Research on Profit Maximization Model for Automatic Pricing and Replenishment Decision of Vegetable and Plant Commodities

Hang Zhao[1], Yikang Li[2], Ying Yang[3], Lin Tang[4,*]

[1]School of Life Science, Yunnan Normal University, Kunming, China, 650500

[2]School of Mathematics, Yunnan Normal University, Kunming, China, 650500

[3]School of Chinese Language and Literture, Yunnan Normal University, Kunming, China, 650500

[4]Faculty of Education, Yunnan Normal University, Kunming, China, 650500

*Corresponding author: maitanweng2@163.com

Abstract. The strategy of automatic pricing and replenishment of fresh commodity sales statistics is a long-standing challenge. Accurate pricing is of great significance for commodity sales, loss reduction and economic benefit improvement. The automatic pricing and replenishment strategy for fresh product sales statistics is a long-standing challenge, and accurate pricing is of great significance for product sales, reducing losses, and improving economic benefits. In order to investigate the pricing relationship between the sales of vegetables within a certain time range and the cost of individual products or categories, as well as the replenishment plan and pricing strategy for the next week, based on the Adaptive Differential Evolution Algorithm theory, the analysis is conducted with individual products and categories as units to achieve the goal function of maximum revenue, and a dynamic programming model is constructed, Obtain the specific replenishment total amount and pricing strategy for each vegetable category in the next week and per day. The results have been tested to have good applicability.

Keywords: Revenue Maximisation Optimisation model; Adaptive Differential Evolution Algorithm; Multi-objective planning model; Cluster analysis.

1. Introduction

Since 2002,fresh food supermarkets have ushered in the golden age, and automatic commodity pricing and replenishment are an important part of the operation of fresh supermarkets. The Multi-product joint replenishment strategy can reduce the corresponding cost [1], and the diversified business strategy can maintain the steady growth of profits [2], but the law and trend of freshness decline of fresh products, the development of appropriate pricing strategy for profit is also an urgent solution [3], and a reasonable sales mix becomes particularly critical.

2. The basic funamental of Adaptive Differential Evolution Algorithm

2.1. The structure of Adaptive Differential Evolution Algorithm

The traditional Evolution Algorithm converges slowly in late evolution and tends to fall into local optimal solutions. This paper improves the evolutionary algorithm based on the ideas proposed by Li Yali et al. [1][4]. To adjust mutation operator and crossover operator, make is adjusted with the change of the number of iterations, and the improved Adaptive Differential Evolution Algorithm (ADEA) firstly on the chromosome coding of genetic algorithm, designed with n number of methods, An individual of length 1+2n is generated, as shown in Figure 1:

1	1	2	1	1	2	0.14	4.36	2.47	2.69	3.46	14.28	16.31

Figure 1. Chromosome (n=6)

Subsequently, the initialisation of the population was carried out with relevant settings, including the chromosome size, the length of the chromosomes and the differential evolution operator of the

ADEA algorithm, which were randomly assigned according to the settings of the parameters.After outputting the results, the optimal number of iterations is judged. If the optimal number of iterations is reached, the optimal result is directly output. Instead, the variation operator and crossover operator need to be adjusted to calculate the fitness value of each chromosome and compare the fitness values of the target vector and the experimental vector. Chromosomes with higher fitness values are then selected to enter the next generation population, and this operation is repeated until the maximum number of iterations is reached. When the maximum number of iterations is reached, the ADEA algorithm stops running and outputs the best fitness, the chromosome corresponding to the best fitness, and the best solution. The specific process is shown in Figure 2:

