Supermarket Replenishment and Pricing Strategy Development Based on Time Series Forecasting and Nonlinear Programming

Jingcong Zhang[1,#], Zehui Li[1,#], Jieyu Wu[1], Jiahao Liu[1], Wanhe Wu[1], Shiqi Li*

[1]College of Science, Xi'an University of Architecture, Xi'an, China

[#]These Authors Contributed Equally to This Work

*Corresponding author: lsq200455@163.com

Keywords: Pricing Strategies, Polynomial Fitting Method, Time Series Model, Differential Evolutionary Algorithm

Abstract: This study aims to provide a comprehensive analysis of the sales characteristics and pricing strategies of vegetable commodities. Considering the high demand for the freshness of vegetable commodities as well as the interrelationships and price fluctuations among different categories, superstores need to take into account a variety of factors, including historical sales volume, seasonal variations, and availability, in order to formulate comprehensive pricing and replenishment strategies. This study examines the relationship between sales volume and cost-plus pricing of each category of vegetables in the off-season and peak seasons, using the categories of vegetables as the basis for classification. A polynomial fitting method was used to obtain the functional relationship between the two and to verify the fitting effect, which showed that the mean value of accuracy exceeded 95.6%. In addition, the sales volume is forecasted by a time series model and the smoothness of the series is verified. Under the premise of maximizing the revenue of the superstore, this study uses the pricing of each category as a decision variable, establishes an optimization model, and obtains the specific pricing strategy of each category by a differential evolutionary algorithm. The finalized pricing strategies were: 6.87 yuan/kg for flowers and leaves, 9.83 yuan/kg for edible mushrooms, 10 yuan/kg for aquatic root meridians, 13.64 yuan/kg for chili peppers, 7 yuan/kg for eggplant, and 10.99 yuan/kg for cauliflower. These results provide an important decision-making basis for the superstores to make more effective pricing and replenishment strategies for vegetables to meet market demand and maximize revenue. Through in-depth analysis of the relationship between sales data and pricing, superstores can better respond to changes in different seasons and market conditions, thereby improving operational efficiency and providing consumers with better quality goods and services.

1. Introduction

As an important agricultural product indispensable to the daily life of urban and rural residents, the efficiency and high quality of vegetable supply are crucial to social stability and the improvement of people's living standards. Ensuring the quality of vegetable supply is an issue that involves a wide range of people's livelihoods, so supermarkets need to take a series of measures to ensure that the supply of vegetables can meet people's demand for fresh and diversified vegetables. This not only requires supermarkets to formulate a scientific replenishment plan based on historical sales data and customer demand, but also needs to change the sales mix and other ways to maximize the supply of vegetables to meet the needs of the residents, and at the same time to achieve the maximization of the profits of the supermarkets themselves.

First of all, supermarkets can make more accurate replenishment plans by deeply analyzing historical sales data and customer demand. By analyzing the sales of different vegetable categories in detail, supermarkets can predict the sales peaks of different seasons and holidays, and then reasonably arrange the replenishment time and quantity. This kind of refined planning not only reduces the risk of inventory backlog but also improves the turnover rate of vegetables and reduces the loss caused by long-time storage of vegetables. In addition, supermarkets can also work closely with suppliers, and timely understanding of the origin of vegetables, growth cycle and other information, in order to

Copyright © (2025) Francis Academic Press, UK **DOI: 10.25236/icmmct.2025.022**

better arrange replenishment plans to ensure that the supply of vegetables can be synchronized with customer demand.

Secondly, supermarkets can change the sales mix to ensure that the supply of vegetables is maximized. Supermarkets can be based on seasonal and customer demand changes, flexible adjustments to the location of the display of vegetables and sales strategy, so that the hot vegetables are more prominent, but also to promote the sales of other categories. For example, for different seasons, supermarkets can increase the display of cool fruits such as watermelon and cucumber in summer, while increasing the supply of hot vegetables such as cabbage and radish in winter. In addition, supermarkets can also introduce pre-packaged vegetables and processed vegetable products to increase the added value of the products and meet the needs of customers who live a fast-paced life, thus further enhancing the sales volume of vegetables and the profits of supermarkets.

