Optimization Of Automatic Pricing and Replenishment Decision Based on Time Series and BP Neural Network Model

Chang Sun [1, *, #], Xiaohan Li [2, #], Zhiheng Lai [2, #]

[1] Business and Intellectual Property Academy, Dalian University of Technology, Panjin, China, 116024

[2] Leicester International Academy, Dalian University of Technology, Panjin, China, 116024

* Corresponding Author Email: 2836797076@qq.com

[#]These authors contributed equally

Abstract. In major fresh food supermarkets, vegetable products often have a short shelf life and poor product quality, which affects their sales. Therefore, it is particularly important to make replenishment decisions for each vegetable category on the same day without knowing the specific individual product and purchase price. This article focuses on the problem of automatic pricing and replenishment decision-making for vegetable products. Based on the fact that vegetable products have a relatively short shelf life, the automatic pricing and replenishment strategies for vegetable products are optimized, indirectly proving the correlation between certain individual products. A mathematical model is constructed using time series and BP neural network models to obtain corresponding replenishment strategies, Finally, provide recommendations on automatic pricing and replenishment decisions for vegetable products. The "cost plus pricing" method involved in this question plays a crucial role in the pricing of vegetables. At the same time, from the perspectives of the demand side and supply side, the limitations of sales time and sales space are also particularly important.

Keywords: Time series prediction model, BP neural network model, Automatic pricing, Replenishment decision.

1. Introduction

In the current vegetable buying and selling market, due to the characteristics of vegetable products such as short shelf life, high seasonal price fluctuations, and sales volume being affected by weather and temperature changes, vegetable supermarkets often need to make replenishment decisions without knowing the exact individual product and purchase price every day. However, due to multiple influencing factors, merchants often find it difficult to make the decision that maximizes profits, Therefore, this article analyzes and studies it to help supermarkets make the best replenishment decision. Previously, research on this issue mostly focused on analyzing the price changes of each individual item separately [1]. However, this article included individual items in the "shopping basket" based on the relevant system and analyzed the relationship between individual items [2]. Some studies focus on studying spatial differences from the perspective of market supply and demand relationships [3]. And analyze the changes in vegetable prices from the perspective of the country's macroeconomic regulation of the market. In summary, research often uses mathematical methods as tools [4], selecting one or more factors as the focus for analysis, reflecting the fluctuations in vegetable prices under different priorities [5].

The purpose of this article is to develop a replenishment strategy that maximizes the revenue of a supermarket by using the sample information of six vegetable categories distributed by a supermarket, as well as the sales flow details and wholesale prices of each product from July 1, 2020 to June 30, 2023, as the specific individual items and purchase prices of vegetable products are unknown. This article analyzes the distribution patterns and interrelationships of sales volume of various categories and individual products of vegetables. In order to establish a comprehensive scientific quantitative analysis system, this article calculates the highest monthly sales of individual products and determines whether they are discounted. Then analyze the category and establish an indicator system or time series model for the relevant data before conducting relevant analysis, processing, and judgment. Next, this article analyzes the quantitative relationship between the sales volume of various vegetable categories and the cost-plus pricing [6] and predicts the daily sales volume of each category in the next week based on historical data. At the same time, an objective function is established to maximize the weekly total revenue of supermarkets, and constraints are established, such as the daily replenishment volume being greater than or equal to the predicted sales volume. Finally, mathematical optimization methods are used to solve the replenishment plan and pricing strategy. Finally, this article establishes the variable "salable items" and establishes a BP neural network model based on its data to predict the sales volume of each item as of July 1, 2023. Then use the moving average method to predict the wholesale price of sellable items and consider the corresponding quantity of item losses after restocking.

2. Research design

2.1. Model Assumptions

(1) Variable type: Two variables should belong to interval or proportional variables.
(2) Linear relationship: There is a certain linear relationship between two variables.
(3) Normal distribution: Two variables should roughly follow a normal distribution.
(4) Data pairing: Each observation data in the dataset includes paired data.
(5) No outliers Q: The dataset should not include extreme outlier data.