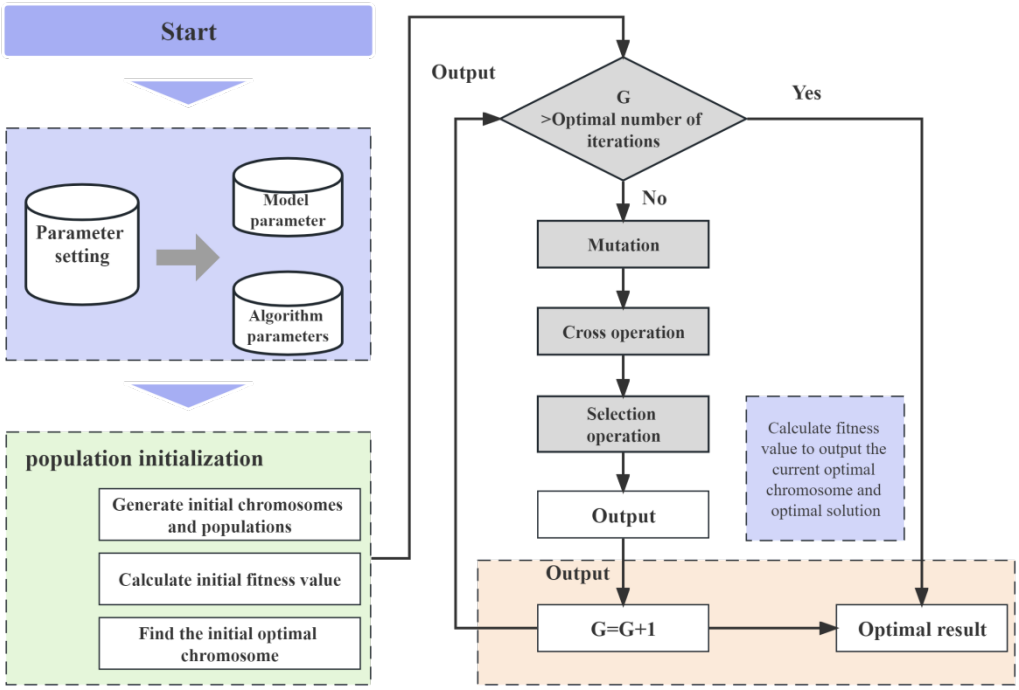

Figure 2. Adaptive Differential Evolution Algorithm (ADEA) model operation flowchart

2.2. The determination of the objective function

By quantifying the preservation situation of each category, analyzing the data such as discount, daily sales volume, unit price of sales, sales type, discount sales, average wholesale price, and transportation loss rate, the profit of the objective function is maximized based on the adaptive differential evolution algorithm model, and the absolute value of the difference between the supply data and the demand data is quantified, and then a multi-objective programming model is obtained. Set the total number of items (available for sale from June 24 to 30, 2023) of each category as N, where n1,n2,... ni is the corresponding sales price (unit price) recorded as Qi (i=1,2,3,…,i),Pi is the degree of discount, Mi is the market demand, Yi is the shipping loss rate (situation), Gi is the purchase price, etc. Existence Li satisfies:

$$L_i = \begin{cases} 1, & \text{Replenishment} \\ 0, & \text{No replenishment} \end{cases} \tag{1}$$

Meanwhile, set the order quantity to xi, and combine the sales price (unit price), daily sales volume, discount, type, etc. Based on ADEA algorithm model, the corresponding objective function is as follows:

$$\begin{aligned}\prod(p_1,T,k_i)=&\sum_{i=1}^{n}\frac{[1+sign(k_iT-t_{di})]}{2}\left\{p_iD_i(p)-\frac{c_iD_i(p)}{k_iT}\frac{c_iD_i(p)}{k_iT_i}\left[\frac{1}{2}\theta_i(k_iT-td_i)^2+k_iT\right]\right.\\&\left.-\sum_{i=1}^{n}\frac{h_iD_i(p)}{2k_iT_i}\left[\theta_i td_i(k_iT-td_i)^2+(k_iT)^2\right]-\frac{s_i}{k_iT}\right\}+\sum_{i=1}^{n}\frac{[1-sign(k_iT-t_{di}]}{2}\\&\left\{p_iD_i(p)-c_iD_i(p)-\frac{h_iD_i(p)T_i}{2}-\frac{s_i}{k_iT}\right\}-\frac{S}{T}\\&=L_ix_iQ_i(1-Y_i)-\sum L_ix_iG_i+\sum L_ix_iP_iY_i\end{aligned}\tag{2}$$

In addition, the correlation between supply and specific demand needs to be satisfied [5],which is: $C_i=|M_i—X_i|$, and at the same time satisfy that the number of individual items is greater than 27 and less than 33, it is: $27\le L_i\le 33$, and the minimum value of the order quantity of each individual item is 2.5kg, it is: $2.5\text{kg}\le X_i$. The specific relationship between daily replenishment quantity and pricing strategy of individual products, and the objective function optimized based on ADEA algorithm is obtained as follows:

$$\begin{aligned}\max\left(\prod(p_1,T,k_i)\right)=&\sum_{i=1}^{n}\frac{[1+sign(k_iT-t_{di})]}{2}\left\{p_iD_i(p)-\frac{c_iD_i(p)}{k_iT}\frac{c_iD_i(p)}{k_iT_i}\left[\frac{1}{2}\theta_i(k_iT-td_i)^2+k_iT\right]\right.\\&\left.-\sum_{i=1}^{n}\frac{h_iD_i(p)}{2k_iT_i}\left[\theta_i td_i(k_iT-td_i)^2+(k_iT)^2\right]-\frac{s_i}{k_iT}\right\}+\sum_{i=1}^{n}\frac{[1-sign(k_iT-t_{di}]}{2}\\&\left\{p_iD_i(p)-c_iD_i(p)-\frac{h_iD_i(p)T_i}{2}-\frac{s_i}{k_iT}\right\}-\frac{S}{T}\\&=L_ix_iQ_i(1-Y_i)-\sum L_ix_iG_i+\sum L_ix_iP_iY_i\end{aligned}\tag{3}$$