In conclusion, in the process of ensuring the supply of vegetables, hypermarkets need to take into account market demand, supply chain and sales strategy and other aspects, in order to ensure that the supply of vegetables is efficient, and high quality, to meet the growing demand of residents, but also to maximize their own profits. This not only helps to improve the quality of life of urban and rural residents but also promotes the sustainable development of supermarkets. Therefore, superstores need to continuously optimize operation management and improve supply chain efficiency in order to adapt to changes in the market and the diversification of customer needs and provide society with better quality vegetable products and services.

2. Literature review

At present, many scholars have carried out research on related problems.

For the demand forecasting problem, most of them belong to time series forecasting. There are many mature algorithms available. Yang et al [1] proposed a weighted slow feature analysis-based adaptive feature extraction (WSFA-AFE) method for multivariate time series forecasting. First, the weighted SFA (WSFA) algorithm is developed to extract slow features weighted by contribution; then, an improved adaptive sliding window algorithm is designed for self-judging the historical information of the slow features for inputs; finally, the WSFA-AFE method is applied to different ANN models, and combined with several benchmark datasets and practice datasets of the wastewater treatment process to validate the WSFA-AFE method's out-of-model performance. The results show that the WSFA-AFE method can adaptively extract feature variables from multivariate time series, which leads to better predictive modeling performance of multivariate time series for ANNs. In addition, the robustness of the proposed method is also demonstrated. Chen et al [2] proposed a hybrid ARIMA-LR model based on an autoregressive integrated moving average model (ARIMA) and linear regression model (LR) using an improved Bayesian combination model. Through the actual prediction of civil aviation cargo volume, it is found that the ARIMA-LR hybrid model not only better adapts to the changes caused by unexpected events, but also has a higher overall prediction accuracy than the ARIMA model and the LR model. The three indicators of the ARIMA-LR hybrid model, namely, mean absolute error (MAE), mean squared error (MSE), and mean absolute percentage error (MAPE), are respectively 1.0% lower than those of the ARIMA model by 1.06, 29.02, and 0.03; and compared with the LR model by 3.00, 92.00, and 0.06, respectively. Fu et al [3] proposed a self-attentive architecture with an information-interaction module known as a mix former for multivariate chaotic time series forecasting tasks. First, based on a rethinking of the multivariate data structure, a set of cross-convolution operators capable of automatically updating the parameters are utilized to compensate for the lack of phase space reconstruction, and an explicitly designed feature reconstruction module is proposed. Then, for the problem of interaction and fusion between series information and channel information, an information interaction module is proposed to realize feature communication by expanding and contracting dimensions. By breaking the communication barrier between series information and channel information, the feature representation capability is enhanced.Finally, we construct the mixformer that combines local sparse features and global contextual features.

In their study of nonlinear programming, Molina-Pérez et al [4] proposed a new proposal based on two basic strategies to improve the performance of differential evolutionary algorithms for solving MINLP problems. The first strategy considers a set of "well-adapted-unfeasible solutions" that help to explore promising regions from infeasible contours. It reduces the vulnerability of solutions to be attracted to larger discontinuous feasible components with unpromising objective function values. The second is composite trial vector generation to improve combinatorial exploration while ensuring strong convergence to the final solution. Sixteen well-known MINLP problems are used in several experiments to evaluate the performance of the proposed algorithm and compare it with state-of-the-art EA. The results provided by the proposal show better performance in terms of quality, robustness and computational cost. Huang et al [5] called the fuzzy interval credibility constraint (FIC) model combining fuzzy interval set (FIS) and interval nonlinear programming (INP) as the FIC-INP model and applied it to two cases. After building the FIC-INP model, the upper and lower bounds on the cost of disinfectant enhancers can be obtained for different combinations of upper and lower confidence limits for the two cases. The results show that the upper and lower booster costs increase as the lower confidence limit increases. The upper and lower booster costs increase with the number of boosters. For lower bound constraint confidence levels in the range of 0.6 ~ 0.9, the upper and lower booster costs increase as the trapezoidal distribution increases/widens, and the booster cost interval increases as the interval uncertainty increases. When the lower bound constraint confidence level is 0.5, the upper and lower booster costs decrease under a shrinking/narrowing trapezoidal distribution, and the booster cost interval increases under a trapezoidal distribution with larger intervals. Interval uncertainty has a higher impact than fuzzy uncertainty. The obtained results can provide more information for managers to develop a booster solution under fuzzy and interval uncertainty. Kırdar et al [6] addressed both objectives through a unified approach that integrates two decision support methods: regression and optimization. In the first stage, factors affecting the time to reach the curing temperature are identified and they are related using a multiple linear regression model. In the second stage, the regression model from the first stage is used and two objectives are considered to determine the efficient placement of parts in the autoclave: minimizing the duration of the heating phase and the maximum time delay between parts to reach the curing temperature. The former corresponds to increasing productivity and the latter to minimizing quality loss. Then Kırdar et al. proposed a bi-objective mixed integer nonlinear programming model and its equivalent linear model to generate the effective boundaries. In addition, a multi-objective evolutionary algorithm and its mechanism for solving the problem are proposed in order to obtain a solution faster. The method is applied to a real case of a composite plant. The estimated values of the regression model are very close to the realized values and considerable gains are observed for both objectives using the optimization tool.

