Vegetable pricing and replenishment decisions based on the optimization model

Haoru Jiang[1, *, #], Xin Chen[2, #], Junyao Fang[3, #]

[1]College of Integrated Circuit Science and Engineering (Industry-Education Integration School), Nanjing University of Posts and Telecommunications, Nanjing, China, 210023

[2]School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, 210023

[3]Portland Institute, Nanjing University of Posts and Telecommunications, Nanjing, China, 210023

*Corresponding author: simple_huracan@163.com

#These authors contributed equally.

Abstract. Vegetables are one of the essential foods in human daily life. Due to the short shelf life, it is difficult to preserve, the supermarket needs to timely and reasonable purchases and pricing. This paper aims to establish a model that can predict vegetable price fluctuations and develop reasonable pricing and replenishment strategies. According to the data given from a wet market, this paper studies the pricing and replenishment decisions of supermarket vegetables. By establishing the SARIMA model and Markov forecast model, the wholesale price of the same project on July 1 is calculated based on the wholesale price forecast of the previous week, and reasonable market demand is predicted. It is found that the predicted results have a reasonable error with the actual situation, and the accuracy is relatively high. The establishment of these two models is of great significance for market regulation. The model has good predictability and practical value, with good dynamic predictive ability, suitable for a reasonable allocation of resources in supermarkets and other trading places to obtain maximum benefit.

Keywords: ARIMA, Markov prediction, Optimization model.

1. Introduction

Vegetable commodities are one of the necessary foodstuffs in human daily life, and they are also hot commodities in supermarkets, farmers' markets, and other retail places. The shelf life and preservation conditions of vegetables cause them to be easily damaged, which also makes the supermarkets need to purchase and replenish vegetables in a timely and reasonable manner. In order to obtain the optimal profit, the wholesale and sales of vegetables need to be planned according to people's needs. In order to obtain the optimal profit, both wholesale and sales of vegetables need to be planned according to people's demand [1]. Therefore, by considering historical sales data, accurately predicting the trend of vegetable demand [2], and establishing an efficient supply chain network, the pricing and replenishment decisions of vegetable products can be improved to achieve the goals of maximizing sales and minimizing costs [3].

Mao Lisha in the "Supply Chain Perspective of Vegetable Wholesale Market Pricing Strategy and Production and Marketing Mode Research" in the article mainly use ARIMA model [4] for vegetable wholesale market pricing prediction. From the point of view of the modeling in this paper, the model takes into account the factors that affect the prediction results, and the data is real and reliable, which has a high reference value. In addition, from the prediction results of the model, the price prediction value is not much different from the actual value, the prediction trend is generally consistent with the actual, and the price difference is controlled within one yuan, with a certain degree of error [5].

ARIMA models are affected by seasonality and are very sensitive to outliers [6], while they may face difficulties in modeling and predicting dynamic and fast-changing systems, such as the New Crown epidemic [7]. Markov models [8], on the other hand, can better handle nonlinear relationships [9], and as vegetable prices fluctuate significantly in the short term due to weather disasters, disease outbreaks, or other emergencies, Markov models may perform more robustly to such anomalies [10]. Thus, under the influence of the new crown epidemic, for the short-term vegetable price prediction of about one week in this paper, the Markov model has high practicality and accuracy. At the same time, the Markov prediction model established in this paper can achieve better prediction results for time-related applications, such as stock market fluctuations, weather forecasting, etc., and has a wide range of applications and high feasibility.

2. Model

2.1. SARIMA Model

The ARIMA model, first proposed by Box and Jen-kins in the 1970s, is called the differential autoregressive moving average model and is denoted as ARIMA (p, d, q), where p and q represent the order of autoregressive and moving average respectively. d is the trend difference order, and the basic model structure is as follows:

$$\Psi(\mathrm{B})\nabla^{d} x_{t}=\Theta(B)\varepsilon_{t} \tag{1}$$

where: $\Psi(\mathrm{B})=1-\Psi_{1}\mathrm{B}-\cdots-\Psi_{p}B\nabla^{d}=(1-B)^{d}$ represents as a p-order autoregressive coefficient polynomial, B is the delay operator. ∇^{d} is for d-order difference. $\Theta(\mathrm{B})=1-\theta_{1}B-\cdots-\theta_{q}B^{q}$ is for q-order moving average coefficient polynomial.

The ARIMA model can be used to deal with the unsteady temporal order without seasonal effect column modeling, while in real life most of the time sequential columns exist in a given week period, for both seasonal effect and long-term potential effect of both SARIMA can be used for sequences with complex cross-cross interactions model.

