ISSN 2616-5775 Vol. 6, Issue 13: 138-144, DOI: 10.25236/AJCIS.2023.061320

Research on vegetable pricing and replenishment strategy based on time series model and particle swarm optimization

Mengxiao Li, Wei Li, Zhixuan Zhao

School of Mathematics and Statistics, Liaoning University, Shenyang, 110036, China

Abstract: *The pricing and replenishment strategy of vegetables is very important for supermarkets, and accurate prediction and appropriate pricing replenishment strategies are of great significance for saving costs, improving supply chain management, and reducing the loss rate of vegetables. In order to accurately predict the pricing strategy and replenishment of vegetables every day in the coming week, based on the time series prediction model and particle swarm optimization algorithm, combined with the advantages of MATLAB in processing cleaning data, the pricing and replenishment strategy model that maximizes the revenue of supermarkets was constructed. Finally, based on the data, it is concluded that the maximum income of the supermarket in the coming week is 17108 yuan, and the maximum income per day is 2108.20 yuan.*

Keywords: *Time series, Particle swarm optimization algorithm, Vegetable replenishment*

1. Introduction

In daily life, vegetables are seasonal commodities, and in the process of vegetable production and sales, reasonable price control to meet the living needs of the people while improving the income of operators is very important. In simultaneous fresh supermarkets, vegetable products have a short shelf life, and the quality of vegetables deteriorates with the increase of sales time,[1-2].

With the increase of the cost of living, price and freshness have become the main factors for consumers to measure the quality of products, and this study establishes an appropriate mathematical model based on the sales details of different vegetable categories at different times, including wholesale prices and loss rates, to help merchants solve the problem of pricing and replenishment of vegetable commodities [3,4].

2. Materials and Methods

2.1 Data acquisition and preprocessing

The data mainly comes from market research and statistics, and all data are true and credible.

After analyzing the data, it is found that there are more outliers and missing values in the data, so the data is preprocessed, the negative rows are deleted first, the number of returns is small and it is not convenient to carry out data statistics, so it can be deleted, and then the outliers of the data are cleaned up, and in MATLAB, the data that do not meet the quartile in the data are cleaned up by the method of linear interpolation, so as to make the statistics of the data more accurate.

2.2 Method introduction

Based on the survey data, the time series forecasting is more accurate and effective in finding out the average sales volume and average selling price of different types of vegetables from the time variables [5]. The particle swarm optimization algorithm has few parameters, high operation efficiency, fast convergence speed, and is easy to find the global optimal solution.

ISSN 2616-5775 Vol. 6, Issue 13: 138-144, DOI: 10.25236/AJCIS.2023.061320

3. Model building and solving

3.1 Analyze the purchase strategy by category

After analyzing that the sales of vegetables are limited by seasons, in order to formulate the total replenishment volume and pricing strategy from July 1 to 7, 2023, the time series is used to forecast the sales volume.

3.1.1 Time series models

Step1: Sequence analysis

(1) Construct the sequence

Therefore, the corresponding time series are defined according to the question data:

$$x(t) \tag{1}$$

(2) Check whether the sequence is a white noise sequence

Null hypothesis:

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_m = 0, \quad \forall m \geq 1 \tag{2}$$

Alternative hypotheses:

$$H_1 : \exists \rho_k \neq 0, \ \forall m \geq 1, k \leq m \tag{3}$$

Q Statistics and LB statistics are used to test hypotheses, respectively.

First, construct Q the statistic:

$$Q = n\sum_{k=1}^{m} \hat{\rho}_k^2 \tag{4}$$

Where: is the n number of periods of observation of the series m, is the specified delay period, and is the $\hat{\rho}$ sample correlation coefficient. And if the null hypothesis is true, it is $\hat{\rho}_k$ independently identically distributed and approximately obeys the normal distribution. The $\hat{\rho}_k$ standard normal transformation is then performed

$$\sqrt{n}\hat{\rho}_k \sim N(0,1) \tag{5}$$

So there is

$$n\hat{\rho}_k^2 \sim \chi^2(1), \forall k \neq 0 \tag{6}$$

Therefore, we can obtain Q a chi-square distribution of m the statistics with approximately negative obedience degrees of freedom:

$$Q = n\sum_{k=1}^{m} \hat{\rho}_k^2 \sim \chi^2(m) \tag{7}$$

Judgment on white noise sequences: Q When the quantile of the chi-square distribution m with a statistic greater $1-\alpha$ than the degrees of freedom or the value of the statistic P is less than α that, the null hypothesis is rejected at the confidence level and the sequence is $1-\alpha$ considered to be a non-white noise sequence, otherwise, the null hypothesis is accepted that the sequence is a white noise sequence.

LB The process of constructing statistics is as follows:

In the case of small sample statistics, the Q test of statistics is not accurate, in this question to deal

ISSN 2616-5775 Vol. 6, Issue 13: 138-144, DOI: 10.25236/AJCIS.2023.061320

with short-term and long-term data, using statistics to calculate long-term data, short-term data to Q test with LB statistics.

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right) \tag{8}$$

Where: is the n number of periods of observation of the series m, is the specified delay period, and is the $\hat{\rho}$ sample correlation coefficient. The test method is the same as Q the statistic.

If the test results show that the sequence is a non-white noise sequence, it means that there is some factor affecting the sequence in the data information, that is, there may be seasonal effects, trend effects and random effects in the series, and the corresponding parameters need to be supplemented.

If the test results show that the sequence is a white noise sequence, then the subsequent prediction is irregular, and the random walk model can be used to make a rough prediction, or the expectation can be used instead.

Step 2: ARIMA model

The sequence is fitted by ARIMA, and the fitting model of the residual sequence is in the following form:

$$\begin{cases} \Phi(B)\nabla^d \varepsilon(t) = \Theta(B)\xi(t) \\ E(\xi(t)) = 0, Var(\xi(t)) = \sigma_\xi^2, E(\xi(s)\xi(t)) = 0, s \neq t \\ E(\varepsilon(s)\xi(t)) = 0, \forall s < t \end{cases} \tag{9}$$

Where: $\nabla^d = (1-B)^d$;is the d order of difference, is the autoregressive coefficient polynomial of the stationary reversible model, is the moving average coefficient polynomial of $\Phi(B) = 1-\phi_1 B - \cdots - \phi_p B^p$ the stationary reversible $ARMA(p,q)$ model $\Theta(B) = 1-\theta_1 B - \cdots - \theta_q B^q$, and is the $ARMA(p,q)$ $\{\xi(t)\}$ zero-mean white noise sequence.

According to the fitting model of the residual series, the modified moving average model is as follows:

$$MA_N(t) = \varepsilon(t) + \alpha_N \times x(t) + \alpha_N(1-\alpha_N) \times x(t-1) + \alpha_N(1-\alpha_N)^2 \times x(t-2) + \cdots \tag{10}$$

The recursive formula is:

$$MA_N(t) = \begin{cases} x(1) + \varepsilon(1) & t = 1 \\ \varepsilon(t) + \alpha_N \times x(t) + (1-\alpha_N) \times MA(t-1) & t > 1 \end{cases} \tag{11}$$

$$\varepsilon(t) = \frac{\Theta(B)}{\nabla^d \Phi(B)}\xi(t) \tag{12}$$

Table 1: Forecast sales by category

category	Cauliflower	Mosaic and leafy	Chili peppers	Nightshades	Edible fungus	Aquatic rhizomes
Forecasted sales volume (kg)	44.41	153.39	56.90	25.65	32.10	11.37

Using the above model to forecast the time series using weekly data for the 16 weeks from March 2021 to June 2021, the forecasted sales volume for each category from July 1 to July 7 is shown in Table 1.

Taking cauliflower vegetables as an example, the time series sales forecast is shown in Figure 1.