This shows that the relevant research methods are more mature. However, there is a lack of detailed research on vegetable replenishment and pricing in supermarkets. Therefore, this paper focuses on the problem of vegetable replenishment and pricing strategy development in supermarkets.

3. Exploring the relationship between volume and pricing

Cost-plus pricing is based on product cost, product price = product cost + product cost Cost margin, i.e., product price = product cost (1 + margin).

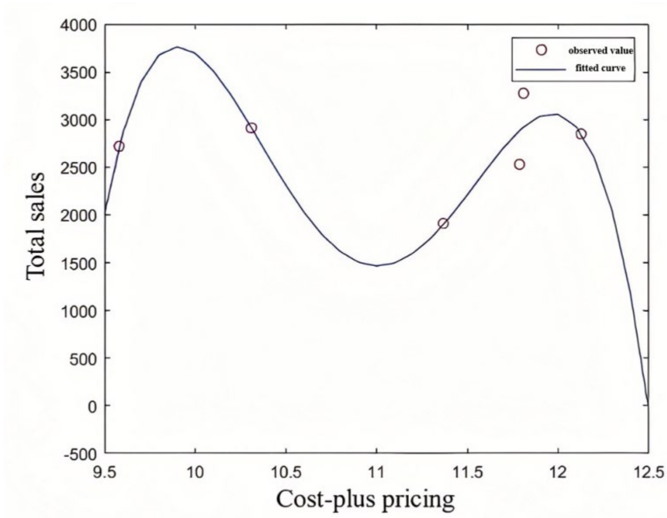

Fig.1 Relationship between sales volume and cost-plus pricing of edible mushrooms during peak selling season and the fitted equation

In this paper, the product price used is the unit price of sales that have been given: in terms of category, based on the law of sales, respectively, six categories of vegetables' off-peak and peak season sales and pricing relationship graph, to get the fitting equation, respectively, sales off-peak and peak season sales of the total amount of sales and cost-plus pricing to be fitted. The fitted graphs for edible mushrooms and eggplant categories with the fitted equations are shown in Figure 1.

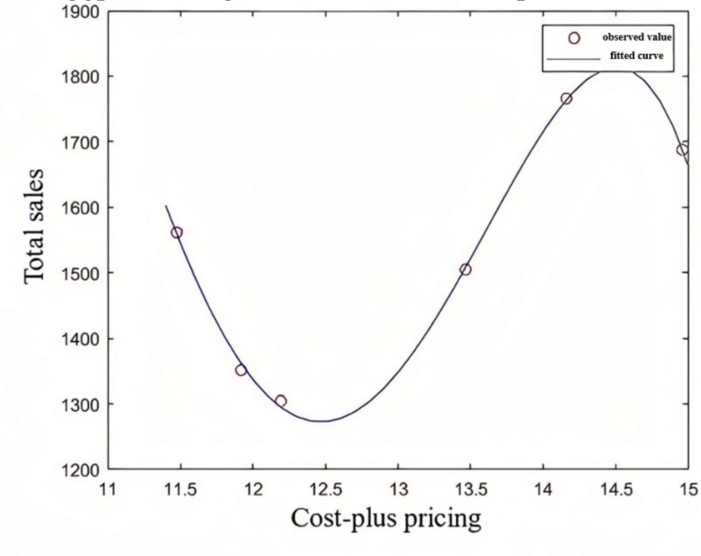

Fig.2 Relationship between sales volume and cost-plus pricing of edible mushrooms during the off-season and the fitted equation