2.2. Symbol Description

The table1. Symbol Description is described below:

Table 1. Symbol Description

symbol	explain	unit	symbol	explain	unit
B_t	Total profit	element	a_{3i}	The coefficient of the third power of the selling price	/
B_i	Profit of each individual product	element	a_{0i}	The coefficient of the zero power of the selling price	/
n_i	Sales volume of each category	Kilogram	wp_i	Wholesale prices for each category	element
p_i	Selling prices for each category	element	N_i	Predicted sales volume of each individual product	individual
a_{1i}	Primary coefficient of selling price	/	α_i	Loss rate of each individual product	%
a_{2i}	Coefficient of the quadratic term of the selling price	/	sp_i	Price of each item	element
rp_i	Wholesale prices of each individual product	element			

3. Empirical analysis

This article first preprocesses the obtained data: merge the data so that the relevant data can display both sales information and product name and category information in a new file. Since we need to find the distribution pattern and mutual influence of sales volume for each category and individual product, we should first find the distribution pattern of sales volume for each category and individual product. Based on the available data, this article speculates that the distribution of sales volume should be related to time and discounts. Next, this article will calculate the sales volume of each category in each month, and the number of discounts in each month. Then, this article will visualize the obtained data, as shown in the following figure1. Monthly sales volume of each category over time per year and figure 2. Changes in the total number of monthly discounts per year:

