Research on Active Pricing and Replenishment Decision Making of Vegetable Commodities Based on Particle Swarm Algorithm

Xutao Huang[1, *, #], Yingying Zheng[2, #], Fan Liang[3, #]

[1]School of Management, Shenzhen University, Shenzhen, China, 518060

[2]School of Mathematics Science, Shenzhen University, Shenzhen, China, 518060

[3]School of Economics, Shenzhen University, Shenzhen, China, 518060

*Corresponding author: huangxutao2021@email.szu.edu.cn

#These authors contributed equally.

Abstract. With economic development, people's eating habits are becoming more green and healthy, which provides a wide market for the sale of vegetable products. Due to the short shelf life and easily damaged characteristics of vegetable products, most of them cannot be sold the next day. As a result, supermarkets need to replenish their stock on a daily basis based on historical sales data. To help supermarkets formulate pricing and replenishment strategies to maximise sales, this paper takes two approaches based on domestic and international studies. First, this paper draws a scatter plot of total sales and markup rate and fits the functional relationship and concludes that the correlation between the two is not strong; it adopts a time series forecasting model to determine the cost price of each vegetable category for the next seven days, and solves the total daily replenishment and pricing for the next seven days using particle swarm algorithm through the establishment of a profit objective function. Second, considering the limited selling space of supermarkets and satisfying the market demand as much as possible, this paper introduces the constraint that the sellable single product categories are within 20, and solves for the replenishment and pricing of single products on 1 July 2023 in supermarkets. This paper provides workable ideas for pricing and replenishment decisions for vegetable products in supermarkets.

Keywords: Time Series Prediction, Particle Swarm Algorithm, Goal Programming.

1. Introduction

With the improvement of people's living standards, people are more in pursuit of a green and healthy lifestyle, eating habits have changed greatly, the demand for vegetable commodities is growing, but also more attention to the quality and variety of vegetable commodities, vegetable commodities in the daily sales of supermarkets accounted for a large proportion of the super is an indispensable part of the super business. However, vegetable commodities are typical perishable products, its preservation cycle is short, the decay rate is fast, usually not sold on the day, the next day will not be able to sell; in the process of transporting vegetable commodities are also very easy to damage, resulting in a certain amount of waste; vegetable commodities are many varieties of commodities, the origin and harvest date are different, susceptible to the impact of seasonal factors, merchants are usually required to be in the specific types of vegetable demand and the price of the purchase of the same day are not sure! The unscientific replenishment and pricing strategy may cause the problem of mismatch between supply and demand, which will have a certain impact on the profitability of supermarkets. Therefore, this paper focuses on the analysis of market demand based on the historical sales data of supermarkets, and researches and formulates the dynamic replenishment and pricing strategy of supermarkets, which can help supermarkets achieve the maximum profit to the greatest extent.

By reviewing the existing literature, for the pricing and replenishment study of fresh vegetable commodities, Zhao Ling and Liu Zhixue [1] analyse the structural part of the optimal value function to represent the optimal pricing and inventory strategy, and at the same time consider the fixed cost and customer returns to establish a dynamic inventory and pricing model, and the numerical analysis using heuristic algorithm solution concludes that the optimal replenishment level and pricing will change monotonically with the customer returns and fixed cost The conclusion. Zhang Jinlong et al [2], based on the consideration of fresh fruits and vegetables and other perishable new products, proposed that the demand for such products is composed of attempted purchases and repeated purchases, and used the Bass model including the price effect to describe the dynamic evolution process of the number of attempted purchases, and then introduced the repeated purchase rate to construct the demand function of perishable new products, and then analysed the relationship between the demand function and the deterioration rate, and established a dynamic inventory and pricing model of perishable new products with dynamic bulk replenishment using a heuristic algorithm. Then, by analysing the relationship between the demand function and the deterioration rate, we establish a joint decision model of pricing and dynamic batch replenishment of new products, and derive the necessity of joint decision of pricing and replenishment, as well as the U-shaped correlation relationship between the optimal price and the product diffusion rate, and the repetitive buying behaviour that decreases and then increases. Song Hongfang et al [3], based on the dynamic pricing problem of goods under rapid replenishment with fixed ordering costs, used Markov decision making and Bellman's equation to construct a dynamic planning model for profit maximisation, and used heuristic algorithms to solve optimal pricing and optimal inventory levels.

