Research On Supermarket Replenishment Decision Based on ARIMA Time Series Prediction and Combination Optimization Model

Peng Fang *, Xutao Duan #, Chengyuan Li #

College of Engineering, Qufu Normal University, Shandong, China, 276826

* Corresponding Author Email: 19956249075@163.com

Abstract. In order to maximize the daily profit of supermarkets, it is necessary to predict the sales volume of each category of vegetables in advance during replenishment to determine the variable cost of replenishment volume. This article uses the ARIMA time series prediction model and a nonlinear programming model with penalty functions to provide a combination optimization decision-making method. Under constraints such as supermarket space constraints and dish richness, it effectively solves the problem of specific daily replenishment total and maximum total profit per day for different products in supermarkets, as well as the problem of maximizing supermarket revenue.

Keywords: normalization BIC; ARIMA model; constraint factor; combination optimization model.

1. Introduction

The supply level and price level of vegetables have an important impact on agricultural economic benefits, national economic health and people 's health and stability [1]. The preservation period of general vegetable products is short. Therefore, fresh vegetable supermarkets will inevitably take corresponding insurance measures to extend the preservation period of vegetables, but their decay rate is still very high [2].Therefore, the sales will be based on historical sales data and the demand for vegetables to carry out daily replenishment. According to the analysis of relevant data, the market 's demand for different types of vegetables varies with different times of the week, and is also affected by holidays and weather conditions. Cao Qingkui et al.constructed a retailer 's inventory cost model under time-varying demand. Through a series of analyses, they concluded that retailers should adopt an equal-period replenishment model under stable demand conditions and a non-equal-period model under time-varying demand conditions [3].

Accurate market demand analysis is crucial to the supermarket 's replenishment. If the forecast is not accurate or the forecast model is not optimized in the replenishment process, it will inevitably bring some losses to the future profit of the supermarket [4]. In order to improve the accuracy of the prediction model, this paper makes a dynamic planning of a vegetable pricing, thus reducing the profit loss caused by static planning, and calculates the average cost markup pricing according to the wholesale price of different types of vegetables every day. On this basis, the corresponding penalty function linear regression equation is constructed to predict the law of cost pricing, so as to make the cost pricing more reasonable. In addition, in order to maximize the profit of vegetables placed in supermarkets, this paper also provides the placement plan of vegetable categories, in order to break through the limitation of supermarket space and increase the amount of vegetables placed as much as possible.

2. Supermarket Vegetable Daily Replenishment Forecast and Combination Optimization Model

2.1. Model Construction

ARIMA time series prediction model is a linear prediction model [5-9].For the daily replenishment of vegetables, the future sales of vegetables can be accurately predicted by observing the historical sales volume of each vegetable category and its error sequence curve. However, in daily life, the time series data of the sales volume of each category of vegetables are often not linear, and contain strong nonlinear characteristics and seasonal components. Therefore, before using the ARIMA time series prediction model, it is necessary to first decompose the huge seasonality and select the optimal prediction period, so as to predict the daily replenishment data with the best practical application guidance value. The dynamic optimization penalty function is constructed by using the relationship between the daily replenishment quantity and the average cost pricing of vegetables. The objective function is the maximum daily total profit of the supermarket and the main constraint factors are extracted : the number of items constraint factor, the display quantity constraint factor, the cost constraint factor and the loss rate constraint factor. In summary, a nonlinear programming problem with penalty function is constructed based on the daily replenishment forecast and combination optimization model of supermarket vegetables [10].

2.2. Solution Procedure

The whole model solving process is divided into the following steps :

Step1:Selection of the optimal fitting period

Before predicting the sales volume of various categories of vegetables, based on the purpose of predicting the most accurate sales results, this paper selects four different periods of day, week, month and quarter respectively, and uses the Akaike information criterion and goodness of fit R^2 to select the optimal fitting period. The specific solution formula is as follows:

$$\text{BIC}=\ln(T)\,(\tau) - 2\ln(\phi) \tag{1}$$

In formula (1), τ is the number of parameters in the model. ϕ is the maximum likelihood function value of the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{2}$$

Among them, y represents the sales volume of vegetables, the number of samples is represented by T, the number of parameters in the model is represented by n, reflecting the complexity of the model, and the maximum likelihood function value of the model reflects the degree of data interpretation (fitting) of the model.

Step2:Seasonal decomposition

Using SPSS software to decompose the experimental data seasonally : define the date and time according to the optimal fitting cycle, click ' analysis-time series prediction-seasonal decomposition ', and further analyze the seasonal bias according to the results.

