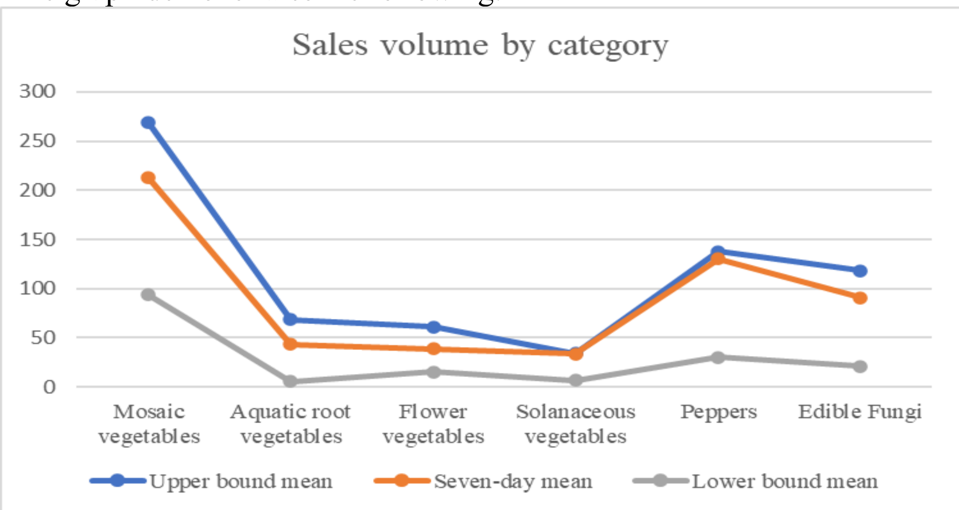


Figure 7. Sales Volume Comparison Chart

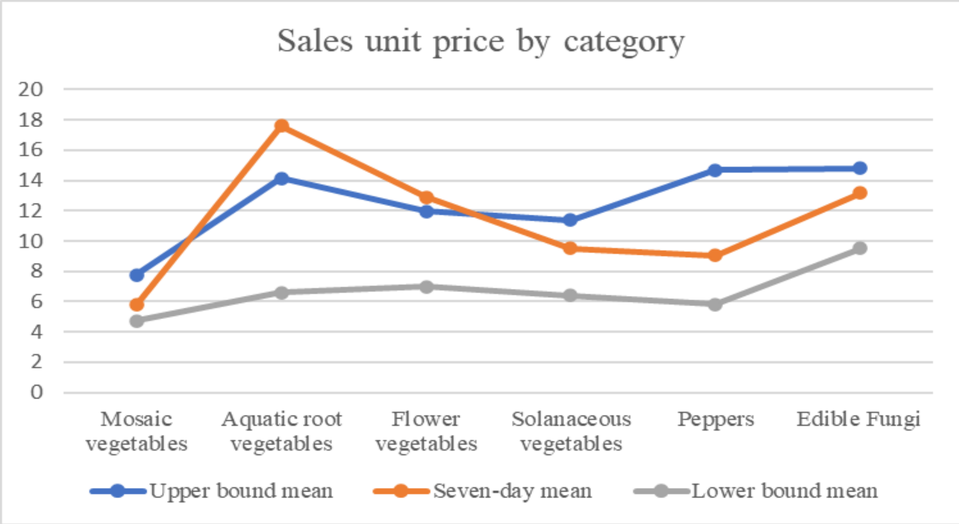


Figure 8. Comparison of unit sales prices

Where the upper bound of the general fluctuation is the mean+standard deviation of the three years of historical data, and the lower bound of the general fluctuation is the mean -standard deviation of the three years of historical data.

From Figure 7 and 8, It can be found that our predicted daily sales volume and sales unit price for the vegetable category are roughly within the mean±standard deviation of the three-year historical data. For the aquatic roots and tubers category, the unit sales price is slightly keep away from the historical average unit sales price profile ($17.58 for the next seven days, which is greater than the historical average±standard deviation of 14.14, which is $3.44 higher), suggesting that it is on the high side of the one-year price profile in early July. For other vegetable categories, the unit price and sales volume basically fit the historical data average, which means that the model predicts accurate and reasonable results.

4. Conclusions

In supermarket replenishment and pricing, predicting future sales demand and unit price is crucial. This paper uses a nonlinear planning model based on revenue maximization, constrained by ARIMA-predicted variance fluctuation range, to derive restocking volume and pricing strategy for the next seven days via particle swarm algorithm. Except for edible mushrooms, sales volume of all vegetable categories is increasing, suggesting rising demand, and supermarkets can accordingly increase future pricing. Focus should be placed on replenishing flowers, leaves, chili peppers, and edible mushrooms, with a relatively obvious demand increase. Finally, ARIMA effectively utilizes historical sales data, providing accurate predictions that fit the data and its changes, offering a reliable reference for future supermarket replenishment and pricing.

References

[1] Kayikci Y, Demir S, Mangla S K, Subramanian, N., & Koc, B. Data-driven optimal dynamic pricing strategy for reducing perishable food waste at retailers[J]. Journal of Cleaner Production, 2022, 344: 131068.

[2] Bridgemohan P, Mohammed M. Cold Chain Operations in the Caribbean: Opportunities and Challenges[J]. Cold Chain Management for the Fresh Produce Industry in the Developing World, 2021: 165-173.

[3] Priyadarshi R, Panigrahi A, Routroy S, & Garg, G. K. Demand forecasting at retail stage for selected vegetables: a performance analysis[J]. Journal of Modelling in Management, 2019, 14(4): 1042-1063.

[4] He Z, Yu S. Application of LightGBM and LSTM combined model in vegetable sales forecast[C]//Journal of Physics: Conference Series. IOP Publishing, 2020, 1693(1): 012110.

[5] Kohli S, Godwin G T, Urolagin S. Sales prediction using linear and KNN regression[M]//Advances in Machine Learning and Computational Intelligence: Proceedings of ICMLCI 2019. Singapore: Springer Singapore, 2020: 321-329.

[6] Ranjitha P, Spandana M. Predictive analysis for big mart sales using machine learning algorithms[C]//2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2021: 1416-1421.

[7] Minghua W, Qiaolin Z, Zhijian Y, & Jingui, Z. Prediction model of agricultural product's price based on the improved BP neural network[C]//2012 7th International Conference on Computer Science & Education (ICCSE). IEEE, 2012: 613-617.

[8] Guo Q, Luo C, Wei Q. Prediction and research on vegetable price based on genetic algorithm and Neural network model[J]. Asian Agricultural Research, 2011, 3(1812-2016-143577): 148-150.

[9] Ensafi Y, Amin S H, Zhang G, & Shah, B. Time-series forecasting of seasonal items sales using machine learning–A comparative analysis[J]. International Journal of Information Management Data Insights, 2022, 2(1): 100058.

[10] Ho S L, Xie M. The use of ARIMA models for reliability forecasting and analysis[J]. Computers & industrial engineering, 1998, 35(1-2): 213-216.

[11] Wang C C, Chien C H, Trappey A J C. On the application of ARIMA and LSTM to predict order demand based on short lead time and on-time delivery requirements[J]. Processes, 2021, 9(7): 1157.

[12] Zhang X, Sun W, Xue M, & Lin, A. Probability-optimal leader comprehensive learning particle swarm optimization with Bayesian iteration[J]. Applied Soft Computing, 2021, 103: 107132.