This paper aims to help supermarkets develop optimal replenishment and pricing strategies for vegetable products as the main research direction, and gives the replenishment and pricing strategies for vegetable products from two directions to provide reference for the operation of supermarkets. The research directions of this paper are as follows:

Direction 1, based on the supermarket's historical sales data to study the relationship between the total sales volume of vegetable products and cost-plus pricing, and based on the resulting relationship to give the supermarket the total replenishment and pricing strategy for each category in the coming week. Vegetable pricing is generally based on the cost-plus pricing method, where customer demand for individual vegetable products is sensitive to price fluctuations. By using this pricing method, supermarkets can better regulate prices to keep them at a stable level, and discounts are also required for individual products that are damaged or in poor condition. The development of better replenishment and pricing strategy usually requires us to fully study the historical sales data of supermarkets of vegetables, this paper first by drawing the total sales of each vegetable category and cost-plus pricing scatterplot, the use of the principle of least squares based on the plot of the scatterplot to fit into a functional relationship, to analyse the relationship between the total sales of each vegetable category and cost-plus pricing. Vegetable commodities are usually affected by seasonal factors, showing a cyclical pattern, so this paper adopts the time series prediction model to determine the cost price of each category in the next seven days, through the relationship obtained by fitting the function to establish the profit objective function, based on the conditions of maximum revenue of the supermarket, the use of particle swarm algorithm to solve the replenishment volume of each vegetable category per day on 1-7 July 2023 and the corresponding pricing strategy. .

Direction two, taking into account the problem of limited sales space of vegetable category in the supermarket, this paper introduces the two constraints of the saleable single species category and the number of single products on the shelves to be controlled within 20, and then gives the pricing and replenishment strategy of the supermarket. In this paper, the particle swarm algorithm is used to solve the nonlinear programming problem under the above constraints, which gives the single items of the superstore to be purchased on July 1, as well as the replenishment quantity and pricing decision of each single item.

2. Preparatory phase of modelling

2.1. Description of symbols

The notation used in this paper to construct and solve the model is shown in Table 1.

Table1 Symbol Description

Notation	Instructions	Unit
W	Profitability of goods	Yuan
P	Pricing of commodities	Yuan
R	Wholesale price of goods	Yuan
Q	Daily replenishment of commodities	kg
S	Sales of goods	kg
I	Mark-up rates for commodities	
L	Wear and tear rate of commodities	

Note: Variables that are not declared are governed by their specific description where the symbol appears.

2.2. Model assumption

(1) It is assumed that the merchant does not know exactly what the specific individual items and purchase prices are on the day of replenishment;
(2) Market conditions are assumed to be relatively stable during the observation period;
(3) Single instantaneous replenishment, no stock-outs allowed.

2.3. Data preprocessing

(1) Data consolidation
In this paper, the data files are merged by applying the merge function of the pandas library in Python, and due to the relatively large amount of data, the merged data are classified and summarised according to a single day by using the groupby() function in Python, which facilitates the analysis and processing of the data.
(2) Cost-plus pricing
Cost-plus pricing = unit cost * (1 + mark-up rate)
Unit cost = (total fixed cost + total variable cost) / sales volume = wholesale price * (1 + discount rate)
Mark-up rate = (selling price - wholesale price) / wholesale price *100%
Then there is:

$$\text{Cost} - \text{plus pricing} = R(1+L)\left(1+\frac{P-R}{R}\right) \tag{1}$$

3. Direction 1 modelling and solving

3.1. Scatterplotting and fitting functional relationships

In this paper, the cost-plus pricing of each vegetable category is taken as the x-axis variable and the total daily sales volume of each vegetable category is taken as the y-axis variable, a scatter plot is drawn to first determine the relationship between the two, and then based on the principle of the least squares method, a regression is fitted to the scatter plot to obtain the functional relationship between the two. As a result, six scatter plots of the relationship between cost-plus pricing and daily sales volume were obtained for each vegetable category, of which the scatter plot for chilli is shown in Figure 1.