Where the pricing of each category in each month is the average of the pricing of the individual items sold in that month Corresponding to foliage, edible mushrooms, aquatic roots and tubers, peppers, eggplants, and cauliflower, respectively.

The equations for sales volume and cost-plus pricing of edible mushrooms during peak and off-season selling seasons were obtained as shown in Figures 1 and 2:

$$s_2 = -1696.68{p_2}^4 + 74363.35{p_2}^3 - 1218594.21{p_2}^2 + 8848271.48p_2 - 24016149.49 \quad (1)$$

$$s_2' = -24.78{p_2}'^4 + 1209.17{p_2}'^3 - 21812.07{p_2}'^2 + 172235.02p_2' - 499837.50 \quad (2)$$

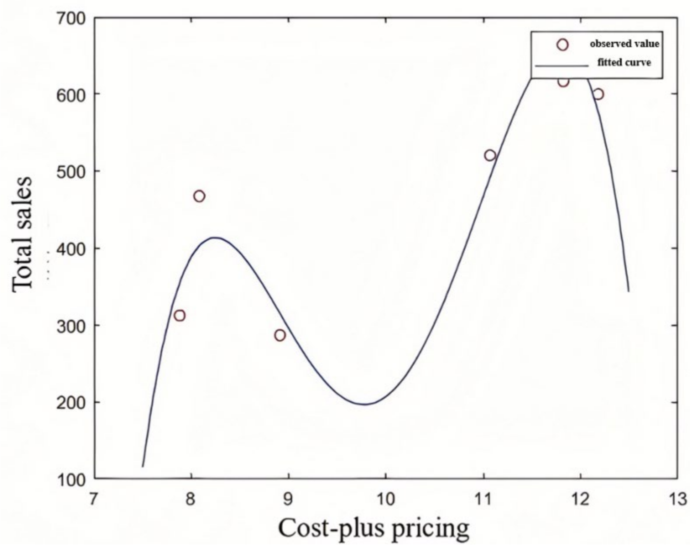

Fig.3 Relationship between sales volume of eggplant and cost-plus pricing during peak selling season and the fitted equation

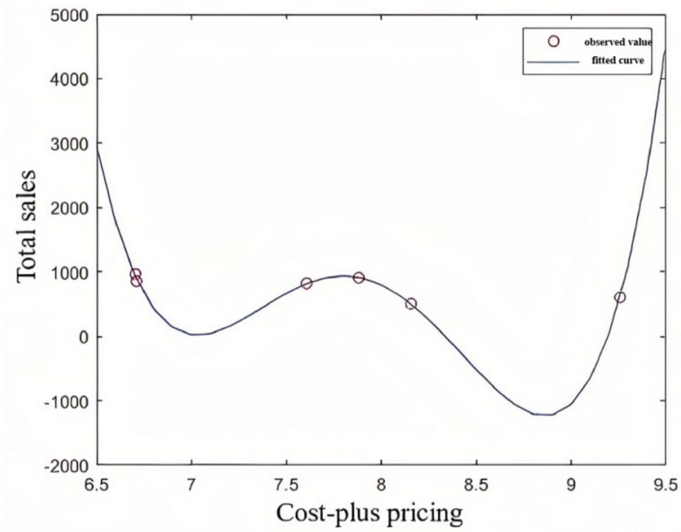

Fig.4 Relationship between sales volume and cost-plus pricing of eggplant in the off-season of sales and the fitted equations

The equations for sales volume and cost-plus pricing of eggplant in peak and off-season sales seasons are obtained in Figures 3 and 4:

$$s_5 = -32.60{p_5}^4 + 1296.40{p_5}^3 - 19124.81{p_5}^2 - 124068.37p_5 - 298398.07 \quad (3)$$

$$s_5' = 2101.41{p_5'}^4 - 66389.10{p_5'}^3 + 782980.37{p_5'}^2 - 4086054.33p_5' + 7962355.95 \quad (4)$$

In order to verify the reasonableness of the regression equation, this paper analyzes the error of the regression equation:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|s_i - \hat{s}_i|}{s_i} \quad (5)$$

The results of the error analysis are shown in Figure 5.