Step3:Selection of the optimal fitting model

ARIMA(p,d,q) The time series prediction model is mainly composed of three parts : autoregressive model (AR), moving average model (MA) and differential inverse process (I), where p represents the order of autoregressive model, q represents the order of moving average model, and d represents the order of difference. The specific principal formula is as follows:

$$(1 - \sum_{i=1}^{p} \alpha_i L^i)(1 - L)^d y_t = \alpha_0 + (1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_t \tag{3}$$

Where α is the smoothing parameter of the level, β is the smoothing parameter of the trend, and L is the average sales lag factor.

Because the time series of most vegetable sales are not stable, it is necessary to use SPSS expert modeler to select ARIMA (p, d, q). After the results are derived by SPSS, the prediction results of the minimum normalized BIC and the optimal goodness of fit R^2 are selected for analysis.

Step4:Residual verification

In order to test the accuracy and reliability of the prediction results, it is necessary to test whether the sales volume residual is a random independent white noise sequence. Generally speaking, if a model has a good fitting effect, the residual of the model satisfies the normal distribution with a mean value of 0, and for any lag order, the residual sequence is a white noise sequence. The results are analyzed according to the residual ACF diagram and the residual PACF diagram.

Step5:Construct a dynamic optimization penalty function

The average cost markup pricing is calculated according to the wholesale price of each category of vegetables:

$$R_i = O_i(1 + m_i) \tag{4}$$

Among them, O_i is the average wholesale price of each category of vegetables, R_i is the average cost markup pricing of each category of vegetables, and m_i is the average cost markup rate of each category of vegetables. Finally, according to the calculated average cost markup pricing and vegetable sales, the dynamic optimization penalty function with the average cost markup pricing as the independent variable and the vegetable sales as the dependent variable is obtained.

Step6:The inverse iterates the objective function.

Due to the complex relationship between each variable taking the maximum daily total profit of the supermarket as the objective function, in order to obtain the objective equation of maximum profit with penalty function, an inverse iterative process was used to ultimately obtain the specific objective function.

The preliminary objective function is as follows:

$$SP = PS_i * R_i - Pr_i * O_i \tag{5}$$

Among them, SP represents the daily maximum total profit of Supermarket, and PS_i is the total daily sales volume of category i vegetables, Pr_i is the total daily replenishment amount of category i vegetables.

After determining the preliminary objective function, the inverse iteration process begins. First, the total daily replenishment volume of category i vegetables in the objective function is solved. Based on daily life, it can be seen that supermarkets always encounter situations where all vegetables are not sold out on the same day when purchasing, and most vegetables have a short shelf life. In order to ensure the quality of vegetables, they can only choose to sell them at a discount or discard them, which will inevitably lead to losses, taking into account the above, it can be concluded that the total daily replenishment of category i vegetables.

$$Pr_i = \frac{PS_i}{1-b_i} \tag{6}$$

Where b_i is the average loss rate of each category of vegetables. After determining the daily replenishment of vegetables, we solve the initial objective function and the only missing average wholesale price:

$$O_i = \frac{\sum_{j=1}^{n}(C_j * K_j)}{K_n} \tag{7}$$

C_j is the cost of each single vegetable, K_n is the total weight of each single vegetable, and K_j is the weight of each single vegetable.

Based on all the above formulas, the final objective function for the maximum daily total profit of Supermarkets is derived as follows:

$$SP_j = \sum_{j=1}^{t}\left[f(m_i) * O_i * (1 + m_i) - \frac{PS_i}{1-b_i} * \frac{\sum_{q=1}^{n}(C_q * K_q)}{K_n}\right] \tag{8}$$

Where t is the number of vegetable categories in the supermarket, and finally based on the constraint factor, the constraint equation is constructed, and the optimal replenishment plan of the supermarket can be obtained.

3. Experiments and Analysis

3.1. Sales forecast

Extract the relevant data based on the replenishment decision of Supermarket, sort out the sales data of Supermarket vegetables in recent years, and compare it with the replenishment scheme based on the prediction model.

Step1:Selection of the optimal fitting period

Predict the total sales volume of vegetables in the supermarket based on four different cycles: day, week, month, and quarter. Table 1 shows the goodness of fit stationary R^2 and normalized BIC for different periods.