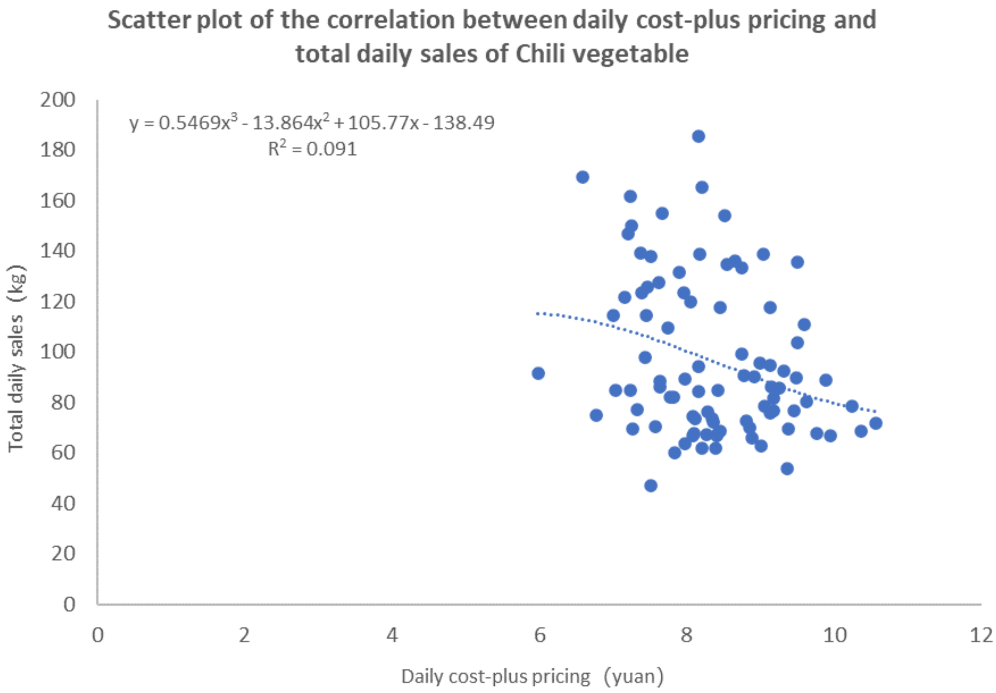

Figure1 Scatterplot of cost-plus pricing of chilli with total daily sales volume

Observing Figure 1, it can be seen that the result of the fitted function for the pepper class is a decreasing function in the defined range, which is in line with the general understanding of price and sales volume, and by analysing the sample points, it can be seen that the determined price corresponds to a large difference in the total amount of daily sales, for example, the total amount of daily sales is distributed in the range of 50-200kg when the cost plus price is 8 yuan, and most of the sales are distributed in the range of 50-100kg.

Looking at the six scatter plots, it can be seen that the goodness of fit R^2 of the regression function is small and poorly fitted, considering that the reason may be that there are too few indicator factors and it is difficult to accurately measure the sales volume of cost-plus pricing of a single indicator.

3.2. Forecast of cost price

(1) Weak smoothness condition for time series:

①$E(X_t) = E(X_{t-s}) = \mu$ (Mean value is a fixed constant)

②$Var(X_t) = Var(X_{t-s}) = \sigma^2$ (Variance exists and is constant)

③$Cov(X_t, X_{t-s}) = \gamma_s$ (Covariance is only related to the interval s, not t)

(2) ARIMA(p, d, p) model is as follows:

$$y'_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i\, y'_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \beta_i\, \varepsilon_{t-i} \tag{2}$$

$$\text{Both } y'_t = \Delta^d y_t = (1-L)^d y_t \tag{3}$$

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right)(1-L)^d y_t = \alpha_0 + \left(1 + \sum_{i=1}^{q} \beta_i L^i\right)\varepsilon_t \tag{4}$$

where p is the autoregressive order, d is the difference order, p is the moving average order, L is the lag operator, ε is the perturbation term, α is the autoregressive coefficient, and β is the moving average coefficient.