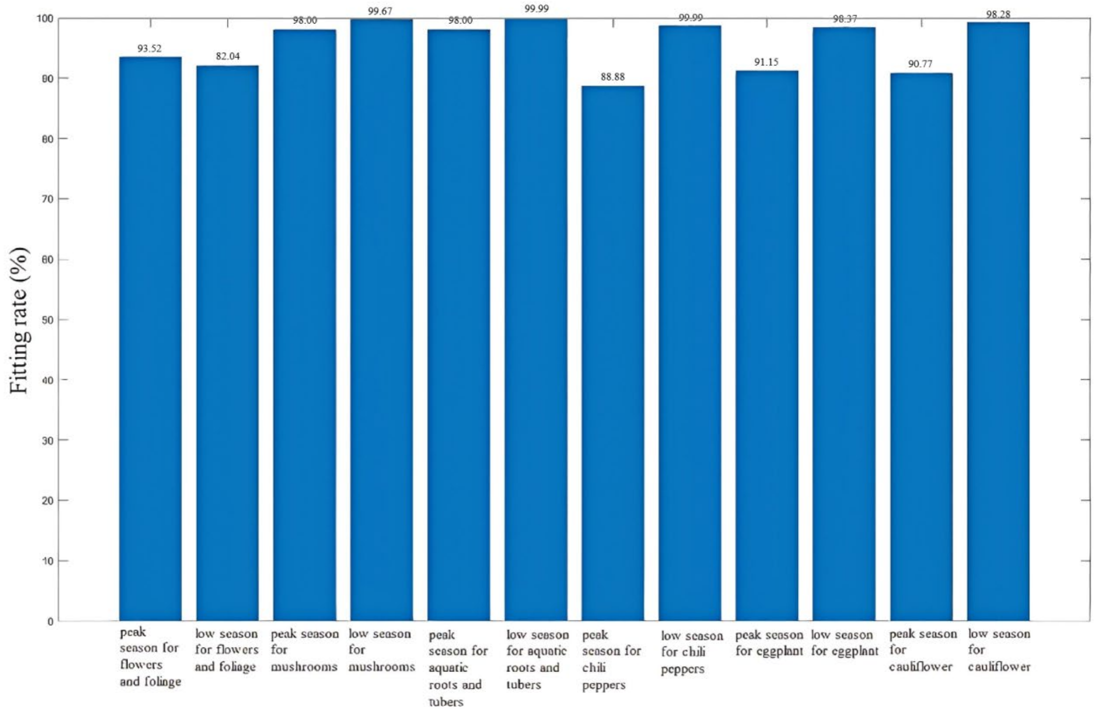

Fig.5 Fitting accuracy of each category at different stages

As can be seen in Figure 5, the average value of the prediction accuracy of each category at different stages is more than 95.6%, indicating that the fitting effect is more accurate according to the classification of sales off-peak and peak seasons given in this paper.

4. Replenishment determination based on time series forecasting

Before using time series for forecasting, the time series Z needs to be tested for smoothness:

Original hypothesis H_0: the sequence Z is smooth

Alternative hypothesis H_1: There is an upward or downward trend in the sequence Z

Set the time series as 2020.7.1 to 2020.7.7, 2021.7.1 to 2021.7.7, and 2022.7.1 to 2022.7.7, import the sales volume of each category for each day, set the confidence level as 95%, and analyze the smoothing analysis by using SPSS, and obtain the serial autocorrelation plot (ACF) and serial partial autocorrelation plot (PACF) respectively. As shown in Figure 6.

