Research on pricing and replenishment strategy of superstore goods based on linear regression and gray prediction models

Wangyi Jia*

School of Mathematics and Statistics, Kashi University, Kashi prefecture, Xinjiang Uygur Autonomous Region, China, 844008

* Corresponding Author Email: wyj666wyj666wyj@163.com

Abstract. The pricing and replenishment strategy of various commodities in superstores have a significant impact on the revenue. This paper aims to address the automatic decision-making problem for pricing and replenishment of vegetable commodities in superstores. It applies the linear regression model to predict commodity prices, aiming to maximize the revenue of superstores. Additionally, it utilizes the gray prediction model to forecast replenishment volume. Based on marketing data from 2020 to 2023 for vegetable goods at a specific superstore, this study provides the replenishment quantity and pricing strategy for each vegetable category from July 1-7, 2023, as well as for each individual vegetable item on July 1, 2023.

Keywords: linear regression, gray prediction, MATLAB software, pricing and replenishment.

1. Introduction

In fresh produce superstores, the freshness period of general vegetable commodities is relatively short, and their quality deteriorates with an increase in sales time. Most varieties cannot be resold on the next day if they are not sold on that day. Therefore, supermarkets usually replenish their stocks daily based on historical sales and demand for each commodity to maximize revenue. Since supermarkets sell many different varieties of vegetables from various origins, merchants must make replenishment decisions for each vegetable category on the same day without knowing the specific product or purchase price.

Fewer studies have been conducted on commodity pricing in fresh superstores, despite several methods being used for commodity pricing and replenishment.Zhang X et al.0 Using time series model and gray prediction model and integrated MATLAB software to predict and analyze product sales, and found the cyclical pattern of product sales.Ruihai L et al.[2] Pricing and batch decision making of perishable commodities were investigated and strategies for pricing and replenishment of commodities under different conditions of selling price, reference price, product freshness and display inventory were derived.Liu L et al.[3] Demonstrated the existence of an optimal PT investment and fresh food pricing in a study of optimal fresh food technology investment and pricing policy.Qu T et al.[4] Investigated the problem of demand forecasting and price optimization for a semi-luxury supermarket and predicted weekly demand based on machine learning algorithms of regression trees and random forests.Xu et al.[5] Studied the horizontal replenishment strategy of fresh produce in supermarket chains, and provided a reference for supermarket chains to cope with shortage or surplus of fresh produce.Tang L et al.[6] Provided a multi-stage inventory model based on shelf life for fresh food in dynamic pricing and inventory temperature decision making for variable demand fresh food.Junxuan L et al.[7] Investigated dynamic inventory allocation for seasonal goods in Dillard's and developed an efficient heuristic algorithm for dynamic inventory allocation.Deniz B et al.[8] Comparison of inventory policies for perishable goods and proved that a fixed order quantity policy is optimal under certain conditions.Liang D et al.[9] Under the new information principle of gray system theory, the new information improved GM(1,1) model is used to predict the raw materials of a sales dessert, which solves the waste caused by excessive order quantity and the loss caused by insufficient order quantity.

2. Modeling

2.1. Linear regression model

2.1.1 Model analysis

To obtain the total daily replenishment volume and pricing strategy for each vegetable category, this paper forecasts the total daily replenishment volume and pricing strategy for each vegetable category from July 1-7, 2023 as an example. The relationship between sales volume and cost-plus pricing is analyzed to establish the correlation between replenishment volume and sales volume in each vegetable category.

2.1.2 Modeling

Let Q_t be defined as the total sales volume of each type of single vegetable commodity, and let P_t be defined as the average sales price of each category in a certain week be determined, obtaining the relationship formula between the two.

$$Q_t = \hat{a} + \hat{b} P_t + \varepsilon_t \tag{1}$$

where ε_t is the random error.

The replenishment volume of the superstore in a day can be divided into two parts: the daily sales volume and the portion of the replenishment volume that is lost, assuming Z_t represents the replenishment volume and λ represents the average loss rate of each commodity during the period ($0<\lambda<1$), which can be obtained through a relationship equation.

$$Z_t = Q_t + \lambda Z_t \tag{2}$$

Further simplification yields

$$Z_t = \frac{Q_t}{1-\lambda} \tag{3}$$

From equation (1)

$$Z_t = \frac{a + bP_t}{1-\lambda} \tag{4}$$

The above formula can be incorporated into the profit expression, The profit is supposed to be $\pi(P_t)$,assuming that the profit is calculated as the daily sales of goods sold minus the cost of replenishment and lost goods. By using the aforementioned formula, we can obtain an accurate expression for profit.

$$\pi(P_t) = (a + bP_t)\left(P_t + \frac{2C_t}{1-\lambda} + 1\right) + \varepsilon_t \tag{5}$$

2.2. Gray prediction model

2.2.1 Model analysis

Based on the available data, the varieties of each individual vegetable product can be screened out from June 24-30, 2023. Due to the small sample size, this falls under the problem of cycle prediction for a single sample. Therefore, the gray prediction model can be used to predict the replenishment quantity and pricing strategy for each individual vegetable product on July 1.