3.3. Q-Test

$H_0 : \rho_1 = \cdots = \rho_s = 0$

$H_1 : \rho_i (i = 1,2,\cdots,s)$ At least one of them is not 0 ,

Under the H_0 the condition that holds, construct the statistic:

$$Q = T\left(T + 2T + 2\sum_{k=1}^{s}\frac{r_k^2}{T-k} \sim \chi_{s-n}^2\right) \tag{5}$$

Where T is the sample size, n is the number of unknown parameters in the model, k is the number of sums, and s is taken as 18 in SPSS.

3.4. Forecasts of wholesale prices

In this paper, the wholesale price of each category in the data file from April to June 2023 is extracted as a time series, first we judge whether the time series is smooth or not, and after differential smoothing of the unstable time series, we use SPSS 25.0 expert modeller [4] to make the ACF plot, PACF plot and Q-test, and from this, we deduce the time series model and make a prediction of the wholesale price of the next 7 days, such as the prediction results of flower and leaf category in Figure 2, Figure 3.

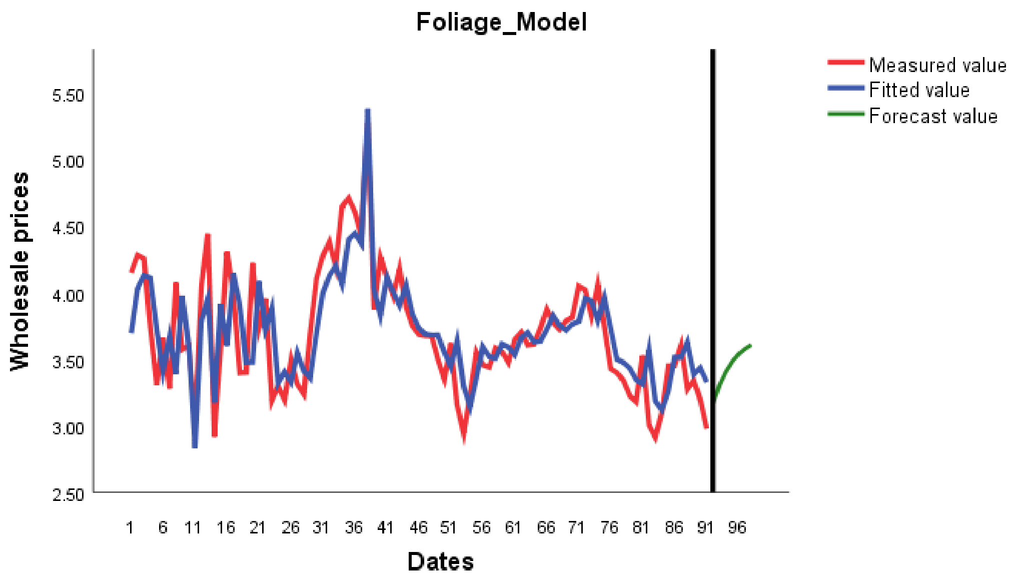

Figure 2 Time series forecast of wholesale price of foliage and flowers

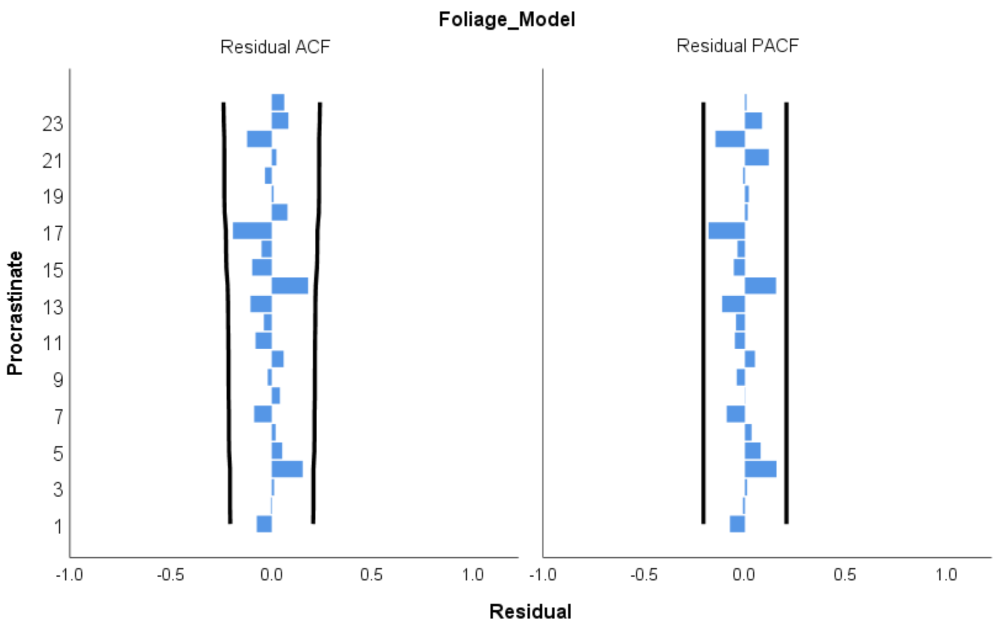

Figure 3 Predicted ACF and PACF maps for foliar species

Observing the ACF and PACF plots of the residuals, it can be seen that the autocorrelation coefficients and partial autocorrelation coefficients of all lagged orders are not significantly different from 0. In addition, this paper performs the Q-test on the residuals, and the probability p-value obtained is 0.500, which results in a result greater than 0.05, and the original hypothesis cannot be rejected under the condition of 99% of the confidence level, i.e, it can be assumed that the residuals are white noise and therefore the ARIMA (1, 0, 0) model is able to identify the time series well.