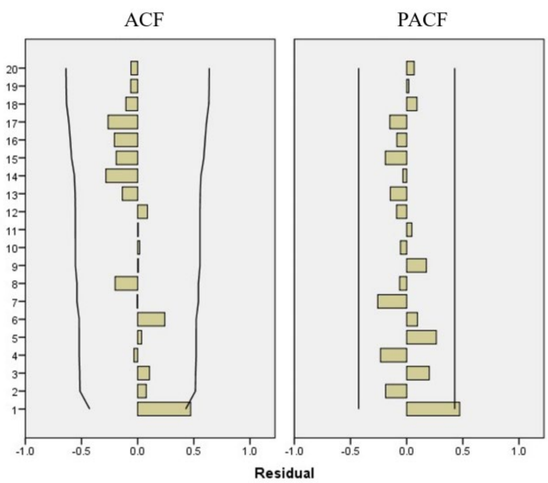

Fig.6 Sequential residual ACF and residual PACF

From Figure 6, the data are all within the confidence interval. Therefore, the original hypothesis H_0 cannot be rejected, which supports to some extent that the model is trendless, i.e., the sequence Z is smooth.

Based on the verified smoothness of the series, this paper uses AR modeling to predict the daily replenishment of each category for a one-week period from 2023.7.1 to 2023.7.7:

$$z_t = \theta_0 + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots\cdots + \theta_p z_{t-p} + \varepsilon_t \tag{6}$$

Based on the above equation, the total daily replenishment for each of the six categories of vegetables per day was obtained as shown in Table 1:

Table 1 Total daily replenishment for each category of vegetables from 2023.7.1-2023.7.7

	July 1	July 2	July 3	July 4	July 5	July 6	July 7
philodendron	118.78	137.36	144.51	133.34	130.20	136.82	138.06
edible mushroom	37.91	33.13	33.16	34.90	33.71	33.87	34.20
Aquatic rhizomes	40.69	40.98	41.26	41.10	41.12	41.13	41.12
capsicum	41.09	45.31	44.17	42.70	43.37	43.84	43.50
eggplant	16.56	14.52	14.17	14.87	14.89	14.66	14.69
cauliflower	42.60	43.90	44.49	43.99	44.16	44.15	44.13

5. Pricing strategy development based on nonlinear programming models

In order to maximize the revenue of the superstore, nonlinear planning is formulated based on the obtained daily replenishment of each category of vegetables:

1) Define the decision variable: this paper sets the decision variable as p_i, the pricing of the ith category

2) Establish the objective function: in order to make the superstore profit maximization, that is, the required revenue is maximized:

$$f = max \sum_{i=1}^{6} p_i \cdot s_i \cdot (1 - l_i) \tag{7}$$

Where l_i represents the attrition rate of the ith category:

$$l_i = \frac{1}{n} \sum_{j=1}^{n} l_{i_j} \tag{8}$$

$(1 - l_i)$ represents the real rate of return, to maximize the product of the return and the real rate of return, i.e., to maximize the real return.

3) Construct the constraints: the sum of the week's sales is not more than the total amount of replenishment for the week:

$$\sum s_i \leq \sum_{t=1}^{7} q_{i_t} \tag{9}$$

Where q_{i_t} is the replenishment quantity of the ith category on day t;

The quartiles are denoted by Q_1, Q_2, Q_3 and each range includes 25% of the data, in this paper, we take the upper quartile and the lower quartile, as shown in Figure 7:

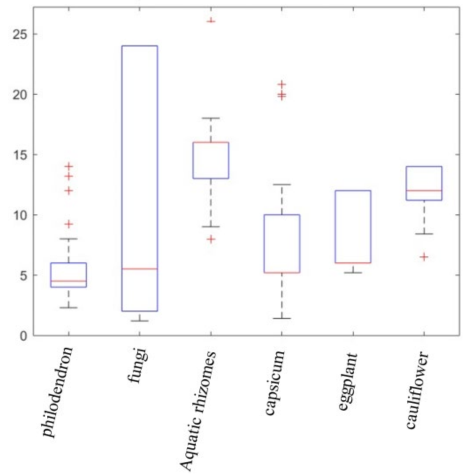

Fig.7 Box plots of the distribution of unit prices of vegetables sold by category