ISSN 2616-5775 Vol. 6, Issue 13: 138-144, DOI: 10.25236/AJCIS.2023.061320

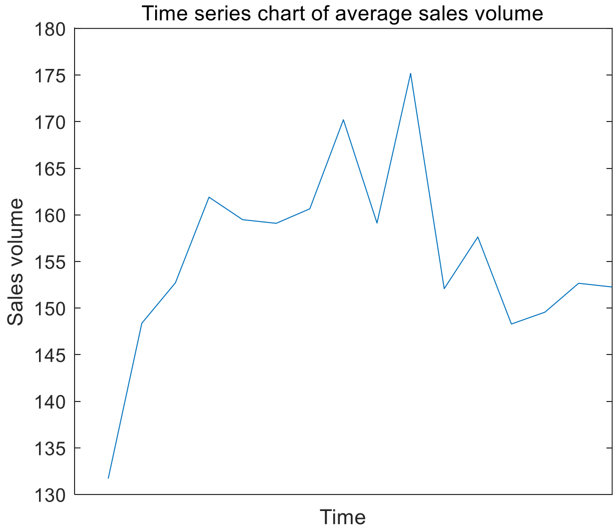

Figure 1: Time series forecast plot

3.1.2 Multivariate nonlinear programming models

Firstly, taking the markup rate of the total replenishment amount and the cost as the decision variables, the objective function of subtracting the total cost from the total selling price of the six categories is established to maximize the benefit, as follows:

$$\max \sum_{i=1}^{6} x_i y_i a_i - \frac{x_i y_i}{(1-z_i)}, \quad i = 1, 2, \cdots, 6 \tag{13}$$

Among them x_i, the average daily sales volume i of the first category of vegetables, the average cost of the first category of vegetables, the markup rate of the cost, and the y_i average i daily loss rate of the a_i first z_i i category of vegetables.

Then, because the selling price cannot be too high or too low, the markup rate cannot be higher than 1.5 and not lower than 1.05, so that the constraints are as follows:

$$1.05 \le a_i \le 1.5 \tag{14}$$

Secondly, the replenishment volume cannot be too much or too little, so it is stipulated that the replenishment volume cannot be higher than 150% of the forecast value and cannot be lower than 50% of the forecast value, so the constraints are as follows:

$$50\% \cdot s_i \le \frac{x_i y_i}{(1-z_i)} \le 150\% \cdot s_i, \quad i = 1, 2, \cdots, 6, \tag{15}$$

This s_i represents the i forecast value of the daily sales volume of the first category of vegetables.

Based on the above analysis, we establish the following multivariate nonlinear programming model:

$$\max \sum_{i=1}^{6} x_i y_i a_i - \frac{x_i y_i}{(1-z_i)}, \quad i = 1, 2, \cdots, 6 \tag{16}$$

$$s.t. \begin{cases} 1.05 \le a_i \le 1.5, \\ 50\% \cdot s_i \le \dfrac{x_i y_i}{(1-z_i)} \le 150\% \cdot s_i \end{cases} (i = 1, 2, \cdots, 6) \tag{17}$$

Using the value of the time series prediction as the substitution constraint, the optimal income of s_i

ISSN 2616-5775 Vol. 6, Issue 13: 138-144, DOI: 10.25236/AJCIS.2023.061320

17,108 yuan can be obtained by solving the nonlinear function, and the sales price, total replenishment and income of each category of vegetables from July 1 to 7, 2023 are obtained as shown in Table 2. According to the multivariate nonlinear programming, the purchase price and selling price of aquatic root vegetables are higher, and the purchase price and selling price of mosaic vegetables are lower.

Table 2: Selling price, replenishment volume and income table of each category of vegetables

category	Sales price (yuan).	Replenishment amount (kg).	Income (yuan).	Cost (yuan).
Cauliflower	10.95	100.36	1010.45	7.46
Mosaic and leafy	6.04	448.81	2538.29	3.01
Chili peppers	7.50	285.11	1538.74	3.37
Nightshades	7.62	59.56	403.35	4.17
Edible fungus	6.78	148.9	526.52	2.54
Aquatic rhizomes	13.02	43.64	536.46	9.57

As can be seen from the table, the price of mosaic vegetables is lower, but the sales volume is large, the market demand is large, so the replenishment volume is also higher, and it is expected to achieve higher income in this week. Aquatic root vegetables are priced higher, have the lowest restocks, and have smaller gains during the week.