The optimization model for maximizing returns is as follows:

$$\max L_i=\sum_{n}^{i=1}\left|M_i-X_i\right|\tag{4}$$

Finally, based on the Revenue Maximization Optimization model, the Multi-objective Programming model is obtained as follows:

$$s.t.\begin{cases}L_i=\begin{matrix}0,\ \text{No replenishment}\\1,\quad\ \text{Replenishment}\end{matrix}\\27\le\sum_{i=1}^{n}L_i\le 33\qquad(i=1,2,3,\text{L}\ ,n)\\2.5\le X_i\end{cases}\tag{5}$$

3. Results

3.1. The establishment of Cluster analysis model

The cluster model of vegetable commodities was implemented in Python software.On the basis of the existing data, this paper cluster the data of vegetable commodities as shown in Figure 3.Based on the data of sales volume, unit price, type of sales, wholesale price and transportation loss rate of vegetable commodities from the third quarter of 2020 to the first half of 2023, the quality control of the data and the concrete display of the difference data were carried out, and the correlation between

different categories and different items of vegetable commodities was obtained by combining the clustering results. Based on the specific sales volume and quarterly sales, a correlation model was built to visualize the school competition rules of vegetable commodities, as shown in Figure 4:

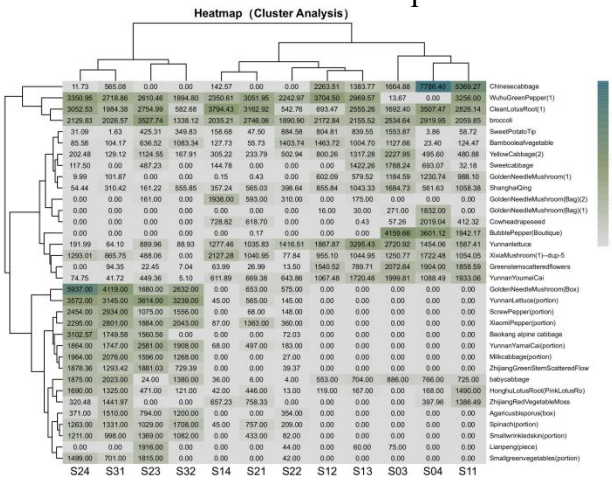

Figure 3. Heatmap for Cluster Analysis Based on Large Sample Data

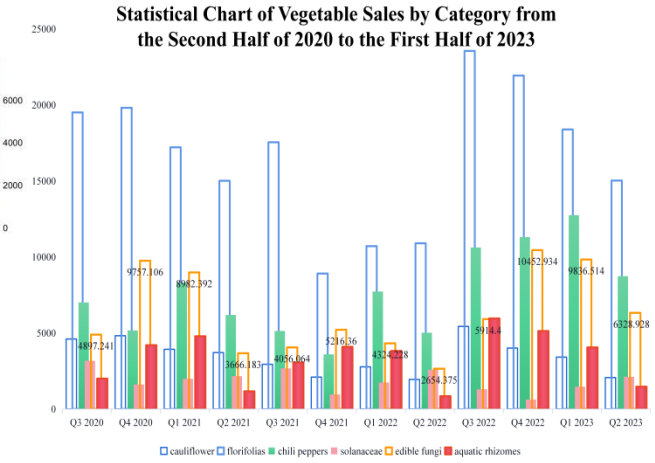

Figure 4. Sales Volume Statistics of Vegetables by Category, H2 2020 - H1 2023

3.2. Analysis of experimental results

According to the results of cluster analysis and correlation analysis, the data were screened, and the sales volume of fruit and vegetable products was 0.5kg as the dividing line, and the data frequency of each interval was divided and counted, and the mean value, variance, range and coefficient of variation of product sales were calculated for comprehensive evaluation.

3.2.1 Short term analysis of vegetable commodity sales volume

Short-term analysis was conducted based on the category sales volume of vegetable commodities (in daily units), and the results were shown in Table 1:

Table.1. Results of short-term analyses of category sales volume of vegetable commodities

category	*mean value*	*variance*	*range*	*coefficient of variation*	*25% quantile*	*75% quantile*
Cauliflower	38.49442488	515.0575977	186.155	0.589562787	23.305	48.521
Anthophyllum	182.9686433	7430.307272	1234.175	0.471114767	128.701	223.512
Capsicum	84.41348295	2855.409198	598.165	0.63302718	51.243	102.131
Solanacea	20.67445346	181.8251844	118.931	0.652218301	11.668	26.834
Edible mushroom	70.12601382	2351.269269	508.124	0.691467911	39.259	89.191
Aquatic rhizomes	37.40216866	983.2733472	295.866	0.838378919	15.053	50.818

Short-term analysis was conducted based on the sales volume of vegetable commodities (in daily units), and the results were shown in Table 2:

Table.2. Results of short-term analyses of unit sales of vegetable commodities

category	*mean value*	*variance*	*range*	*coefficient of variation*	*25% quantile*	*75% quantile*
Spinach	4.83005648	461.1158727	16.9938	0.026678028	1.311158333	6.786941667
Flowering cabbage	4.16362685	564.6375527	18.0719	0.034246325	0.553975	5.90495
Red pepper	3.20173611	412.4091494	16.852	0.038060939	0.123641667	3.895583333
Sweet potato tips	4.7772527	1210.117162	22.8784	0.043695466	0.009875	8.387891667
Yellow cabbage	7.3962870	1659.213511	36.067067	0.033047474	2.480475	9.471175
……	……	……	……	……	……	……
Milk cabbage	5.3854703	368.4356783	12.4414	0.021387405	2.601408333	7.631533333

3.2.2 Long term analysis of vegetable commodity sales volume

Long-term analysis was conducted based on the category sales volume of vegetable commodities (in monthly units), and the results are shown in Table 3:

Table.3. Results of long-term analyses of category sales volumes of vegetable commodities

category	*mean value*	*variance*	*range*	*coefficient of variation*	*25% quantile*	*75% quantile*
Cauliflower	10437.5474	15620030.39	14903.23	0.3787	7407.37	13575.5789
Anthophyllum	29974.1612	106812946.38	39848.15	0.3448	21510.00	36502.4834
Capsicum	20948.1394	120845214.7	58200.73	0.5248	14199.74	23694.9366
Solanacea	5309.0072	5305756.74	9078.0198	0.4339	4189.4207	6702.3334
Edible mushroom	17211.0498	55225912.14	28596.528	0.4318	11872.9761	19549.4450
Aquatic rhizomes	9724.7195	25356817.08	21157.20	0.5178	6430.1376	12210.3853

Long-term analysis was conducted based on the sales volume of vegetable products (in monthly units). The results are shown in Table 4:

Table.4. Results of long-term analyses of unit sales of vegetable commodities

category	*mean value*	*variance*	*range*	*coefficient of variation*	*25% quantile*	*75% quantile*
White mushroom	81.30555	4714.21825	284	0.832659407	30.75	105.25
Spinach	144.90169	13833.4761	509.814	0.800340826	39.33475	203.60825
Flowering cabbage	124.9088	16939.1265	542.157	1.027389737	16.61925	177.1485
Wuhu green pepper	782.3425	239727.219	1632.92	0.617084313	556.40625	1153.02125
Broccoli	764.923	72338.5523	1233.99	0.3466969	595.199	858.4495
……	……	……	……	……	……	……
Bamboo leaf	201.1323	47126.5940	708.283	1.064226102	9.46125	386.89425

3.2.3 Inspection Results

Through the data results in Table 1 to Table 4, it can be seen that the results of the long-term analysis are more stable in terms of volatility and lower dispersion compared to the results of the short-term analysis. Based on this this paper selects the data results of long term analysis and performs correlation analysis as shown in Figure 5-6:

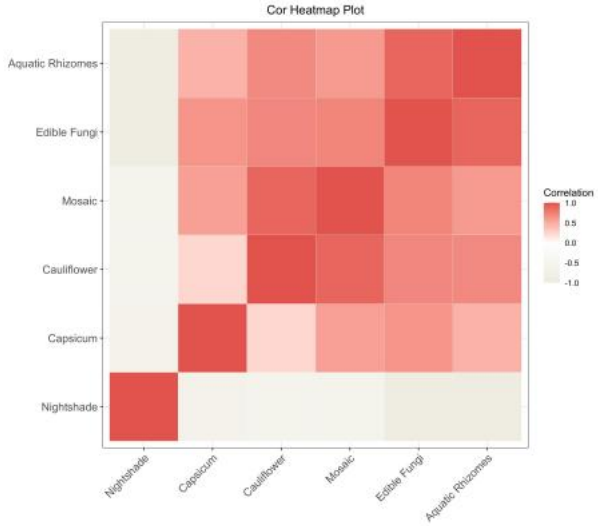

Figure 5. Heat map of correlation coefficient of each vegetable category

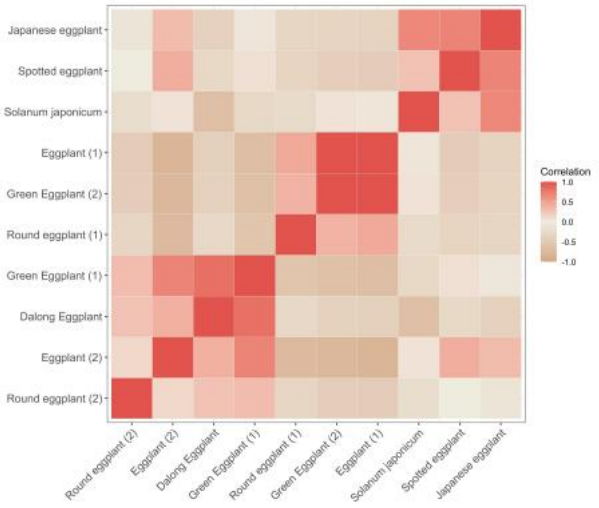

Figure 6. Heat map of correlation coefficient of each vegetable (nightshade)