In the SARIMA model, seasonal and trend parameters are important components. Seasonal parameters were used to describe the data over a given season, while trend parameters describe the overall trend of the data over time.

During the modeling process, the SARIMA model first tests the stationarity of the time series data to ensure that the data fits the assumptions of the ARIMA model. Then, a complete SARIMA model was obtained by modeling the seasonality and trend.

The advantage of the SARIMA model is that it can consider both the seasonality and trend of the data, improving the accuracy and stability of the prediction. In addition, the SARIMA model can be used in combination with other models, such as combining with neural network model, and support vector machine model, to further improve the prediction performance.

It should be noted that the SARIMA model has a limited scope of application, which requires the obvious seasonality and trend of the data. For data that do not meet these requirements, additional time series analysis methods may be required.

Here is a simplified formula derivation process:

First, the time series data satisfy the stationarity condition that its mean and variance do not change over time.

Differential operations were performed on the time series data to eliminate trending and seasonal effects. Autocorrelation and partial autocorrelation were analyzed to determine the parameters of the ARIMA model.

The parameters P, D, Q, p, d, q, and m of the SARIMA model were determined based on the results of the autocorrelation and partial autocorrelation analysis.

The SARIMA model was constructed with the following formula:

$$\mathrm{X}(t)=\sum_{i=1}^{p}\phi_{i}X(t-i)+\sum_{j=1}^{q}\theta_{j}X(t-j)+\sum_{k=1}^{m}\omega_{k}X(t-k) \tag{2}$$

where: $X(t)$ represents the time series data, and ϕ_i, θ_j and ω_k represent the parameters of the ARIMA model and the SARIMA model, respectively.

The model was fit and predicted, and its parameters were solved by minimizing the sum of square prediction errors.

2.2. Markov Prediction Model

Markov prediction model is a stochastic process model based on Markov chains for predicting future states or trends. It predicts future states based on current state and historical data through a probability transfer matrix. Markov models are suitable for sequences with discrete states, such as weather forecasts, stock prices, etc. In a Markov model, the future state is only related to the current state and not to other historical states, i.e. "no aftereffect". Markov models can help us understand and predict the dynamic behavior of a system so that we can make better decisions.

The Markov refinement of the prediction results is as follows: The model predicts the future development of the system based on the transfer probabilities between the states. $n-t$ The vector of state probabilities after a few moments is.

$$X(\mathrm{n}) = x(t)P(1)^{n-t} \tag{3}$$

where: $x(t)$ is the state probability vector at the initial moment t; $P(1)$ is the further state transfer probability matrix.

Classification of states. The sequence is divided into states based on relative values, denoted as $E_1,\ E_2,\cdots,\ E_\mathrm{i},\cdots,E_n,(E_i \in [M_k, N_k], k = 1,2,\cdots,m)$ grey elements M_k, and N_k represents the upper and lower bounds of the k state, respectively.

Construct the state transfer probability matrix. The probability of a data sequence transferring from state E_i to state E_j after the step m is denoted:

$$P_{\mathrm{ij}}^{(m)} = \frac{n_{ij}^{(m)}}{N_\mathrm{i}} \tag{4}$$

where: $n_{ij}^{(m)}$ is the number of times the state E_i has been transferred to the state E_j after m; N_i is the number of times the state E_i has appeared.

Calculate each data in the data array in 1) to obtain the state transfer probability matrix:

$$P^{(m)} = \begin{pmatrix} P_{11}^{(\mathrm{m})} & P_{12}^{(m)} & \cdots & P_{1n}^{(m)} \\ P_{21}^{(m)} & P_{22}^{(m)} & \cdots & P_{2n}^{(m)} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n1}^{(m)} & \cdots & P_{nn}^{(m)} \end{pmatrix} \tag{5}$$

Determining state steering. The closest k times to the prediction time (where k the number of states divided) are selected according to their distance from the prediction time, and then the number of steering steps are $1-k$ steps, respectively.

Determine the predicted value. After determining the state steering, the interval of variation of the relative value of the predicted value $[M_k, N_k]$ is also determined:

$$\widehat{Y} = \frac{1}{2}(M_k + N_k)\widehat{x}_i^{(0)} \tag{6}$$

where: $\widehat{x}_i^{(0)}$ is the predicted model value at i; M_k、N_k is the upper and lower bounds of the state interval.

3. Results

A seasonal index forecasting model was developed using Matlab and Python software to predict the wholesale prices of each vegetable category.

3.1. Analysis of experimental results by SARIMA Model

The predicted values have a good fit to the true values and the model is accurate enough to use this model for wholesale price prediction.