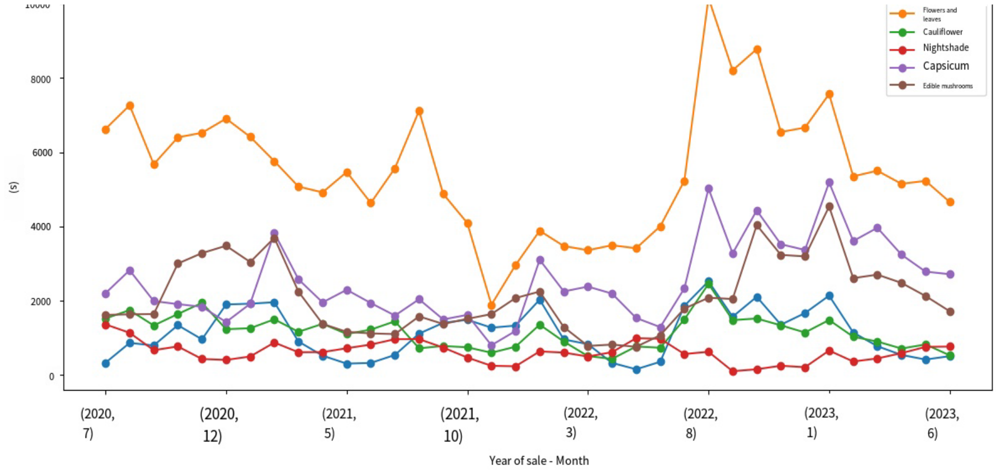

Figure 1. Monthly sales volume of each category over time per year

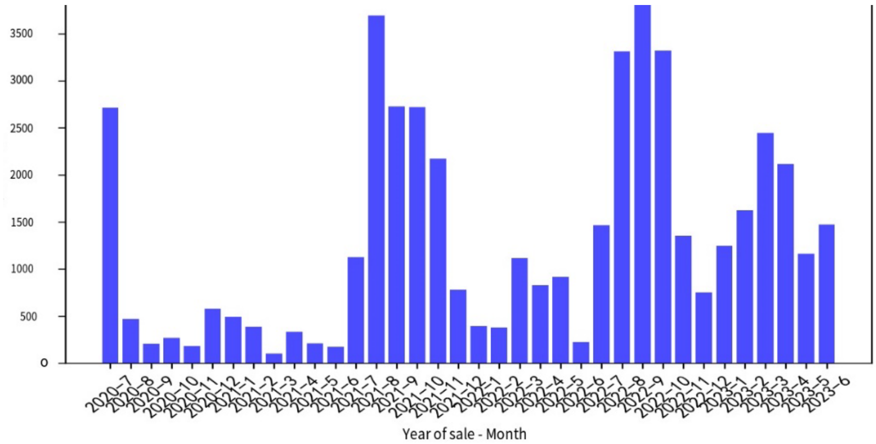

Figure 2. Changes in the total number of monthly discounts per year

This article found that when the total number of discounts for various vegetable products increases, there is generally a situation of sales growth. When the number of discounts decreases, there is also a situation of sales decline, which is in line with market rules. Based on the experimental data, this article also found that there are more discounts in summer. It is speculated that fresh vegetables are not easy to store and can only be sold at a discount if they are not sold out on the same day.

Then, based on the changes in sales volume of each category over time, this article adopts the Pearson correlation coefficient analysis method [7] to calculate the correlation coefficients between each category. Then visualize it as follows figure 3. Category sales correlation heat map:

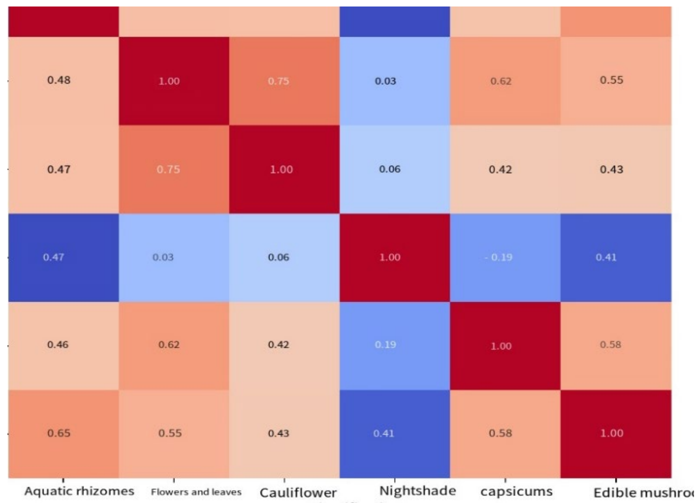

Figure 3. Category sales correlation heat map

The judgment method adopted in this article is: two groups of categories with an absolute correlation value greater than 0.7 are considered to have strong correlation, those between 0.5 and 0.7 are considered to have strong correlation, Those between 0.5 and 0.7 are considered to have strong correlation, those between 0.3 and 0.5 are considered to have weak correlation, those between 0.5 and 0.5 are considered to have weak correlation, and those with an absolute value below 0.3 are considered to have no correlation. In addition, for individual items, this article introduces the concept of a shopping basket. In this article, shopping baskets are used to store single item sales data. When the sales time interval between two single items is within 10 seconds, they are considered to be sold in the same batch, and then they will be placed in the same shopping basket. This article considers items from the same shopping basket as potentially related. This article first filters the shopping basket, and only single item shopping baskets will be removed. There may be duplicate items in the shopping basket, and duplicate items will also be removed. Finally, this article conducted statistical analysis on the combinations of various shopping baskets and found a large number of repeated combinations, indirectly proving the association of certain individual items [8]. The visualization results of this article are illustrated as follows figure 4. Number of occurrences of the first twenty combinations:

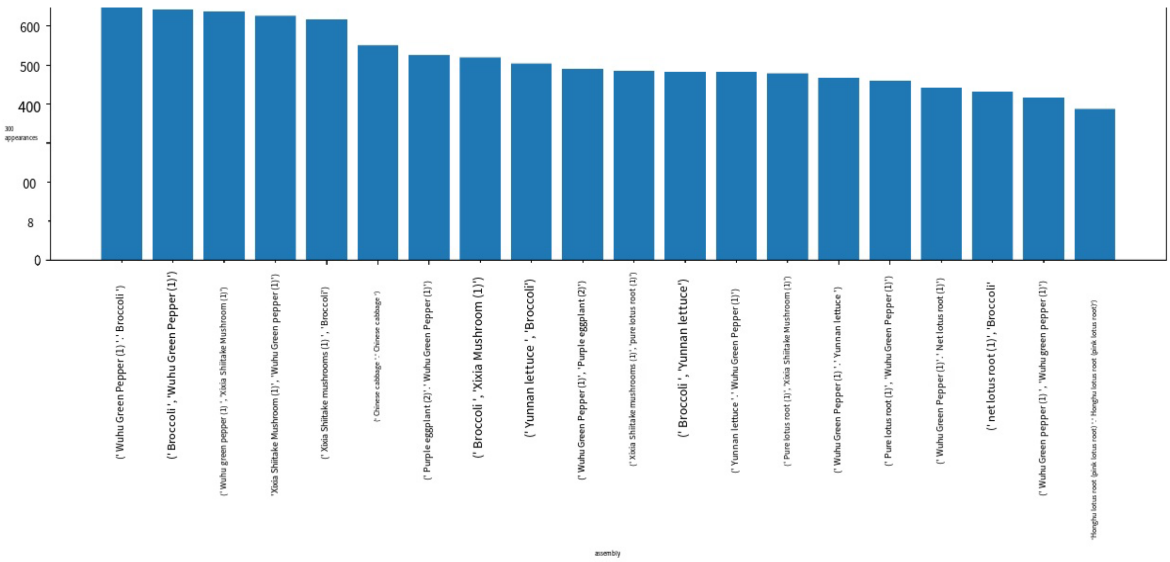

Figure 4. Number of occurrences of the first twenty combinations

Next, this paper aims to establish the relationship between the total sales volume of each category and cost-plus pricing. Therefore, this article chooses a regression model to find the functional relationship between the total sales volume of each category and cost-plus pricing.

Cost plus pricing refers to pricing=cost * (1+profit margin) and setting prices in this way can ensure that vendors earn a certain profit when completing a sale. However, this article considers that consumers' most intuitive perception of a product is its selling price, and they only have a vague understanding of the cost price of vegetables. In order to simplify the model, this article equates the problem to analyzing the relationship between the total sales volume of each category and the selling price. In order to facilitate the analysis of categories, we make the average unit price of all individual items in a certain category on that day the unit price of the category, and the sales volume is calculated based on the total amount on that day [9].

Because the relationship between total sales and average unit price will be used in subsequent articles, polynomial regression is adopted here to facilitate the finding of the functional relationship between the two. Considering the accuracy of the model and preventing overfitting, this article sets the highest coefficient of the polynomial to 3. The resulting relationship is:

$$n_i = a_{1i} \cdot p_i + a_{2i} \cdot p_i^2 + a_{3i} \cdot p_i^3 + a_{0i} \tag{1}$$

Among them, n_i indicates the quantity of each category, p_i represents the selling price of each category in yuan, a_{1i} represents the coefficient of the primary term of the selling price, a_{2i} the coefficient representing the quadratic term of the selling price, a_{3i} the coefficient representing the cubic power of the selling price, a_{0i} represents the zero orders coefficient of the selling price.

The wholesale prices of each category are also calculated based on the average price. Then, based on the average prices of each category over the past three years, predict the wholesale prices for that week using a time series prediction model [10]. After data analysis and calculation, the following table 2. Calculation Results' results are obtained:

Table 2. Calculation Results

	Aquatic rhizomes	**Florifolias**	**florescent vegetables**	**Solanaceae**	**Chili peppers**	**edible fungi**
July 1st	12.06	3.43	7.84	4.42	5.99	4.71
July 2nd	12.05	3.45	7.84	4.42	5.90	4.68
July 3rd	12.06	3.45	7.84	4.42	5.93	4.69
July 4th	12.05	3.45	7.84	4.42	5.92	4.69
July 5th	12.05	3.45	7.84	4.42	5.92	4.69
July 6th	12.05	3.45	7.84	4.42	5.92	4.69
July 7th	12.05	3.45	7.84	4.42	5.92	4.69

