

# Research On Vegetable Replenishment and Pricing Optimization Decision Based on Time Series Prediction and Monte Carlo Simulation

Songling Fan \*

School of Computer, Electronics and Information, Guangxi University, Nanning, China, 530004

\* Corresponding Author Email: 2107310126@st.gxu.edu.cn

**Abstract.** For large fresh supermarket, how to reasonably specify the replenishment and pricing strategy of vegetable commodities is very important. In order to give a reasonable replenishment decision, this paper conducts correlation analysis at first, including cost plus pricing strategy and correlation test. And then analyzes the relationship between sales volume and selling price by using neural network regression and scatter plot matrix, and finds that there is no obvious functional correlation between them. The paper use time series forecasting methods, including simple seasonal analysis and Winter multiplication model, to forecast sales volume and wholesale cost. Finally, the linear programming model was used to determine the total amount of replenishment and markup rate of each category of vegetables in the next week, and the corresponding income situation was obtained. The results of comprehensive analysis show that the replenishment volume, markup rate and income of different categories have dynamic fluctuations, reflecting the changes of demand and market.

**Keywords:** Vegetable sales, seasonal models, linear programming, Monte Carlo Method.

## 1. Introduction

Due to the short shelf life of vegetables and the deterioration of products with the increase of time, supermarkets need to make replenishment decisions on the day according to historical sales and demand, but they cannot clearly know the price of incoming items and each vegetable when purchasing [1]. Vegetable pricing adopts cost-plus pricing, and the spoiled goods will be sold at a discount. Therefore, it is particularly important to forecast market demand, which can support supermarkets to make better replenishment and pricing decisions. From the perspective of demand, vegetable sales will be affected by time; From the perspective of supply, the supply of vegetables from April to October is large, but it is limited by sales space, so it is necessary to optimize and choose a reasonable sales mix. In order to give a reasonable replenishment decision, we first consulted some relevant literature. Fengyan Fan uses outlier detection to process data and uses machine learning to predict fresh food prices [2]. Junhua Han systematically introduced the details of cost-plus pricing in his article [3]. To solve our problem, firstly, the relationship between the total sales volume of different categories of vegetables and the cost-plus pricing is analyzed, and the relationship between the two and the total sales volume is preliminarily explored by using linear regression method. Secondly, the nonlinear relationship between them is explored by using neural network regression method [4]. Then the quantity of replenishment and markup rate are taken as variables, and the maximum profit is taken as the target. The linear programming model is established, and the wholesale price and sales volume are predicted by time series forecasting [5].

## 2. The analysis and solution of problem

### 2.1. The correlation analysis

Let's start with cost-plus pricing. Cost-plus pricing is a commodity pricing strategy commonly used to determine the selling price of a product or service. By looking at the data, the core idea of this pricing strategy is to add the cost of the product or service to a specific percentage (i.e., the margin

markup rate) to determine the final selling price, and this margin markup rate = (selling price/wholesale price - 1) \* 100%.

Now let's do a correlation test. First, plot the scatter plot of the matrix. According to the requirements of the question, the analysis is carried out by category, and the daily sales sum data of each category obtained from the first question is used to calculate the daily selling price of each category accordingly. Then, for each vegetable category, the data of two variables, the average daily selling price and the total daily sales volume, were extracted and made into a scatter plot matrix for correlation analysis to observe whether there is a trend correlation between the two key factors, sales volume and selling price.

In view of the possible nonlinear relationship, the neural network regression method was further used to study the complex relationship between the two variables, and the fitting curves and correlation coefficient R2 of the two relationships were obtained respectively. The specific results are shown in Figure 1, Figure 2 and Table 1.