Table 1 Goodness of fit stationary R^2 and normalized BIC for different periods

Fitting statistics	A period of one day	A period of one week	A period of one month	A period of one quarter
R^2	0.548	0.811	0.632	0.741
normalized BIC	7908.466	7.856	20.015	7.705

It can be seen from Table 1 that the goodness of fit R^2 is only about 0.6 with days as the cycle or months as the cycle, and the normalized BIC is large. The normalized BIC with one day as the cycle even reaches 7908.466, which is not conducive to the choice of time prediction model. Although the goodness of fit R^2 and normalized BIC with quarter as the cycle can better predict the sales volume of various vegetables, the actual reference significance is poor due to the small amount of data. Based on the above, we decided to predict the sales volume of various types of vegetables and select the model in a weekly cycle.

Step2:Seasonal decomposition

Seasonal decomposition is performed based on weekly data, and the results are visualized as shown in Figure 1:

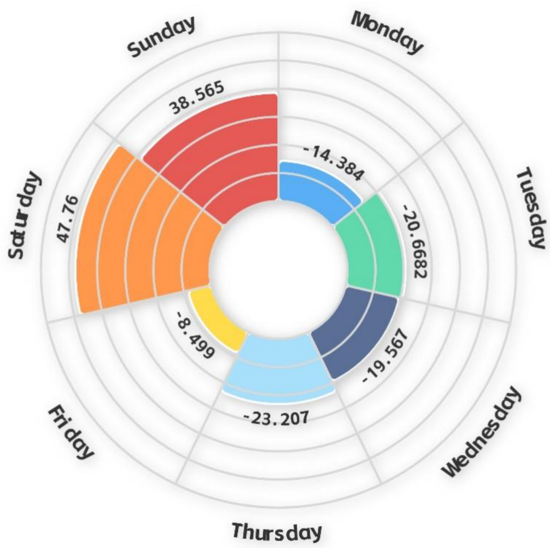

Figure 1 Seasonal factors of vegetable sales

From Figure 1, it can be seen that the seasonal factor from Monday to Friday is negative, which reflects that the sales volume of flower and leaf vegetables from Monday to Friday is lower than the annual average sales volume, while the seasonal factor on Saturday and Sunday is much greater than zero. This reflects that the sales volume of flower and leaf vegetables on Saturday and Sunday is much higher than the annual average sales volume, and the seasonal factor sizes of the six categories of vegetables also well explain the cyclical distribution patterns of different categories of vegetables。

Step3:Residual test and analysis of prediction results

Based on the vegetable sales data of the supermarket in recent years, residual testing was conducted using SPSS, and the test results are shown in Table 2 and Figure 2:

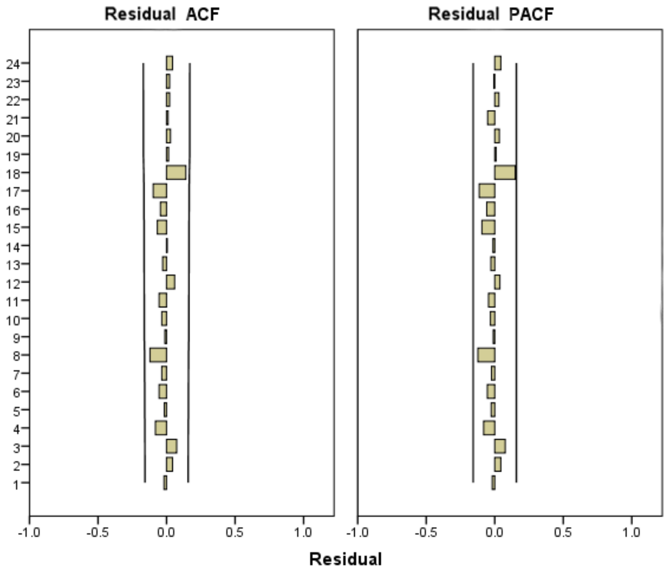

Figure 2 Residual ACF diagram and residual PACF diagram

Table 2 Model fitting degree statistics

Model	Number of forecast variables	R^2	Statistics	DF	Significance	Outlying values
ARIMA(0,1,1)	0	0.811	13.255	17	0.719	4

It can be seen from Figure 2 that the autocorrelation coefficient and partial autocorrelation coefficient of all lag orders of the sales volume of this category of vegetables are not significantly different from 0 at the 95 % confidence level. In order to further prove that the residual of the sales volume of this category is a white noise sequence, it can be seen from Table 2 that the p value is 0.719, which is much larger than 0.05, and the null hypothesis cannot be rejected. Therefore, it is concluded that the residual of this category of vegetables is a white noise sequence, and the sales volume of this category of vegetables tends to be stable after difference.