This article establishes the following optimization model based on the seven day wholesale price obtained and the functional relationship between sales volume and selling price of each category obtained earlier:

$$max\ B_z = n_i \times (p_i - wp_i) \tag{2}$$

$$\begin{cases} p_i < p_{i\_max} \\ p_i > p_{i\_min} \end{cases} \tag{3}$$

Among B_z represents total profit, wp represents wholesale prices for each category.Based on the above model, the results obtained after relevant calculations in the table 3. Pricing Results are as follows:

Pricing:

Table 3. Pricing Results

Aquatic rhizomes	**Florifolias**	**florescent vegetables**	**Solanaceae**	**Chili peppers**	**edible fungi**
18.54	5.33	11.73	11.70	19.96	13.53

The table 4. Total Replenishment Results is described below:

Table 4. Total Replenishment Results

Aquatic rhizomes	**Florifolias**	**florescent vegetables**	**Solanaceae**	**Chili peppers**	**edible fungi**
78.83	1208.60	228.34	106.98	705.45	309.77

This article can also provide the replenishment quantity and pricing strategy for individual products on July 1st. Firstly, we obtain information about all units with sales data, store it in variables of available units, and save it in a file. Li Zai predicted the sales volume of each individual product on July 1st by establishing a BP neural network model. Next, this article predicts the wholesale price of sellable items. Finally, consider the loss rate of each available item, and after considering replenishment, there will be a corresponding amount of item loss.

Based on the above ideas, we establish the following equation:

$$= N_i * (1 - \alpha_i) * (sp_i - rp_i) \tag{4}$$

Among B_i represents the profit of each individual product, N_i represents the predicted sales volume of each individual product, α_i represents the loss rate of each individual product, sp_i represents the selling price of each item, rp_i represents the wholesale price of each individual product.

Because if the purchased goods are not sold out on the same day, it will also cause losses. Therefore, if the purchased quantity exceeds the sales volume:

$$B_i = N_i * (1 - \alpha_i) * (sp_i - rp_i) - (rN_i - N_i) * rp_i \tag{5}$$

If the predicted sales volume of each sellable item is less than 2.5 kilograms, the purchase volume is calculated as 2.5 kilograms. If the predicted sales volume is greater than or equal to 2.5. The purchase volume is calculated as 2.5 kilograms. If the predicted sales volume is greater than or equal to 2.5 kilograms, it is calculated as the predicted sales volume.

According to the above formula, if the predicted sales volume of each sellable item is less than 2.5 kilograms, the purchase volume is calculated as 2.5 kilograms [11]. If the predicted sales volume is greater than or equal to 2.5. The purchase volume is calculated as 2.5 kilograms [11]. If the predicted sales volume is greater than or equal to 2.5 kilograms, it is calculated as the predicted sales volume.

If analyzed according to the actual situation, due to the discount rate, some individual products may have a profit lower than 0. If the number of units with profit greater than or equal to 0 is greater than or equal to 33, the 33 units with the highest profit will be selected. If the number of units with profit greater than or equal to 0 is less than 33 but greater than or equal to 27, all units with profit greater than or equal to 0 will be selected. If the number of units with profit greater than or equal to 0 is less than 27, Then select a few items with a profit less than 0 until 27 items with a profit greater than or equal to 0 are selected.