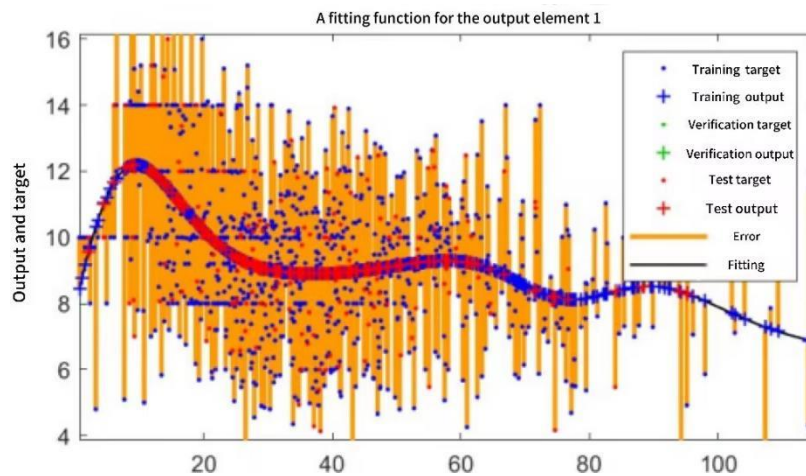


Figure 1 Sales volume - price curve fitting plot

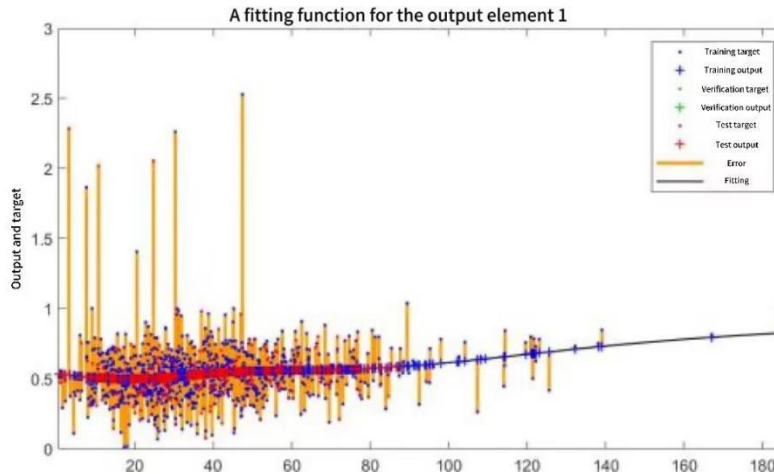


Figure 2 Sales volume - mark-up rate curve fitting plot

Table.1. Training results

	Volume - Markup rate	Volume - Price
R <sup>2</sup>	0.42	0.19

From the direct observation of the scatter plot, there is no clear functional correlation between the two variables. At the same time, it is found that the R2 of the two training sets are 0.42 and 0.19 respectively, considering that there is no obvious functional relationship between sales volume and cost-plus pricing for this topic and the data in this topic. In order to verify this point precisely, Matlab tools are then used to fit multiple groups of functions to the data of two variables, including linear,

logarithmic, exponential and other functions, but the results of correlation coefficient and significance test are not ideal. After comprehensive analysis, it can be concluded that at the current category level, it is difficult to establish an accurate mathematical correlation model between sales volume and selling price, which may be due to the influence of individual factors of different varieties on the results after the summary of multiple categories.

## 2.2. Time series prediction

Time series prediction is a method to predict the change law of data over time [6-7]. The title of this paper states that the sales volume of vegetables on the demand side has a certain relationship with time. In addition, through consulting the line chart of vegetable prices in China over the past ten years in Baidu Index, it is found that under normal circumstances, vegetable prices also change with time. Therefore, for the two unknowns of sales volume and wholesale cost, we use the method of time series prediction to determine them.

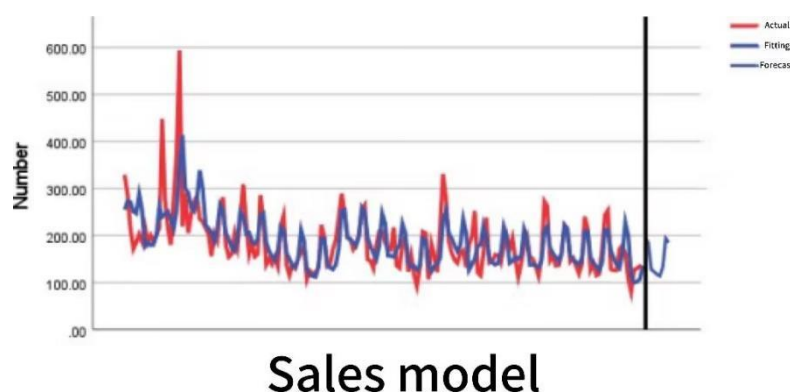
Simple seasonality [8]. Simple seasonal analysis is a method used to study seasonal fluctuations in time series data. It aims to identify and quantify seasonal patterns in data, such as periodic changes from year to year, quarter to quarter, or month to month, and is an important pre-processing step in time series forecasting.