The predicted values of daily wholesale prices for each category obtained from the above model are shown in Table.1.

Table.1. Forecasts of daily wholesale prices by category

Date	cauliflower	philodendron	chilli	eggplant	edible mushroom	Aquatic rhizomes
2023-7-1	7.86	3.13	4.68	4.936	7.83	12.65
2023-7-2	7.89	3.12	4.69	4.937	7.75	12.63
2023-7-3	7.92	3.12	4.65	4.938	7.68	12.61
2023-7-4	7.91	3.11	4.63	4.939	7.52	12.55
2023-7-5	7.90	3.10	4.60	4.940	7.43	12.51
2023-7-6	7.92	3.08	4.61	4.941	7.39	12.50
2023-7-7	7.92	3.09	4.62	4.945	4.25	12.48

It can be seen that the prices of cauliflower, pepper, and cigars rose slowly, while the price of fungi plunged on July 7, and the wholesale prices of other dishes generally showed a downward trend. Predicting the rising and falling trend of vegetable wholesale prices, you can develop a reasonable procurement plan.

An optimization model with profit as the objective is developed. The daily sales volume is the number of vegetables sold by the supermarket in a day, and the daily replenishment volume is the number of vegetables purchased by the supermarket from suppliers every day to meet the sales demand. These two indicators are related to each other when building the model. The total daily replenishment amount of the supermarket is determined by the daily sales volume, and the supermarket should ensure that the replenishment amount is sufficient while avoiding excessive loss due to overstocking.

3.2. Analysis of experimental results by Markov Model

To give the replenishment quantity and pricing strategy of a single product on 1 July to achieve maximum benefit, the purchase price on that day, and the wastage rate need to be considered. Then it is necessary to use the sales data of the week before that day as the basis for the prediction of the wholesale price of individual products and market demand, if the seasonal index prediction model is used, the amount of data is too large, the prediction bias is large, so the Markov prediction model is introduced [7], the use of the Markov chain of the absence of posterior validity, using the previous week's data prediction only, the prediction bias is small, and it is more accurate.

According to the model, the optimal combination is 29 vegetables and the maximum benefit gained by the superstore is $1,940, some of its pricing strategies (please see the annex for details) are shown in Table. 2 and Table.3.

Table.2. Pricing Strategy for Selected Individual Products

Item Name	Price	Discounted	Item Name	Price	Discounted
Colourful Peppers(2)	23.17	Yes	Chrysanthemum coronarium	25.86	Yes
Yunnan oilseedrape (portion)	7.61	Yes	cabbage (round cabbage)	8.29	Yes
Yunnan lettuce	8.74	Yes	Ginger, Garlic & Millet Pepper Combo (Small portion)	8.11	Yes
Yunnan lettuce (portion)	4.30	Yes	Peppers (portions)	6.03	Yes
Round Eggplant(2)	8.86	Yes	snow fungus (Tremella fuciformis)	8.19	Yes

Table.3. Table of replenishment quantities of selected individual products

Product Name	replenishment	Item Name	replenishment
Colourful Peppers(2)	2.50	Chrysanthemum coronarium	2.50
Yunnan oilseed rape (portion)	19.16	cabbage (round cabbage)	5.34
Yunnan lettuce	3.08	Ginger, Garlic & Millet Pepper Combo (Small portion)	6.30
Yunnan lettuce (portion)	29.06	Peppers (portions)	19.29
Round Eggplant(2)	2.50	snow fungus (Tremella fuciformis)	5.34

To better display the data, the image is shown in Figure.1.