The results of the time series fitting models are shown in Table 2, the model type for all six vegetable categories is ARIMA [5], where the difference order is 1 for cauliflower and edible mushrooms, and 0 for all vegetable categories except edible mushrooms, which has a moving average order of 1. The autoregressive orders for aquatic roots and tubers, foliage, cauliflower, eggplant, pepper, and edible mushrooms are 3, 1, 0, 0, 1, 1, and 0, respectively.

Table2 Time Series Models by Category

Kind	Aquatic rhizomes	Philodendron	Cauliflower	Eggplant	Chilli	Edible fungi
Model type	ARIMA (3, 0, 0)	ARIMA (1, 0, 0)	ARIMA (0, 1, 0)	ARIMA (1, 0, 0)	ARIMA (1, 0, 0)	ARIMA (0, 1, 1)

The predicted cost price of each vegetable category from 1 to 7 July is shown in Table 3, and the cost price of each category was relatively stable over the seven-day period, with no significant fluctuations.

Table 3 Projected cost price by category

Cost (yuan/kg) / Dates	Aquatic rhizomes	Philodendron	Cauliflower	Eggplant	Chilli	Edible mushroom
1 July	12.37	3.17	7.90	4.95	4.76	6.82
2 July	12.11	3.31	7.90	4.96	4.80	6.79
3 July	12.05	3.41	7.90	4.97	4.83	6.77
4 July	12.05	3.48	7.90	4.98	4.83	6.74
5 July	11.92	3.54	7.90	4.98	4.84	6.71
6 July	11.88	3.58	7.90	4.99	4.85	6.69
7 July	11.83	3.61	7.90	5.00	4.86	6.66

3.5. Objective function

In order to find the total daily replenishment and pricing strategy for each vegetable category for the next 1-7 days so that the supermarket's profit is maximised, this paper establishes a planning model for each category with the aim of maximising profit.