The following figure5. Profit of each individual product shows the profit of some available items:

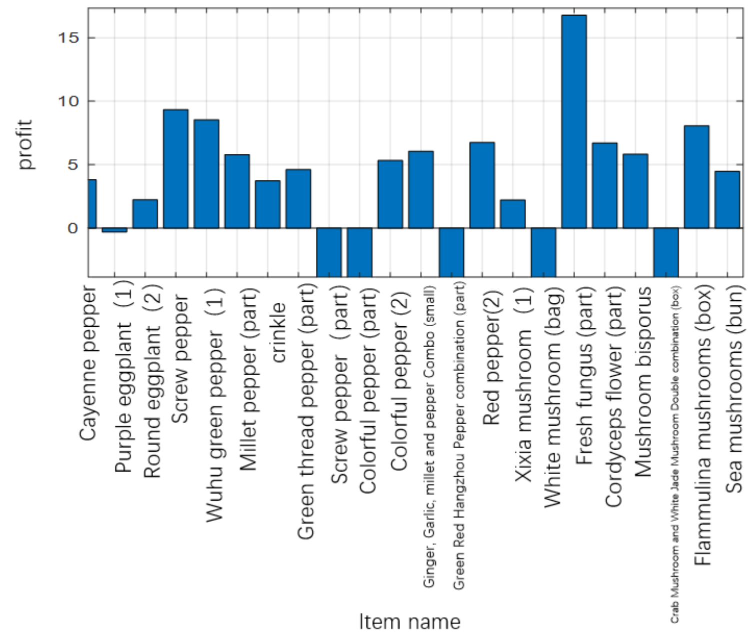

Figure 5. Profit of each individual product

This is based on the above approach to obtain the predicted profit values for each individual product. Finally, the replenishment strategy is obtained as the table 5. Predicted Profit Value of Single Product:

Table 5. Predicted Profit Value of Single Product

Product Name	Predicted price (yuan)	Predicted sales volume (yuan)	Replenishment quantity (kg)	Product Name	Predicted price (yuan)	Predicted sales volume (yuan)	Replenishment quantity (kg)
Purple Eggplant (2)	8.71	3.75	3.75	Green Pepper (portion)	4.10	3.50	3.50
Crab flavored mushroom and white jade mushroom combo (box)	8.03	3.50	3.50	Screw pepper	10.47	2.27	2.50
White Jade Mushroom (Bag)	7.01	3.48	3.48	Echinacea officinalis	15.50	1.75	2.50
Xiaomi Pepper (portion)	5.08	3.50	3.50	Little Green Vegetables (1)	5.70	1.84	2.50
Long line eggplant	10.46	2.80	2.80	Yunnan Lettuce (portion)	4.47	3.50	3.50
Cordyceps Flower (portion)	5.35	3.50	3.50	Fresh agaric (portion)	2.19	3.50	3.50
Golden Needle Mushroom (Box)	3.76	3.51	3.51	Green Red Hangzhou Pepper Combination Package (portion)	3.83	3.50	3.50
Agaricus bisporus (box)	5.34	3.50	3.50	Amaranth	4.87	1.72	2.50
Mushroom (portion)	3.57	3.50	3.50	baby cabbage	5.02	3.50	3.50
Purple Eggplant (1)	11.09	1.71	2.50	Shanghai Qing	7.52	1.59	2.50
Round Eggplant (2)	7.35	2.14	2.50	Bamboo leaf vegetable	6.20	1.06	2.50
Small wrinkled skin (portion)	3.39	3.50	3.50	Yunnan lettuce	8.29	0.93	2.50
Combination of ginger, garlic, millet, and pepper (small portion)	4.04	3.50	3.50	Combination of ginger, garlic, millet, and pepper (small portion)	4.03	3.50	3.50
Screw Pepper (portion)	4.93	3.51	3.51	broccoli	9.97	1.25	2.50
Seafood Mushroom (Bun)	3.23	3.49	3.49	Amaranth	4.86	1.72	2.50
Yunnan Youmai Cai	4.03	3.50	3.50	Milk cabbage	4.95	1.11	2.50