Constraints on the pricing of each category of vegetables:

$$Q_{i_1} \leq p_i \leq Q_{i_3} \tag{10}$$

$$s.t.\begin{cases} \sum s_i \leq \sum_{t=1}^{7} q_{i_t} \\ Q_{i_1} \leq p_i \leq Q_{i_3} \end{cases} \tag{11}$$

Construct a nonlinear plan based on the objective function and constraints:

$$f = max \sum_{i=1}^{6} p_i \cdot s_i \cdot (1 - l_i) \tag{12}$$

$$s.t.\begin{cases} \sum s_i \leq \sum_{t=1}^{7} q_{i_t} \\ Q_{i_1} \leq p_i \leq Q_{i_3} \end{cases} \tag{13}$$

For nonlinear planning models, intelligent optimization algorithms are a common way to solve them. In this paper, the model is solved using a differential evolutionary algorithm.

Population initialization

The population size M is chosen as 100, and M individuals are randomly and uniformly generated in the solution space.

$$X_i(0) = \left(\mathrm{x}_{i,1}(0), x_{i,2}(0), x_{i,3}(0), \ldots, x_{i,n}(0)\right), i = 1,2,3, \ldots, M \tag{14}$$

Among them.

$$x_{i,j}(0) = L_{j\_\min} + \mathrm{rand}(0,1)\left(L_{j\_\max} - L_{j\_\min}\right), i = 1,2,3, \ldots, M, j = 1,2,3, \ldots, n \tag{15}$$

Mutation.

In the g-th iteration, three individuals $X_{p1}(g), X_{p2}(g), X_{p2}(g)$ are randomly selected from the population with $p1 \neq p2 \neq p3 \neq i$, generating a vector of variation:

$$H_i(g) = X_{p1}(g) + F \cdot \left(X_{p2}(g) - X_{p3}(g)\right) \tag{16}$$

Where $\Delta_{p2,p3}(g) = X_{p2}(g) - X_{p3}(g)$ is the difference vector and F is the scaling factor.

The three randomly selected individuals in the variance operator are ranked from best to worst to obtain Xb, Xm, Xw, corresponding to the fitness fb, fm, fw, the variance operator reads:

$$V_i = X_b + F_i(X_m - X_w) \tag{17}$$

Also, the value of F varies adaptively according to the two individuals generating the difference vector:

$$F_i = F_l + (F_u - F_l)\frac{f_m - f_b}{f_w - f_b}, F_l = 0.1, F_u = 0.9 \tag{18}$$

The mutation strategy is:

$$DE/\ rand\ /1: V_i(g) = X_{p1}(g) + F\left(X_{p2}(g) - X_{p3}(g)\right) \tag{19}$$

$$DE/\ best\ /1: V_i(g) = X_{\text{best}}(g) + F\left(X_{p1}(g) - X_{p2}(g)\right) \tag{20}$$

$$DE/\ current\ to\ best/1 : V_i(g) = X_i(g) + F\left(X_{\text{best}}(g) - X_i(g)\right) + F\left(X_{p1}(g) - X_{p2}(g)\right) \tag{21}$$

$$DE/\ best\ /2: V_i(g) = X_{\text{best}}(g) + F\left(X_{p1}(g) - X_{p2}(g)\right) + F\left(X_{p3}(g) - X_{p4}(g)\right) \tag{22}$$

$$DE/\ rand\ /2: V_i(g) = X_{p1}(g) + F\left(X_{p2}(g) - X_{p3}(g)\right) + F\left(X_{p4}(g) - X_{p5}(g)\right) \tag{23}$$

Crossover:

$$v_{i,j} = \begin{cases} h_{i,j}(g), \mathrm{rand}(0,1) \leq cr \\ x_{i,j}(g),\ \text{else} \end{cases} \tag{24}$$

Where cr∈[0,1] is the crossover probability, taken as $cr = 0.7$.

Select:

$$X_i(g+1) = \begin{cases} V_i(g), f(V_i(g)) < f(X_i(g)) \\ X_i(g), \text{ else} \end{cases} \tag{25}$$

The model parameter settings are shown in Table 2.