$$max: W = SP - RS(1 + L)$$
$$St.\text{ Decision variables } P \in [a, b] \tag{6}$$

where a is the lowest price from April 2023 and b is the highest price from April 2023; In particular, S is a function of P. A fitted regression gives.

where the total daily replenishment quantity Q has the following expression

$$Q = S \times (1 + L) \tag{7}$$

3.6. Nonlinear programming solutions

In this paper, the particle swarm algorithm is used in the heuristic algorithm [6] to numerically solve the above objective function. The core idea of the particle swarm algorithm is to use the exchange of information between individuals in the group to make the whole group evolve from disorder to order in the solution space so as to obtain a feasible solution to the problem.

The particle is located at the dth iteration:

$$x(d) = x(d-1) - v(d-1) \cdot t \tag{8}$$

The velocity of the particle at the dth iteration:

$$v(d) = w \cdot \underbrace{v(d-1)}_{\text{Previous step own speed inertia}} + c_1 \cdot r_1 \cdot (\underbrace{pbest(d)}_{\text{Self-perception component}} - x(d)) + c_2 \cdot r_2 \cdot (\underbrace{gbest(d)}_{\text{Social-perception component}} - x(d)) \tag{9}$$

Among them c_1, and c_2 denote individual, social learning factors, respectively; r_1 and r_2 denotes the $[0, 1]$ the random numbers on the w as the inertia weights, and $pbest$, and $gbest$ denote the individual and group historical optimal positions, respectively. The particle swarm algorithm implementation process is shown in Figure 4.