It can be seen from the data results that the correlation of each group is relatively close, the correlation coefficient $r>0.8$, and the confidence $P<0.001$. The correlation matrix determines the degree of linearity by the color depth. As shown in the figure above, the darker the color, the higher the degree of linearity between the indices. Because there are obvious dark blocks, indicating that the correlation coefficient between each index is large, it can be considered that there is a large correlation between each category and single product. At the same time, this paper analyzes the relationship between the sales volume of six categories and obtains the results in Figure 7:

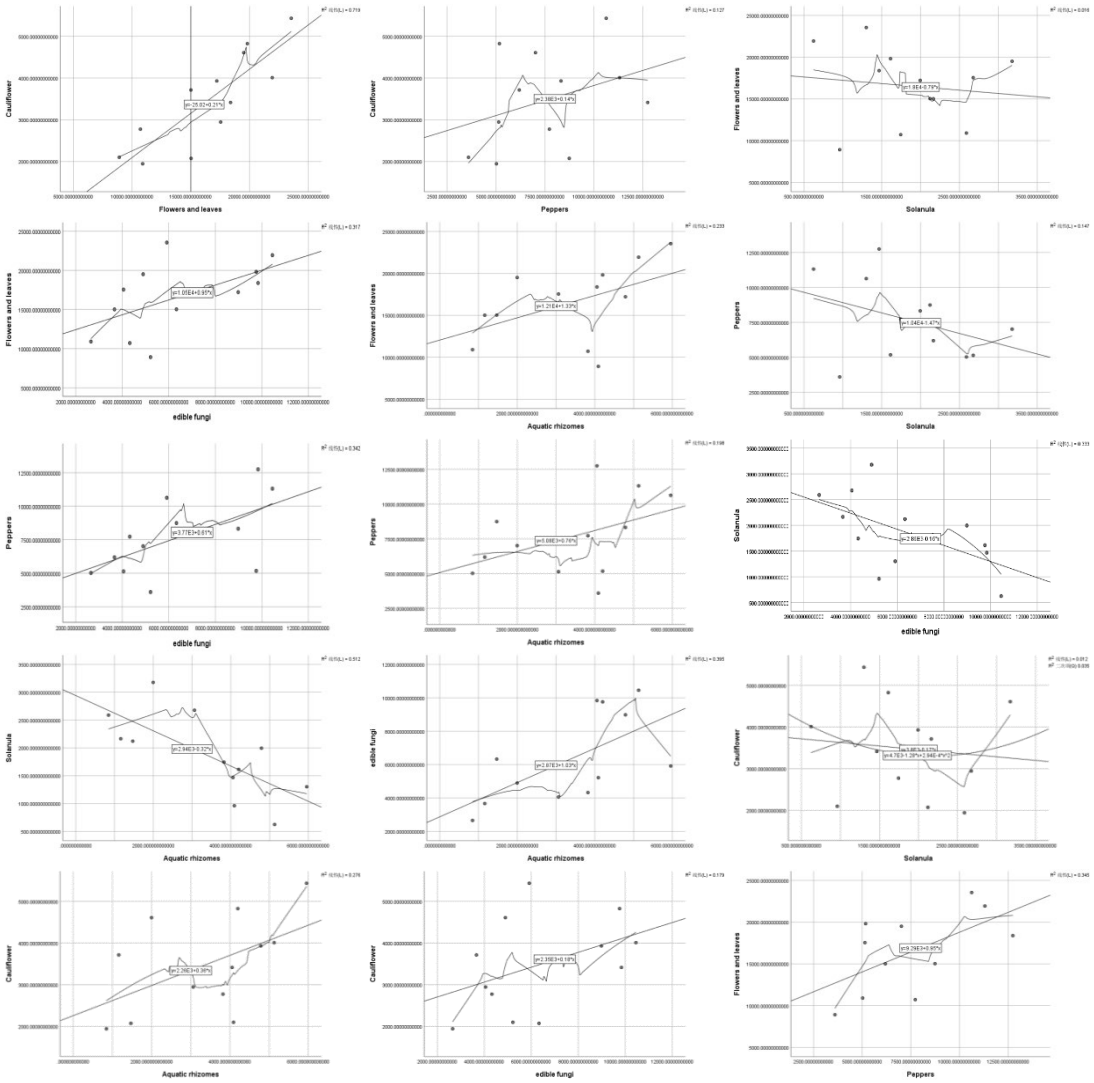

Figure 7. The relationship between the sales volume of vegetable commodities

This indicates that the sales distribution of each product in the same category is unbalanced, and different categories of products show different sales distribution rules in different seasons. Flowers and leaves and cauliflower are the most prosperous in spring, while eggplant, aquatic roots and peppers show high sales vitality in winter and spring, but edible fungi have a low sales volume in summer. In general, flowers, peppers and edible fungi have a high demand and stable sales performance in the market. Affected by the fluctuation of consumer demand and market price, the correlation between each item of the same category is weak, while the consumer groups of Mosaic and cauliflower, Mosaic and chili, Mosaic and edible fungus, and chili and edible fungus have a certain overlap, and there is a strong correlation between the corresponding items of different categories.