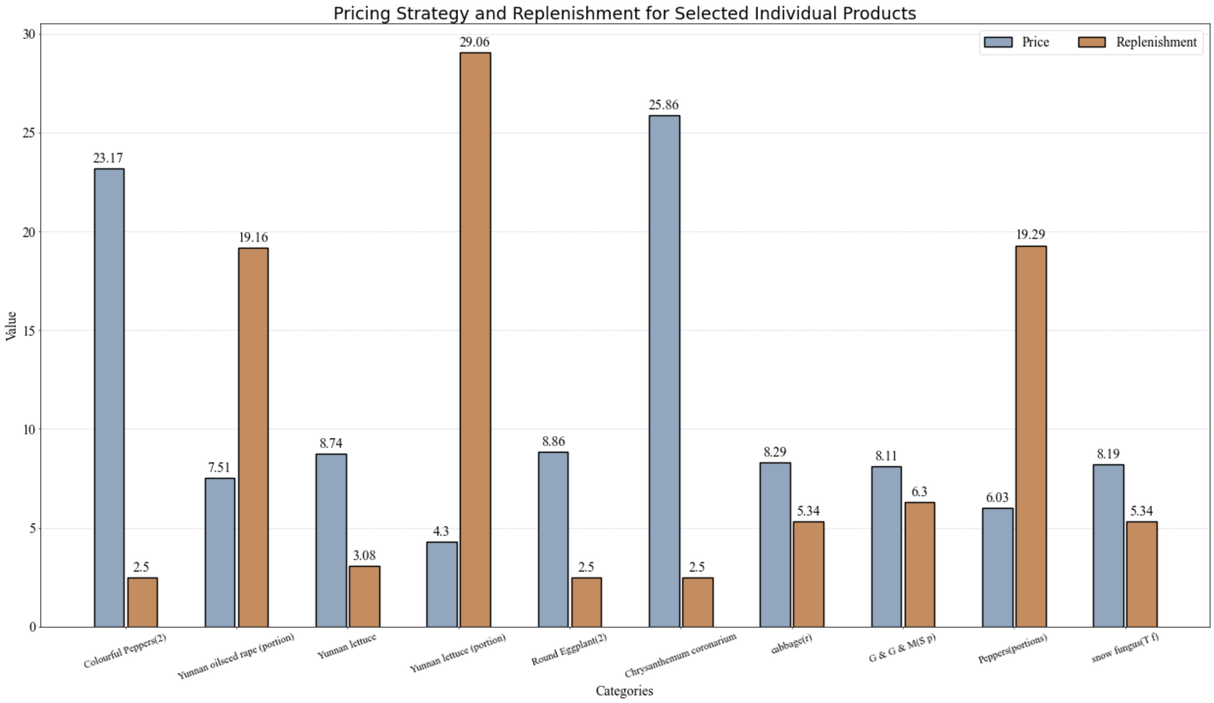

Figure.1. Pricing Strategy and Replenishment for Selected Individual Products

Figure 1 shows the pricing and replenishment strategies of various kinds of vegetables. It can be seen that the pricing is roughly negatively correlated with the replenishment volume. Vegetables with large replenishment quantities generally have low prices, with small profits and quick turnover. For example, Yunnan lettuce; and vegetables with small replenishment quantities are generally priced higher. For example, color pepper. Pricing strategies and replenishment decisions are better solved through the prediction of uncontrollable variables combined with the optimization of decision variables. Meanwhile, for different data requirements, the seasonal index prediction model and Markov prediction model are chosen to predict the data with better accuracy.

4. Conclusions

Consequently, supermarkets are required to make timely and sensible purchases of these goods, considering the challenges in preserving them and preventing spoilage. The procurement and pricing of vegetable commodities need to be reasonable according to market demand and weather conditions, but it is difficult to allocate resources and arrange procurement rationally based on experience alone, sometimes causing huge losses. After collecting enough data, the powerful data processing function of Markov and ARIMA model is used to forecast. In this paper, the data is processed and analyzed, and combined with the background environment, different models are used to process, and the predicted price is in high agreement with the actual price. In the actual comparison, the model has good predictability and practical value, for the rational allocation of resources in supermarkets and other trading places, in order to obtain the maximum benefit.

Reference

[1] Waugh F V. Quality factors influencing vegetable prices[J]. Journal of farm economics, 1928, 10(2): 185-196.

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

[3] Chong E K P, Lu W S, Żak S H. An Introduction to Optimization: With Applications to Machine Learning[M]. John Wiley & Sons, 2023.

[4] 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: 47-52.

[5] Mao L, Huang Y, Zhang X, et al. ARIMA model forecasting analysis of the prices of multiple vegetables under the impact of COVID-19[J]. PLoS One, 2022, 17(7): e0271594.

[6] Alsharif M H, Younes M K, Kim J. Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea[J]. Symmetry, 2019, 11(2): 240.

[7] Alabdulrazzaq H, Alenezi M N, Rawajfih Y, et al. On the accuracy of ARIMA-based prediction of COVID-19 spread[J]. Results in Physics, 2021, 27: 104509.

[8] Davis M H A. Markov models & optimization[M]. Routledge, 2018.

[9] Ye J, Xu Z, Gou X. An adaptive Grey-Markov model based on parameters Self-optimization with application to passenger flow volume prediction[J]. Expert Systems with Applications, 2022, 202: 117302.

[10] Marfak A, Achak D, Azizi A, et al. The hidden Markov chain modelling of the COVID-19 spreading using Moroccan dataset[J]. Data in brief, 2020, 32: 106067.