3.2 Purchasing strategy based on single products

In order to more accurately predict the purchase of vegetables, the purchase quantity of each single product is predicted every day, and due to the large variety of single products, the number of single products is limited to 24-30 when solving, and the purchase quantity of each single product is not less than 2 kg, and the particle swarm algorithm is used to analyze the daily purchase volume and the purchased single product based on the corresponding conditions.

3.2.1 Particle swarm model

Basic steps of particle swarm modeling:

(1) Initialize the particle population (the population size is n), and initialize the random position and velocity of the particles

(2) Evaluate the fitness of each particle.

(3) Compare the current adaptation value of each particle with the adaptation value corresponding to the global optimal position, if the current adaptation value is higher, the global optimal position is updated with the current particle's position.

(4) Update the velocity and position of each particle according to the formula.

Particle i Position:

$$X_i = (X_{i1}, \ X_{i2}, \ \cdots, \ X_{iD}) \tag{18}$$

Particle i velocity:

$$V_i = (V_{i1}, \ V_{i2}, \ \cdots, \ V_{iD}) \tag{19}$$

The best position that an individual of Particle I has ever experienced:

$$pbest_i = (P_{i1}, P_{i2}, \ \cdots, \ P_{iD}) \tag{20}$$

The best position the population has ever experienced:

$$gbest = (g_1, g_2, \ \cdots, \ g_D) \tag{21}$$

The d-dimension update formula i for particles:

$$V_{id}^k = WV_{id}^{k-1} + c_1 r_1 (pbest_{id} - X_{id}^{k-1}) + c_2 r_2 (gbest_d - X_{id}^{k-1}) \tag{22}$$

(5) Judgment. When the end condition is reached, the algorithm ends. If the end condition is not met, repeat step 2. The algorithm stops when the algorithm reaches the maximum number of iterations or the

ISSN 2616-5775 Vol. 6, Issue 13: 138-144, DOI: 10.25236/AJCIS.2023.061320

increment of the optimal adaptation value is less than a given constant.

3.2.2 Model solution and result analysis

The particle swarm algorithm can obtain the single product replenishment quantity and pricing strategy on July 1, as shown in Table 3.

Table 3: Single product replenishment volume and pricing strategy

Vegetables	Name	Incoming quantity (kg).	Cost (yuan).	Sales price (yuan).
Cauliflower	broccoli	14.77	110.41	10.28
Cauliflower	Edae bok flower	3.67	27.24	10.55
Mosaic and leafy	amaranth	6.60	10.11	2.34
Mosaic and leafy	Takekana	5.98	10.38	5.23
Mosaic and leafy	Shanghai green	3.82	14.48	7.17
Mosaic and leafy	Fungus vegetables	7.39	13.89	9.40
Mosaic and leafy	Baby cabbage	8.66	38.49	6.41
Mosaic and leafy Mosaic and leafy	Milk cabbage Sweet potato tip	2.65 2.72	6.75 5.35	5.17 4.09
Mosaic and leafy	Green Vegetables	3.02	8.07	4.99
Mosaic and leafy	Yunnan lettuce (serving)	19.23	66.25	5.57
Mosaic and leafy	Oilseed lettuce (serving)	6.62	18.92	4.74
Mosaic and leafy	Spinach (serving)	5.22	19.71	6.31
Chili peppers	Wuhu Green Pepper	10.85	41.97	4.67
Chili peppers	Millet pepper (serving)	10.47	22.10	5.81
Chili peppers	Screw pepper (serving)	5.37	24.86	6.14
Chili peppers	Ginger, garlic, millet pepper	9.53	22.12	3.46
Chili peppers Chili peppers	Hangzhou pepper (serving) Red peppers	3.26 2.62	11.68 28.95	5.22 16.91
Nightshades Nightshades	Purple eggplant Long purple eggplant	3.65 2.52	11.94 11.50	6.63 12.24
Edible fungus	Cordyceps flowers	4.15	10.65	3.17
Edible fungus	Bisporus mushroom (box)	7.65	26.34	4.86
Edible fungus	Enoki mushroom (box)	14.88	21.55	1.41
Edible fungus	Seafood mushrooms (pack)	2.71	5.50	2.42
Aquatic rhizomes	Pure lotus root	4.93	50.86	13.31
Aquatic rhizomes	Rhombic	2.70	19.70	7.70