Table 2 Differential evolutionary algorithm parameter settings

Parameter	value
Number of populations	100
Crossing rate	0.7
Variation rate	0.85
Allowable error of convergence	10^{-10}
Convergence tolerance judgment number	1000
Maximum allowable number of iterations	30000

The pricing of vegetables can be obtained from a differential evolutionary algorithm as shown in Table 3.

Table 3 Pricing of various types of vegetables

class	Unit price (yuan/kg)
Foliage	6.87
Edible Mushrooms	9.83
Aquatic Roots	10.00
Peppers	13.64
Eggplant	7.00
Cauliflower	10.99

6. Conclusions

In view of the high demand for freshness of vegetable commodities and the characteristics of interrelationships and price fluctuations that exist between different categories, superstores need to take into account a variety of factors, such as historical sales volume, seasonal variations, and availability, in order to formulate a comprehensive pricing and replenishment strategy that analyzes and forecasts the sales of various categories and individual vegetable items.

In this paper, the categories of vegetables are used as the basis for segmentation, and the relationship between sales volume and cost-plus pricing for each category of vegetables in the off-season and peak seasons is investigated. A polynomial fitting method is used to obtain the functional relationship between the two and to verify the fitting effect, which shows that the mean value of accuracy is more than 95.6%. In addition, in order to forecast the sales volume from July 1 to 7, 2023, a time series model was developed and the series was tested for smoothness using SPSS software. It was found that the autocorrelation and partial autocorrelation plots of the series are within the confidence interval, indicating that the time series is relatively smooth.

Under the premise of ensuring the revenue maximization of the superstore, this paper takes the pricing of each category as a decision variable and establishes an optimization model with the objective of maximizing the revenue of the superstore. The specific pricing strategy of each category is obtained by solving the differential evolution algorithm. The finalized pricing strategies were: 6.87 yuan/kg for flower and leaf, 9.83 yuan/kg for edible mushrooms, 10 yuan/kg for aquatic root meridians, 13.64 yuan/kg for chili peppers, 7 yuan/kg for eggplant, and 10.99 yuan/kg for cauliflower.

These results provide an important decision-making basis for the superstore, enabling it to effectively develop pricing and replenishment strategies for vegetables to meet market demand and maximize revenue. Meanwhile, through in-depth analysis of the relationship between sales data and pricing, superstores can better respond to changes in different seasons and market conditions, thereby improving operational efficiency and providing better quality goods and services to consumers.

References

[1] S. Yang, W. Li, and J. Qiao, “A WSFA-based adaptive feature extraction method for multivariate time series prediction,” Neural Computing and Applications, vol. 36, no. 4, pp. 1959–1972, 2024.

[2] B. Chen, J. Liu, Z. Ruan, M. Yue, H. Long, and W. Yao, “Freight traffic of civil aviation volume forecast based on hybrid ARIMA-LR model,” in International Conference on Smart Transportation and City Engineering (STCE 2022), M. Mikusova, Ed., Chongqing, China: SPIE, Dec. 2022, p. 69.

[3] K. Fu, H. Li, and Y. Bai, “Mixformer: An improved self-attention architecture applied to multivariate chaotic time series prediction,” Expert Systems with Applications, vol. 241, 2024.

[4] D. Molina-Perez, E. Mezura-Montes, E. A. Portilla-Flores, E. Vega-Alvarado, and B. Calva-Yanez, “A differential evolution algorithm for solving mixed-integer nonlinear programming problems,” Swarm and Evolutionary Computation, vol. 84, 2024.

[5] X. Huang and Y. Wang, “A fuzzy-interval nonlinear programming for boosters under uncertainty,” Water Supply, vol. 22, no. 11, pp. 8139–8156, 2022.

[6] G. Kirdar, D. Tezcaner Ozturk, and M. C. Testik, “Regression-based multiobjective optimization approaches for the autoclave composite parts placement problem: a mixed integer nonlinear programming model and an evolutionary algorithm,” Soft Computing, vol. 27, no. 19, pp. 14045–14062, 2023.