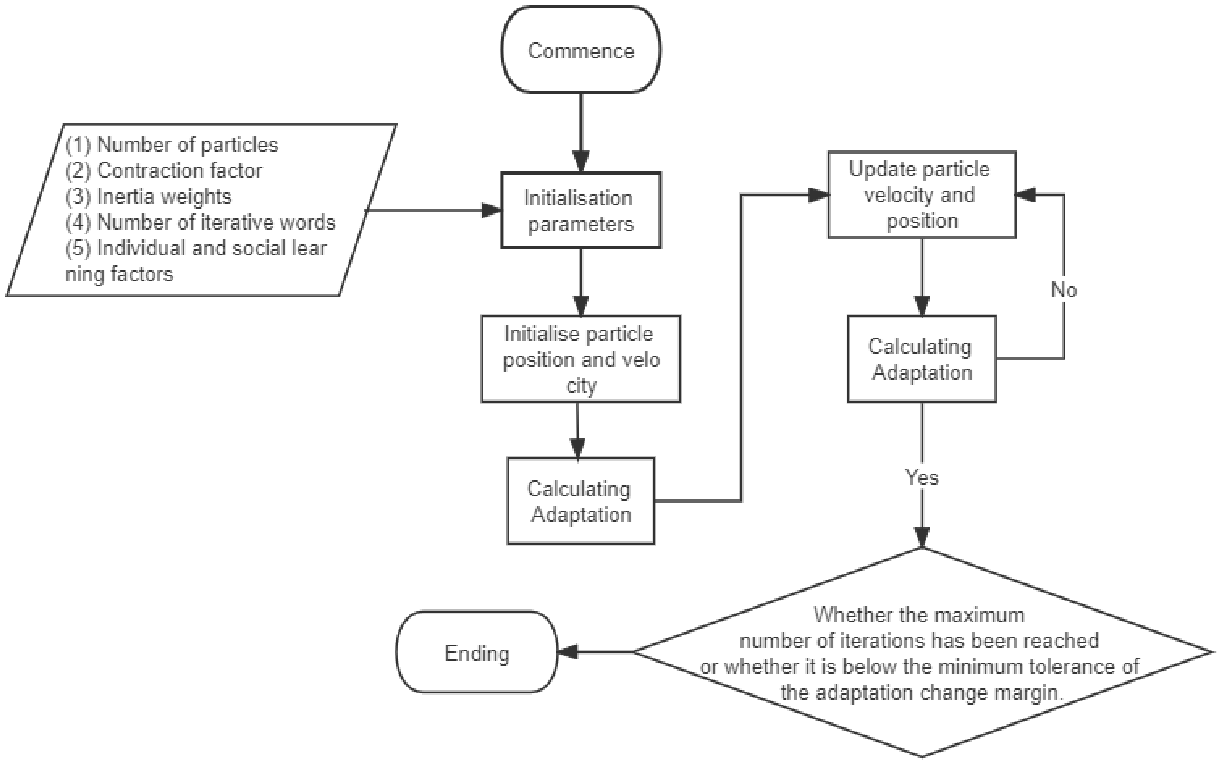

Figure 4 Process for implementing particle swarm algorithms

3.7. Principles of Particle Swarm Algorithm

In this paper, we use Matlab's own particle swarm function particleswarm [7][8] For the solution, the principle adopts the adaptive domain model, according to Qian Feng, "Particle swarm algorithm and its industrial applications", the domain model refers to the particle swarm in the search process only part of the particles around it as the domain particles, this model makes the particle swarm can be divided into several different sub-populations, which is conducive to search in multiple regions, avoiding the algorithm to fall into the local optimum.

Its self-adaptation is reflected in the fact that when the fitness starts to stagnate, the particle swarm search switches from the domain mode to the global mode, and then back to the domain mode when the fitness starts to decrease, in order to avoid falling into a local optimum. When the number of stagnations of the fitness is large enough, the inertia coefficient starts to decrease, thus facilitating the local search.

3.8. Preset parameter

(1) The number of particles is set by default to: min{100.10*nvars }, where nvars is the number of variables;
(2) The inertia weights are set by default in the range [0.1, 1.1] [9];
(3) The individual learning factor is set to 1.49 by default [10];
(4) The social learning factor is set to 1.49 by default;
(5) The scale of the particles in the field is set to 0.25 by default.

3.9. Results

The total daily replenishment and pricing for each vegetable category for 1-7 July 2023 was derived by setting up a profit objective planning function, which was solved using the particle swarm algorithm under profit maximisation. The results for the aquatic roots and tubers category are shown in Table 4.

Table 4 Total daily replenishment and pricing of aquatic roots and tubers 1-7 July 2023

Aquatic rhizomes			
Dates	Total daily replenishment (Q)	Pricing (P)	Profit (W)
1 July	18.77	13.56	52.89
2 July	18.69	13.82	56.51
3 July	19.01	12.82	56.71
4 July	18.67	13.87	57.22
5 July	18.63	13.99	59.14
6 July	18.61	14.02	59.65
7 July	18.60	14.07	60.34

4. Orientation 2 Model building and solving

4.1. Screening of saleable items

In this paper, we use the pivot_table function in Python's pandas library to sort and summarise the data files by date of sale, item name and sales volume, and then select the total daily sales volume of each item from 24-30 June 2023, and then use Excel to sort and filter out the saleable items from 24-30 days.

4.2. model solution

For each individual product, an objective function is defined and solved using the particle swarm algorithm according to the same principle as above.

$$max: W = SP - RS(1 + L)$$

$$St.\text{ Decision variables } P \in [m, n] \tag{10}$$

where m is the lowest price on 24-30 June 2023 and n is the highest price on 24-30 June 2023;
In particular, S is a function of P. This is obtained by fitting a regression to the average daily price and total daily sales data on 24-30 June 2023 for each individual product.

4.3. Results

The overall daily replenishment and pricing strategy for each vegetable item on 1 July 2023 is derived by setting a profit objective planning function for each item and solving it using a particle swarm algorithm under profit maximisation. The results are presented in Table 5.