4. Conclusion

The advantages of the model selected in this article are as follows: it can capture the dependency relationship between time points, reflect the dynamic changes in the time series, and after parameter learning, it can well fit the trend of historical data changes and improve prediction accuracy. The parameters used in this model, such as autoregressive terms, are easy to interpret and reflect the generation mechanism of time series. It can also handle various time series data regardless of their sampling frequency, such as daily data and weekly data. In addition, this model can be used for online learning and prediction as the data is constantly updated, making it suitable for real-time prediction needs. However, it cannot be denied that the model used in this article still has some shortcomings, such as the fact that this model has non-stationary characteristics such as trend and periodic changes, and it is difficult to capture complex nonlinear relationships, which are mainly based on linear assumptions.

The correct understanding of the development laws of objective things is a prerequisite for achieving scientific prediction, only in this way can there be high-precision speculation and judgment on the trends of vegetable pricing and replenishment. Previous studies have shown that vegetable pricing and restocking are mainly influenced by two factors: changes in supply and demand relationships, and the inherent characteristics of vegetables. In terms of factors affecting pricing replenishment, this article makes the following speculations:

For the factors related to holidays, this article believes that consumption patterns during holidays such as the Spring Festival are different from workdays and need to be distinguished and predicted as shown in the figure 6. Weekly average price changes of cabbage, celery, and cauliflower:

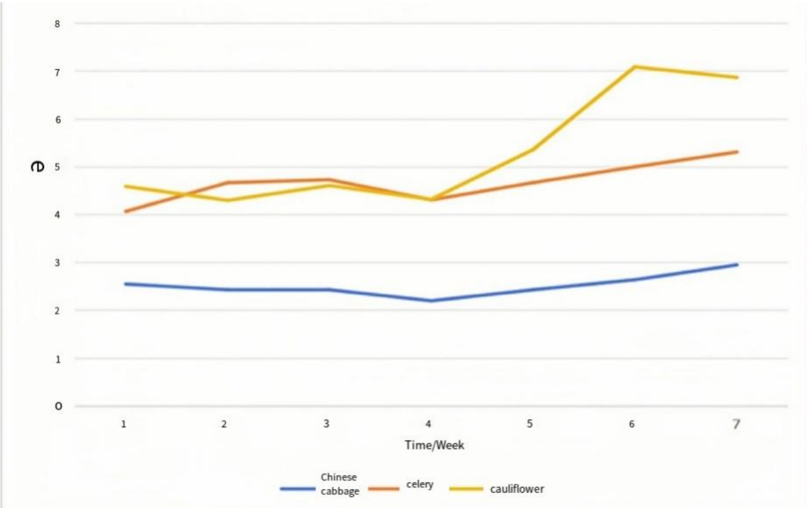

Figure 6. Weekly average price changes of cabbage, celery, and cauliflower

This article analyzes the price changes of three main vegetables in Fujian Province before and after the National Day holiday in 2021. It is found that the average prices of the three vegetables have decreased during the National Day period (4th week). Through market competition principles analysis, if the pricing of supermarkets remains unchanged at this time, it will lead to a significant decrease in sales and loss of profits. Therefore, merchants should adjust their pricing according to the market to ensure sales and profits.

In terms of the quantity of vegetable suppliers, this article believes that the reason why the quantity of vegetable suppliers will have an impact on pricing and replenishment decisions in the short term is that the supplier may suddenly increase or decrease while demand remains roughly unchanged, such as suddenly opening or closing other vegetable supermarkets as shown in the figure 7. Supply Curve and Demand Curve:

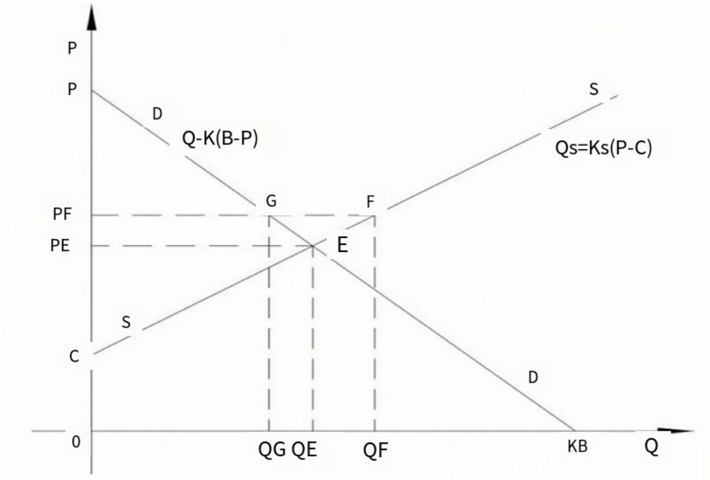

Figure 7. Supply Curve and Demand Curve

According to the supply and demand curve, when the supply side increases, in order to maintain profitability and supply-demand balance, it is necessary to reduce daily replenishment volume. When the supply side decreases, in order to maintain market competitiveness and supply-demand balance, it is necessary to increase the daily replenishment volume. Therefore, it is necessary to constantly understand and master the quantity of vegetable suppliers, and under the conditions of complying with objective laws - supply and demand theory, exert subjective initiative to adjust pricing and replenishment strategies, ultimately achieving maximum profit.

Next, this paper can fuse multiple time series models such as ARIMA and RNN to enhance the overall prediction ability by utilizing the strengths of each model. This text can also try using deep learning networks such as LSTM and CNN to extract time series features and improve traditional time series models. To make the predicted model more accurate, this article continuously adopts optimization algorithms to optimize the hyperparameters of the model and improve the accuracy of the BP neural network prediction model. This paper can use the research results of this article to make more people aware of the importance of automatic pricing and replenishment decisions for vegetable products, and actively contribute the authors' own strength and wisdom to solving such problems.

References

[1] Yan Haowei, Mu Yueying. Research on the Construction of Vegetable Industry Monitoring and Early Warning System Based on Vegetable Market Price Monitoring and Early Warning [J]. Vegetable, 2023 (05): 1-10

[2] Zhu Wenjuan. A Study on the Spatial Differences between Vegetable Price Fluctuation and Consumer Price Elasticity [J]. Price Theory and Practice, 2022 (08): 72-75

[3] Zheng Yuxin Prices fluctuate significantly, and supermarkets are engaged in a vegetable supply guarantee war Economic Observer, 2021-11-01 (005)

[4] Zhang Bi. Characteristics of Vegetable Price Fluctuation in China: An Empirical Analysis Based on the PVAR Model [J]. China Business Review, 2021 (10): 21-24

[5] Zhang Qianqian Research on Vegetable Demand Prediction System Based on SVM [D]. Beijing Jiaotong University, 2015

[6] Han Junhua, Gan Shengdao. Introduction to Cost Plus Pricing Method [J]. Finance and Accounting Monthly, 2012 (22): 74-75.

[7] Cheng Juanjuan. Empirical Study on the Relationship between Scientific Research and Teaching in Universities - Analysis Based on Pearson Correlation Coefficient [J]. China University Science and Technology, 2022 (10): 46-52

[8] Maurissa Research on Pricing Strategies and Production and Sales Models of Vegetable Wholesale Market from the Perspective of Supply Chain [D]. Central South University of Forestry and Technology, 2022

[9] Wang Jue. Post-Keynesian Enterprise Pricing Model - Cost Plus Pricing Principle [J]. Journal of Lanzhou University, 2003 (03): 99-102

[10] Guo Shuang, Zheng Yutong. Analysis of Marketing Strategies for Agricultural Products in Jilin Province Based on Time Series Models [J]. Market Modernization, 2022 (17): 55-57

[11] Cui Yunhao Research on Fresh Vegetable Sales Forecast Based on CNN-PSO-LSTM Combination Model [D]. Anhui Agricultural University, 2022