Winter multiplication model [9]. Winter multiplication model is a time series decomposition method, which is used to decompose time series data into trend, seasonality and residual components. It breaks down raw data into these components to better understand and predict time series.

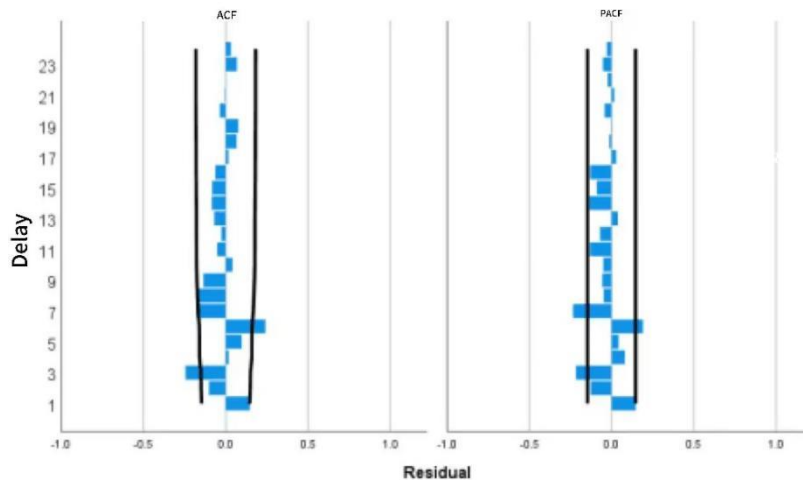
ACF (auto correlation coefficient). ACF is a statistical tool used to measure autocorrelation in time series data. It calculates correlations between data points with different time lags (lags), helping to identify periodicity and trends in time series.

PACF (partial auto correlation coefficient). PACF is often used to determine the order of lag of an auto regressive model (AR). It can help identify lag items in time series to better model and predict data.

Now let's make a sales forecast. Import the sale-time data of vegetable categories obtained in the previous question into SPSS, and the results are shown in Figure 3, Figure 4 and Table 2. It is observed that the images can capture the trend, seasonality and periodicity, and there is a significant correlation structure in the ACF and PACF charts, and the stable R2 is close to 1, so the prediction results can be regarded as good. (Due to the length of the article, only the prediction results of the Mosaic category are shown here.)