Table 5 Total daily replenishment and pricing for each vegetable item

trade name	Shanghai Qing	Yunnan oilseed rape (portion)	Lotus root(1)	Yunnan lettuce (portion)	Peppers (portions)	Little bok choy(1)	......
Pricing P (yuan)	8.00	4.45	15.53	5.73	5.76	5.20	......
Replenishment Q (kg)	3.72	27.66	7.42	39.20	27.07	5.93	......
Profit W (yuan)	14.24	21.45	17.40	81.21	18.41	9.17	......

5. Conclusions

Vegetable products occupy an important position in people's diet structure, but the supply of vegetable products is affected by a variety of variables such as freshness period, transportation loss rate and seasonal factors, how to effectively use the past sales data of vegetable products to predict the replenishment and pricing of supermarkets on the same day, which is of great significance to the operation of supermarkets. This paper aims to help supermarkets formulate pricing and replenishment strategies scientifically and effectively, and based on the existing research at home and abroad, the research on the optimal replenishment and pricing strategies of vegetable products in supermarkets is carried out in two directions. In the first direction, since the pricing of vegetable products usually adopts the "cost-plus pricing" method, this paper first plots the scatter plot of daily total sales and cost-plus pricing of each vegetable category, and concludes that the correlation between daily total sales and cost-plus pricing is not strong; and since the data have the characteristics of time series, this paper uses SPSS25.0 to study the optimal replenishment and pricing strategy. SPSS25.0 Expert Modeller to construct a time series prediction model to determine the cost price of each vegetable category for the next seven days, and then constructed a profit objective function, which was solved numerically using the particle swarm algorithm in the heuristic algorithm to obtain the total daily replenishment volume and pricing strategy based on the objective of the maximum profit of the supermarket for the next 1-7 days. In the second direction, in real life, because the supermarket has a fixed footprint, considering the limited sales space of the supermarket and to meet the market demand of each vegetable category as much as possible, this paper filters out 19 saleable single products based on the ranking of daily sales volume, and constructs an objective planning function based on the replenishment volume and pricing of each single product from 24-30 June 2023, and also uses the particle swarm algorithm to solve numerically to obtain the replenishment volume and pricing strategy of the single product on 1 July 2023 in the supermarket. Through the above research, this paper provides realistic guidance for the replenishment and pricing decisions of vegetable products in superstores.

References

[1] ZHAO Ling,LIU Zhixue. Joint replenishment and pricing strategy considering customer returns and fixed costs[J]. Operations Research and Management,2022,31(06):105-110+124.

[2] ZHANG Jinlong,WU Xiang,XU Haoxuan. A joint decision model for pricing and replenishment of perishable new products[J]. Journal of Systems Engineering,2018,33(01):79-89.

[3] SONG Hongfang,RAN Lun,LI Jinlin et al. Dynamic pricing and inventory joint strategy under fast replenishment mode[J]. Practice and Understanding of Mathematics,2016,46(05):79-86.

[4] Qian Jun,Wang Yongbo. Application comparison of several statistical software to build ARIMA model[J]. Statistics and Decision Making,2018,34(16):80-84.

[5] ZHOU Ming,YAN Zheng,NI Yixin et al. A new ARIMA method for electricity price prediction with error prediction correction[J]. Chinese Journal of Electrical Engineering,2004(12):67-72.

[6] Kennedy J, Eberhart R. Particle swarm optimization[C]//Proceedings of ICNN'95- International Conference on Neural Networks. IEEE,1995

[7] Mezura-Montes, E., and C. A. Coello Coello. "Constraint-handling in nature-inspired numerical optimisation: past, present and future". Swarm and Evolutionary Computation. 2011, pp. 173-194.

[8] Pedersen, M. E. "Good Parameters for Particle Swarm Optimisation." Luxembourg: Hvass Laboratories, 2010

[9] Simplified particle swarm optimisation algorithm based on random inertia weights[J]. Computer Application Research,2014,031(002):361-363,391.

[10] Eberhart R C. Comparinginertia weights and constriction factors in otimisation[C]//Proceedings of the 2000 IEEE Congress on Evolutionary Computation, La Jolla, CA.IEEE, 2000.