Finally, the forecast result of this category of vegetables on the same day is 124.439 kg. The time series of the sales volume of this category of vegetables is shown in Figure 3:

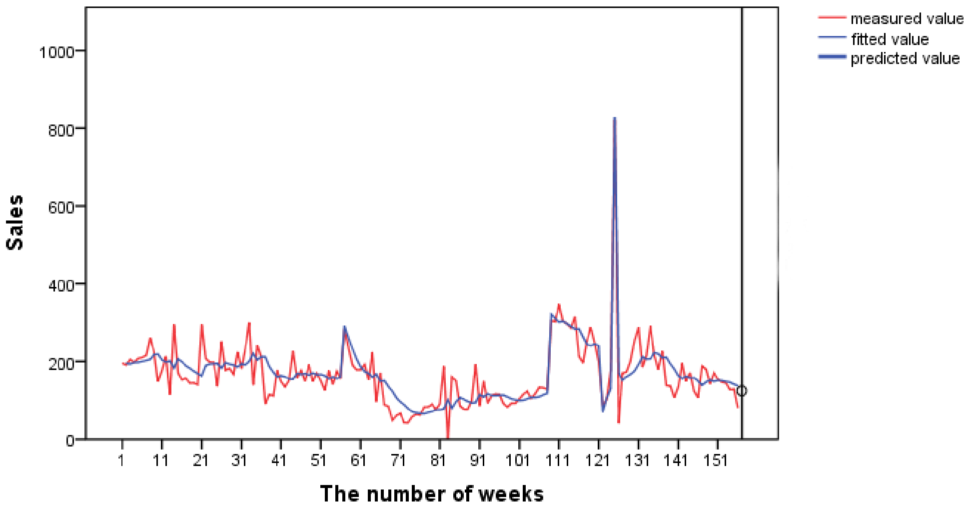

Figure 3 Vegetable sales time series diagram

It can be seen from Figure 3 that most of the fitted sales volume curves basically coincide with the actual sales volume curves without considering the data with strong subjective factors such as holidays and outliers. The model fitting degree is good, and the goodness of fit of the model is 0.811. The performance of the model is excellent, and the prediction results have practical reference application value.

Step4:Construct a dynamic optimization penalty function

This article summarizes and processes the sales data of the supermarket to obtain the average pricing and total sales of different single vegetable products. The data with a total sales volume of 0 is removed, and the processed average pricing and total sales data are used to obtain a penalty function linear regression equation about the average pricing and total sales volume using Matlab fitting toolbox:

$$\mathrm{f(x)} = -11.24 * \mathrm{x} + 1726 \tag{9}$$

From the linear regression equation, it can be seen that when the price of vegetables with smaller average price fluctuates, it will greatly affect the sales volume of vegetables and lead to losses. Therefore, the pricing of vegetables generally does not change significantly, which is also consistent with the market sales law.

Step5:Replenishment scheme under multi-constraint nonlinear programming

Combined with the actual operation situation of the supermarket, several constraint factors that have a great influence on the sales profit of the supermarket are extracted : single product quantity constraint factor, display quantity constraint factor, cost constraint factor and loss rate constraint factor. The constraints are listed as follows:

$$\begin{gathered} \sum_{i=1}^{t}(PS_i O_i * (1+m_i) - PR_I * O_i) \geq 0 \\ 0.0568 \leq b_i \leq 0.1918 \\ 27 \leq t \leq 33 \\ n_i \geq 2.5 \\ P_{ri}(i = 1 \dots \dots t) \geq 0 \end{gathered} \tag{10}$$

In the constraint formula: n_i is the display quantity of each vegetable item; b_i is the loss rate of each vegetable item; t is the number of single vegetables.

Based on the objective equation with penalty function and constraint conditions mentioned above, the multi constraint nonlinear programming model was solved using Matlab. The final solution result was that the optimal number of single item purchases on that day was 29, and the maximum total profit of supermarket vegetables on that day was 11076.28 yuan. The replenishment result is shown in Figure 4:

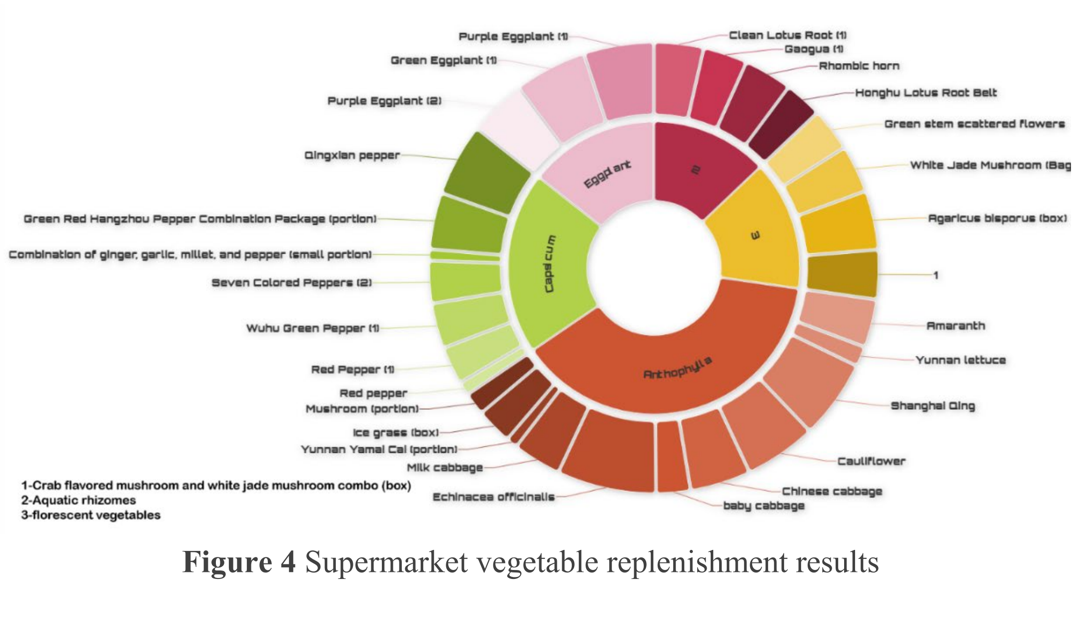

Figure 4 Supermarket vegetable replenishment results

From Figure 4, it can be seen that although only 29 categories are displayed, each individual vegetable is distributed among different categories, which not only meets the market demand for each category of vegetables, but also earns the maximum profit under practical constraints. The model has good expression effect and has certain practical reference application value.

4. Conclusions

In response to the practical problems of optimizing supermarket revenue and space constraints, the article improves the accuracy of data on predicting future vegetable sales volume of supermarkets through seasonal cycle analysis and ARIMA time series prediction model, thereby effectively solving the problem of specific daily replenishment total for different products in supermarkets. After collecting relevant information from multiple aspects, utilizing massive data support and analyzing the relevant data that should be collected by the supermarket, the constraint factors that have a significant impact on the supermarket's profits are extracted. Based on the pricing penalty function, an objective equation is constructed, and a dynamic combination optimization model is used to provide a good combination optimization decision-making method. Finally, the optimal purchase quantity and display plan are given. The experimental results show that the replenishment decision based on the dynamic combination optimization model under the optimal prediction not only brings significant profits to supermarkets but also provides the optimal display plan under space constraints. It also meets the supermarkets' pursuit of revenue and customers' pursuit of vegetable category richness, providing direction for supermarkets to better formulate replenishment and pricing decisions for vegetable products.

References

[1] LishaMao. Research on pricing strategy and production and marketing mode of vegetable wholesale market from the perspective of supply chain [D]. Central South University of Forestry and Technology,2022.

[2] ZingZhuo, GanhuaLong,PeiqiShen.Research status of storage and preservation technology of leafy vegetables at home and abroad [J].Anhui Agricultural Sciences,2013,41(15):6897-6899+6901.

[3] QingkuiCao,BingYu,QinLi Research on optimal decision-making of retailer replenishment time point under time-varying demand [J]. Journal of Hebei University of Engineering (Natural Science Edition),2007(03):64-67.

[4] YiyanHuang,Qixu Pricing decision optimization algorithm based on time-varying inventory model [J]. Statistics and Decision,2006(23):24-25.

[5] Jadhav V, CHINNAPPA R BV, Gaddi G M. Application of ARIMA model for forecasting agricultural prices [J]. 2017.

[6] J. Fattah, L. Ezzine, Z. Aman, H. El Moussami, A. Lachhab, Forecasting of demand using ARIMA model, Int. J. Eng. Bus. Manag. 10 (2018).

[7] Shumway R H, Stoffer D S, Shumway R H, et al. ARIMA models [J]. Time series analysis and its applications: with R examples, 2017: 75-163.

[8] Khandelwal I, Adhikari R, Verma G. Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition [J]. Procedia Computer Science, 2015, 48: 173-179.

[9] BoFan,WenqiSong Research on product sales forecast based on ARIMA model [J]. Industrial Control Computer,2021,34(05):128-129+125.

[10] Kuhn H W, Tucker A W. Nonlinear programming [M]//Traces and emergence of nonlinear programming. Basel: Springer Basel, 2013: 247-258.