**Figure 3** Forecast results of sales volume of Mosaic category

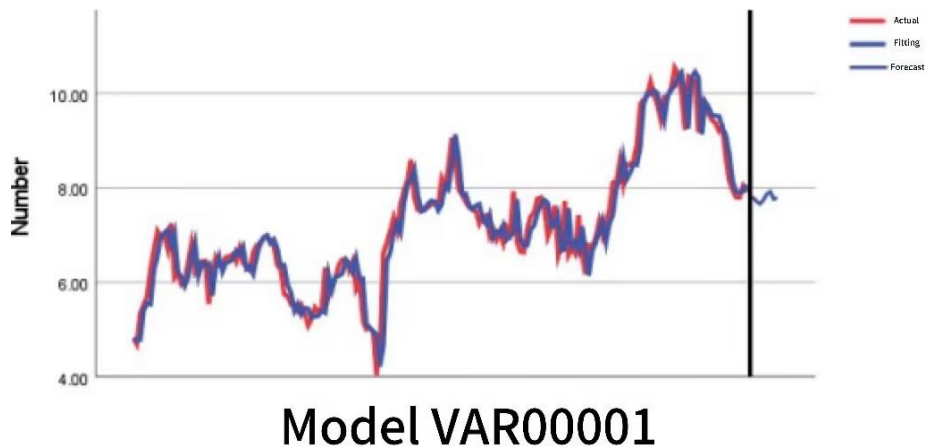


**Figure 4** Residual analysis of Mosaic category

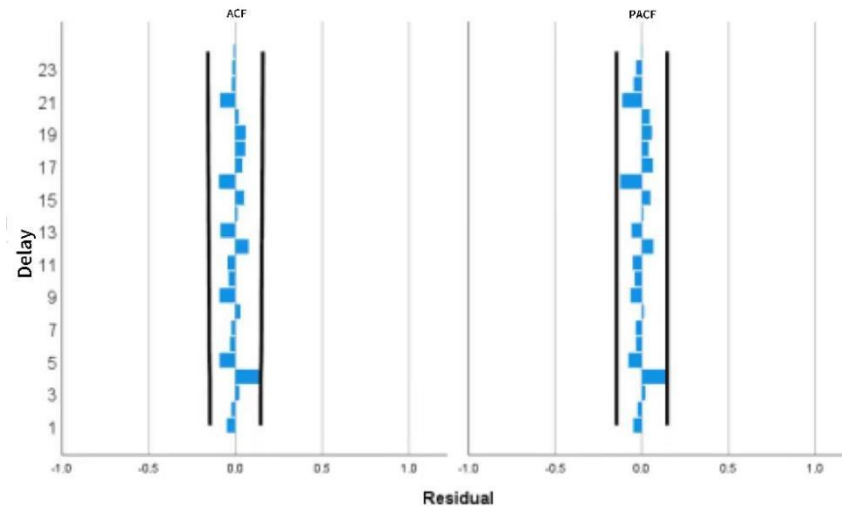
**Table.2.** Results of model fit degree

	Volume
Stationary R <sup>2</sup>	0.527

Next, make a cost forecast. Import the wholesale price-time data of vegetable categories obtained in the previous question into SPSS, and the results are shown in Figure 5, Figure 6 and Table 3. It can be observed that the trend, seasonality and periodicity are captured, and there is a significant correlation structure in the ACF and PACF charts, and the stable R2 is close to 1, so the prediction results can be regarded as good. (Due to the length of the article, only the prediction results of the Mosaic category are shown here.)



**Figure 5** Forecast results of Mosaic cost



**Figure 6** Mosaic cost residual analysis

**Table.3.** Results of model fit degree

	Cost
Stationary R2	0.707

### 2.3. Linear programming

Linear programming is an important branch of mathematical programming in operations research. Its objective function and constraint condition are linear functions [10]. Aiming at the total quantity of replenishment and pricing strategy, taking replenishment quantity and markup rate as variables and maximum profit as target, the following linear programming model is established:

First, define the selling price of the wasting goods:

$$W_1 = P_2NP_3S \tag{1}$$

Selling price of unworn goods:

$$W_2 = 1 - P_2NS \tag{2}$$

Commodity cost price:

$$W_3 = AM \tag{3}$$

Then, income:

$$W = W_1 + W_2 - W_3 = P_2NP_3S + 1 - P_2NS - AM \tag{4}$$

Price:

$$S = A*1 + P_1 \tag{5}$$

Where,  $P_1$  represents the markup rate,  $P_2$  represents the loss rate,  $P_3$  represents the discount,  $A$  represents the cost,  $N$  represents the sales volume, and  $M$  represents the replenishment volume. Here we set the variables:  $P_1, M$  for  $x_1, x_2$ . Then:

$$\begin{aligned} W &= P_2P_3NA(1+x_1) + (1-P_2)NA(1+x_1) - Ax_2 \\ &= (P_2P_3 + 1 - P_2)NAx_1 - Ax_2 + (P_2P_3 + 1 - P_2)NA \end{aligned} \tag{6}$$

Constraints:

1. The addition rate is greater than the minimum value of the historical information and smaller than the maximum value of the historical information, that is,  $MIN \leq x_1 \leq MAX$ .

- 2, the replenishment quantity is greater than or equal to the sales volume, that is,  $x_2 \geq N$ .
3. The replenishment quantity is less than or equal to the maximum historical sales volume, that is,  $x_2 \leq MAX$ .