3.2.4 Sales Analysis of Vegetable Categories

Then, the histogram of the significant difference changes of vegetable commodity categories obtained after principal component analysis was compared and analyzed, as shown in Figure 8:

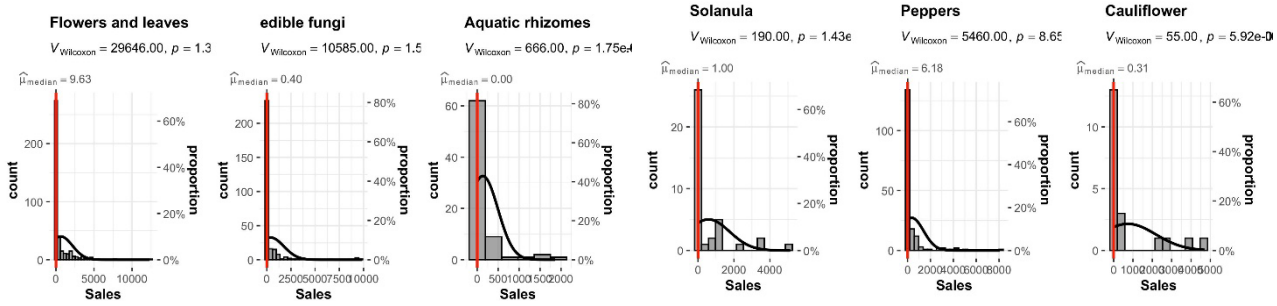

Figure 8. Histogram of frequency distribution of sales of six types of vegetable commodities

Based on the above frequency distribution histogram, it can be intuitively seen that the sales of six categories of vegetable commodities show a downward trend, and the main source of profits mainly depends on the sales of vegetable commodities. In this paper, replenishment and pricing are combined to maximize the revenue of the objective function, the parameter domain of each algorithm is tested, and the parameters are set in this paper. The relationship between total sales volume and time, discount degree, normal average price and loss rate was calculated by repeated calculation of each parameter 40 times.

The specific test of the relationship between total amount of vegetable sales and time is shown in Table 5:

Table.5. Relationship between total amount of vegetable sales and time of in the vegetable category

category	*Relationship between total sales and time*	R^2
cauliflower	$y=4\times10^{-5}x^2-3.1629x+70821$	$R^2=0.5880$
anthophyllum	$y=0.0002x^2-20.708x+461502$	$R^2=0.5851$
capsicum	$y=0.00001x^2-9.7567x+216170$	$R^2=0.1493$
solanacea	$y=0.0004x^2-39.483x+877514$	$R^2=0.5501$
edible mushroom	$y=0.0001x^2-11.792x+261052$	$R^2=0.6820$
aquatic rhizomes	$y=-4\times10^{-5}x^2-3.5571x-79334$	$R^2=0.0161$

The specific test of the relationship between total wastage and the degree of discount in the vegetable category is shown in Table 6:

Table.6. Relationship between total wastage and level of discounting in the vegetable category

category	*Relationship between total sales and time*	R^2
cauliflower	$y=-0.0019x^2-0.0286x+0.1274$	$R^2=0.5091$
anthophyllum	$y=7\times10^{-7}x^2-0.0005x+0.4113$	$R^2=0.5230$
capsicum	$y=1\times10^{-6}x^2-0.0008x+0.2007$	$R^2=0.2889$
solanacea	$y=2\times10^{-7}x^2-0.0003x+0.1476$	$R^2=0.2550$
edible mushroom	$y=-1\times10^{-6}x^2-0.0016x+0.0492$	$R^2=0.0461$
aquatic rhizomes	$y=9\times10^{-7}x^2-0.0006x+0.3146$	$R^2=0.4911$

The specific test of the relationship between total sales and normal average unit price for the vegetable category is shown in Table 7:

Table.7. Relationship between total sales volume and normal average unit price for the vegetable category

category	*Relationship between total sales and time*	R^2
cauliflower	y=-0.0668x^2-1.2704x+57.335	R^2=0.0851
anthophyllum	y=3.5707x^2-58.32x+401.44	R^2=0.5732
capsicum	y=0.4188x^2-12.751x+164.02	R^2=0.0884
solanacea	y=0.3664x^2-8.029x+61.576	R^2=0.58
edible mushroom	y=-0.5805x^2-8.5827x+55.739	R^2=0.0981
aquatic rhizomes	y=0.0122x^2-2.7509x+64.855	R^2=0.4891