It can be seen from the table that in the replenishment single product on July 1, 27 types of single products were selected as purchase single products by solving, which can meet the restrictions on sales space in the title, with the largest variety of flower and leaf single products and moderate prices, while cauliflower vegetables, aquatic root vegetables and nightshade vegetables were replenished in less quantity, among which the net lotus root, long purple eggplant, red pepper, and Zhijiang green stem scattered flowers were sold at a higher price.

In the iterative process, the particle swarm algorithm gradually maximizes the revenue after 200 iterations, in order to seek the replenishment volume at this time, and meet the requirements of 24-30 sellable items, and the order quantity of each single item meets the minimum display quantity of 2.5 kg. From the calculations, it can be seen that the maximum value tends to be 2100 and the optimal value of individual particles is concentrated between 1800 and 2100, and the global optimal value that occurs at 200 iterations is shown in Figure 2.

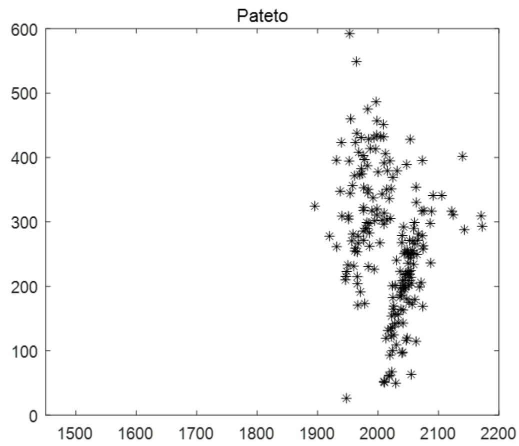

Figure 2: Optimal solution for particle iterations in 200 generations

4. Discussion

The sales volume data of various categories of vegetables under different conditions in the model is insufficient, so more research can be carried out to obtain the sales volume of various vegetables in supermarkets during holidays and different seasons and different weathers, so as to make predictions and seek more reasonable pricing and replenishment decisions. At the same time, the parameters in the model can be appropriately changed, such as increasing or decreasing the purchase volume of a certain category of vegetables, observing the change of the maximum revenue of the supermarket, and then adjusting the model to avoid the situation of undersupply or oversupply, reduce the loss, and maximize the income.

5. Conclusion

After 200 iterations, the maximum benefit of the particle swarm algorithm is around 2100, the error is small, and the model convergence speed is fast, and it is easy to obtain the global optimal solution. In the model solving, the maximum revenue of the supermarket and the corresponding replenishment strategy are solved with high accuracy. The model can be widely extended to prediction problems and optimization problems under multi-constraint conditions, and provides some suggestions for the purchase volume and pricing of vegetables, while meeting the needs of consumers and the maximum revenue of supermarkets.

References

[1] Mao Lisa. Research on pricing strategy and production and marketing model of vegetable wholesale market from the perspective of supply chain [D].Central South University of Forestry and Technology, 2023.
[2] Sun Xiaolu. Insurance pricing of vegetable price index based on policyholder's expectation [J]. Journal of Insurance Vocational College, 2018, 32(04):22-28.
[3] Zheng Lu. Research on dynamic bulk inventory replenishment of community group buying retailers [D]. Zhongnan University of Economics and Law,2022.
[4] Liang Shuting. Research on the inventory strategy of community fresh retail terminals where time and inventory affect the deterioration rate together [D].Beijing Jiaotong University, 2021.
[5] Liu He, Li Yanchun, Du Qinglong, et al. High water cut yield prediction method based on multivariate time series model[J].Journal of China University of Petroleum(Natural Science Edition),2023, 47(05): 103-114.