The final integration into the following model:

$$MAX \quad W = (P_2P_3 + 1 - P_2)Nax_1 - A_1x_2 + (P_2P_3 + 1 - P_2)NA \tag{7}$$

$$s.t. \begin{cases} MIN \leq x_1 \leq MAX \\ N \leq x_2 \leq MAX \end{cases} \tag{8}$$

### 3. Results

#### 3.1. The solution result

The prediction results of the previous time series of wholesale price and sales volume were substituted into the equation as known parameters of each day in the next seven days, and the problem was solved by using Matlab to solve the function of linear programming, and the total replenishment amount and pricing strategy of different categories of vegetables in the next seven days were obtained, as shown in Table 4 below (due to the limitation of the length of the article, only the prediction results of flowers and leaves were shown here).

**Table.4.** Total replenishment, markup rate and income of Mosaic category each day in the coming week

	Total daily replenishment	Markup rate	income
2023/7/1	193.82	0.838779092	506.4904437
2023/7/2	185.1234239	0.838554199	471.9194537
2023/7/3	128.0895729	0.838322984	315.3375102
2023/7/4	123.5287242	0.83867091	283.2790414
2023/7/5	117.5669172	0.838438644	284.2082281
2023/7/6	114.390952	0.837774939	285.9937182
2023/7/7	133.8013057	0.838535513	322.1027761

#### 3.2. Analysis of experimental results

By analyzing the replenishment volume and markup rate of each vegetable category in the next week, we have:

In terms of the total amount of replenishment, flowers and leaves are the largest, nightshade and aquatic roots are smaller. This reflects the difference in sales demand.

In terms of addition rate, flowers and leaves were the highest and aquatic rhizomes were the lowest.

In terms of income, Mosaic is the largest, nightshade and aquatic rhizomes are the smallest, which is related to the total demand for replenishment and the setting of addition rate.

The replenishment volume, markup rate and income of different categories all show a certain dynamic fluctuation, reflecting the demand and market changes.

### 4. Conclusions

According to the characteristics of fresh products, on the basis of cost-plus pricing method, we put forward how to formulate replenishment strategy to meet the needs of supermarkets. To this end, we first draw the scatterplot of total sales and average pricing of each category of vegetables, and the overall trend shows a weak negative correlation. In order to determine the replenishment and pricing strategy for the coming week, based on historical sales data and the relationship between sales volume and time, a simple seasonal model was used to forecast the sales volume and wholesale price for each

category from July 1 to July 7, 2023. Then, the total daily revenue is taken as the objective function and the variable of markup rate is introduced. The replenishment volume and markup rate of each category are calculated with the linear programming model and Monte Carlo algorithm. Here, the total daily replenishment volume and markup rate of cauliflower on July 1, 2023, are 31.11kg and 0.62 respectively. The results are consistent with the negative correlation between sales volume and cost-plus pricing.

## References

- [1] Zewei Chen. Study on inventory replenishment strategy of agricultural product retailers considering different behaviors [D]. Chongqing jiaotong university, 2023.
- [2] Fengyan Fan. Supermarket to live fresh commodity pricing model research [D]. Fujian normal university, 2019.
- [3] Junhua Han, Shengdao Gan. Cost plus pricing review [J]. Journal of accounting monthly, 2012, (22): 74-75.
- [4] Weige Yu, Huarui Wu, Cheng Peng. Research on combination model of short-term vegetable price prediction based on Lasso regression and BP neural network [J]. Intelligent Agriculture (Chinese and English), 2019,2(03):108-117.
- [5] Zhiyan Zhong, Yijun Fan, Haoran Liu. Suppliers based on the linear programming model of optimal strategy study [J]. Journal of enterprise reform and management, 2022, (13): 29-31.
- [6] CHENG J P. Research on vegetable price prediction in Henan Province based on ARMIA-BP neural network [D]. Henan Agricultural University,2019.
- [7] Nan Ding. Vegetable prices based on time series prediction and correlation analysis [D]. Shanxi agricultural university, 2022.
- [8] Guolin Wang. Prediction of hospital discharge based on seasonal exponential smoothing model [J]. Journal of Qiqihar Medical College, 2019,41(07):854-856.
- [9] Wang Haifeng. Prediction of primary premium income of property insurance based on Winters multiplication Model [J]. Shanghai Insurance,2020, (05):54-57.
- [10] Jianhui Tian, Ling Luo, Pan Shang. Construction of CO2 emission prediction model for refining and chemical enterprises based on linear programming technology [J]. World Petroleum Industry,2022,29(05):26-32.