The specific test of the relationship between total sales and wastage rate for the vegetable category is shown in Table 8:

Table.8. Relationship between total sales volume and wastage rate in the vegetable category

category	*Relationship between total sales and time*	R^2
cauliflower	y=-1.0905x^2+27.988x-132.17	R^2=0.5711
anthophyllum	y=6.2495x^2-162.73x+1217.2	R^2=0.1553
capsicum	y=3×10^{-6}x^2-0.0029x+8.1203	R^2=0.2461
solanacea	y=-2.087x^2+31.716x-95.09	R^2=0.2102
edible mushroom	y=-0.056x^2-3.5553x+105.85	R^2=0.4330
aquatic rhizomes	y=-0.002x^2-2.6383x+63.536	R^2=0.1453

Based on the analysis of the relationship between the total sales volume and time, discount level, normal average price, and loss rate in Tables 5-7, it can be concluded that when R^2>0.5, it indicates a strong correlation, and when R^2<0.5, it indicates a low correlation.

3.2.5 Vegetable Commodities Pricing and Replenishment Forecast Results

Based on the relationship analysis between total sales volume and time, discount degree, normal average price and loss rate, this paper predicts the pricing strategy of vegetable commodities in the next week, as shown in Figure 9, and predicts the replenishment volume of vegetable commodities in the next week as shown in Figure 10.

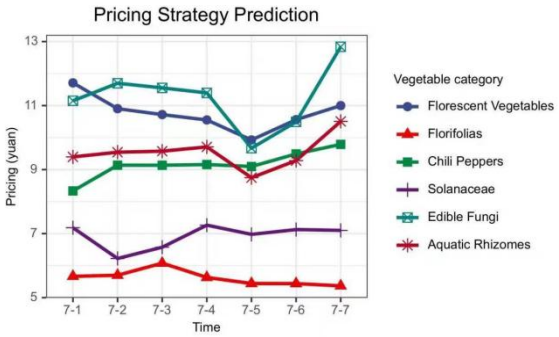

Figure 9. Vegetable commodity next week pricing strategy chart

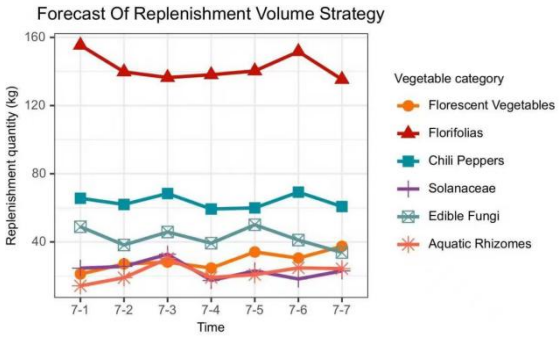

Figure 10. Vegetable commodities next week replenishment chart

It can be seen that cauliflower products reach the maximum profit and minimum standard error in the pricing interval of (9,12) and the replenishment interval of (10,40); The maximum profit and minimum standard error are achieved in the pricing interval of (5,7) and the replenishment interval of (120,160); The maximum profit and minimum standard error are achieved in the pricing interval of (8,10) and the replenishment interval of (50,80). The maximum profit and minimum standard error are achieved in the pricing range of (6,8) and the replenishment range of (10,40) of the nightshade category; The maximum profit and minimum standard error of edible fungi are achieved in the pricing interval

of (9,13) and the replenishment interval of (30,60). Aquatic root products achieve maximum profit and minimum standard error in the pricing interval of (8,11) and replenishment interval of (10,40).

Based on the average profit, the results of ADEA algorithm are generally consistent with the expected results. Based on the standard deviation, ADEA algorithm has a low standard deviation, but the retrieval of ADEA algorithm is fast. In terms of target profit of vegetable category, Mosaic category has the highest profit, chilli category and edible fungus category are roughly the same, but there is a big profit difference compared with Mosaic category, cauliflower category and aquatic root category are relatively flat and general, and nightshade category has the lowest target profit.

Based on the existing data, the sales volume and sales price of single products on a single day on July 1, 2021 and 2022 are analyzed, and the available varieties on July 1 are counted to analyze the income situation. Subsequently, this paper processed the data from June 24 to 30, 2023 to obtain the total number of available items and information details. Combined with quantitative indicators, the sales unit price, discount degree, market demand, transport loss rate and purchase price of each category of single product are obtained respectively, as shown in Table 9:

Table.9. Concrete quantified result

Item Name	*requirement*	*purchase price*	*Sales unit price*	*Transport loss rate (%)*
Spinach (portion)	9.8	4.134897959	5.5102	9.43
Tall Melon (1)	3.004428571	3.236315789	13.193	29.25
Sea mushrooms (bun)	8.857142857	1.953387097	2.74839	0
Sweet potato tips	4.506571429	3.190740741	5.33333	8.42
Honghu lotus root belt	4.033142857	18	20.0909	24.05
rugosa(part)	11.28571429	10.40924051	2.58608	9.43
Yunnan lettuce(part)	32.28571429	2.502743363	4.4615	9.43
Solanum japonicum	4.214428571	6.984931507	12	6.9
Bamboo leaf	13.29671429	4.495348837	3.77395	13.62
Eggplant (2)	10.90285714	10.17432624	6	6.07

In view of the above quantitative results, comprehensively considering the specific situation of category sales and sales of the highest single product as well as the number of samples, this paper extracts three indicators for analysis. After repeated debugging, three types of differences and differences can be roughly distinguished, and finally presented and processed in the form of three-dimensional scatter plots, as shown in Figure 11-14:

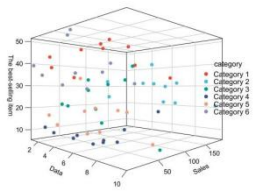

Figure11. Relationship between category sales and top unit sales (1-10 July 2020)

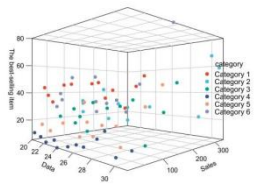

Figure12. Relationship between category sales and top unit sales (1-10 July 2021)

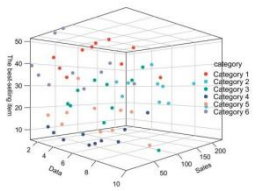

Figure13. Relationship between category sales and top unit sales (1-10 July 2022)

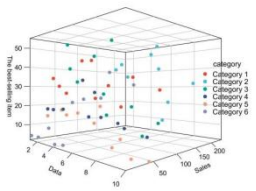

Figure13. Relationship between category sales and top unit sales (20-30 June 2023)

4. Conclusions

The trend of massive data of Vegetable and Plant commodities provides a basis for big data analysis and the establishment of prediction models. However, it is difficult for traditional prediction methods to calculate large amounts of data and the result error is large. Meanwhile, overfitting will also affect

the prediction accuracy and cause decision-making errors. In this paper, Adaptive Differential Evolution Algorithm is used to establish a revenue-maximization optimization model, making full use of Adaptive Differential Evolution Algorithm is powerful data processing function and preventing overfitting function. The experimental results show that the Adaptive Differential Evolution Algorithm model has good predictability and feasibility, and has certain practical application value.At the same time, based on the analysis of the relationship between the total sales volume of vegetables and plant categories, time, discount level, and other factors, the pricing strategy for the next week and the prediction of replenishment volume for the next week are very effective, and have strong practical applications for the operation of fresh supermarkets.

References

[1] Yali Li. Joint replenishment and pricing collaborative decision making considering investment in preservation technology[D]. Chongqing Jiaotong University,2023.

[2] Jing Lu . Research on inventory control and dynamic pricing of fresh agricultural products[D]. Tianjin University,2021.

[3] Si-Hong Gu . A study on dynamic pricing of fresh products in H-retailers considering freshness variation[D]. Donghua University,2023.

[4] Hui-Hong Wu,Yu Lin,Xi Zeng,Shu-Qiu Qian. Adaptive differential evolutionary algorithm and its application to dynamic environmental economic scheduling problems[J]. Computer Application Research,2021,38(05):1443-1448+1454.

[5] Cai X. Research on Profit Maximisation of Community Fresh Food Stores under Nonlinear Inventory Holding Cost[D]. Hefei University of Technology,2021.

[6] Yanbo Li,Ruochen Li,Bo Shi,Junshuo Chen. Energy-efficient optimisation of highway service area self-consistent energy system based on improved simulated annealing genetic algorithm[J]. Journal of Xi'an Jiaotong University:1-10..

[7] Leiyan Xu, Xina Yue, Yeng Cao. Supply and sales balance management strategy of directly operated chain enterprises based on profit maximisation[J]. Journal of Jiangsu College of Economics and Trade, 2023,(02):21-24.

[8] Tianshan Yang,Gonglin Yuan. Research on Dynamic Pricing and Replenishment Strategy of Perishable Goods Considering Green Distribution Technology Investment[J]. China Management Science:1-13.

[9] Yunhao Cui. Research on fresh vegetables sales prediction based on CNN-PSO-LSTM combination model[D]. Anhui Agricultural University,2023.

[10] Shuyun Wang,Yingmei Jiang,Xianjie Wang. A three-level integrated cold chain inventory model based on distribution centres[J]. Journal of Systems Management,2017,26(02):390-398.