The profitability of each vegetable item is ranked based on the data obtained from the model in order to determine which items are selected for sale.

2.2.2 Modeling[10]

1. Grade ratio test

Create a time series of the total sales for each merchandise item between June 24 and 30 as follows

$$X^{(0)} = \left(X^{0}(1), X^{0}(2), \ldots X^{(0)}(7)\right) \tag{6}$$

(1) Find the grade ratio $\lambda(k)$

there are

$$\lambda(k) = \frac{X^{(0)}(k-1)}{X^{(0)}(k)} \tag{7}$$

$$\lambda = \left(\lambda(2), \lambda(3) \ldots \lambda(7)\right) \tag{8}$$

(2) Class ratio judgment

Since all $\lambda(k) \in [\ ]$, k=2,3,.... .7,so we can model GM(1,1) with $X^{(0)}$

2. GM(1,1) modeling

(1) Doing a summation of the original series $X^{(0)}$ yields $X^{(1)}$

(2) Construct the data matrix B and the data vector Y, with

$$B = \begin{pmatrix} -\frac{1}{2}\left(x^{(1)}(1) + x^{(1)}(2)\right) & 1 \\ -\frac{1}{2}\left(x^{(1)}(2) + x^{(1)}(3)\right) & 1 \\ & \\ -\frac{1}{2}\left(x^{(1)}(6) + x^{(1)}(7)\right) & 1 \end{pmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \\ x^{(0)}(7) \end{bmatrix} \tag{9}$$

(3) Calculations

$$\hat{u} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (B^T B)^{-1} B^T Y \tag{10}$$

(4) Modeling

$$\frac{dx^{(1)}(t)}{dt} + \hat{a}x^{(1)}(t) = \hat{b} \tag{11}$$

Solving for this gives

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}}\right)e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}} \tag{12}$$

(5) To generate the sequence prediction value $\hat{X}^{(1)}(k+1)$ and the model reduction value $\hat{X}^{(0)}(k+1)$, so that k = 1,2,3,4,5,6, by (1.16) type of time response function can be calculated

$\hat{X}^{(1)}$,which take $\hat{X}^{(1)}(1)=\hat{X}^{(0)}(1)=X^{(0)}(1)=$, by $\hat{X}^{(0)}(k+1)=\hat{X}^{(1)}(k+1)-\hat{X}^{(1)}(k)$, to obtain k = 1,2,3,4,5,6, get

$\hat{X}^{(0)}=(\hat{x}^{(0)}(1),\hat{x}^{(0)}(2)...\hat{x}^{(0)}(7))$

$$\frac{dx^{(1)}(t)}{dt}+\hat{a}x^{(1)}(t)=\hat{b} \tag{13}$$

$$\hat{x}^{(1)}(k+1)=\left(x^{(0)}(1)-\frac{\hat{b}}{\hat{a}}\right)e^{-\hat{a}k}+\frac{\hat{b}}{\hat{a}} \tag{14}$$

Similar to the previous problem, let A represent the sales volume of each vegetable item, B represent the replenishment volume of each item, C represent the cost of each item, and λ be assumed as the wastage rate.

$$Q_t=Z_t+\lambda Z_t \tag{15}$$

Also, establish a functional relationship between the sales volume and replenishment volume of each vegetable item.

$$Z_t=\frac{Q_t}{1-\lambda} \tag{16}$$

Let the linear fitting equation for the sales volume and sales price of each vegetable item be given.

$$Q_t=a-bP_t \tag{17}$$

Let E be the profit that can be obtained

$$f=P_tQ_t-C_tZ_t \tag{18}$$

3. Result

3.1. Solving linear regression models

As shown in Table 1, The slope parameter b and the intercept parameter a were obtained from the data and formulas, which were input into the linear regression model through MATLAB software. This allowed us to obtain the slope parameter in the equation that represents the relationship between the total sales volume of each vegetable category and the average selling price of each week.

Table 1. Summary of slope and intercept parameters in the fitting equations for each vegetable category

Category Name	The slope parameter in the fitted equation	Intercept parameters in the fitted equations
cauliflower	-20.3768	440.0256
capsicum	-9.7011	759.2965
philodendron	-174.0911	2177.5
eggplant	5.2886	76.3711
edible mushroom	-41.6307	943.2337
Aquatic rhizomes	-18.2238	464.2724

As shown in Table 2, The total amount of replenishment and pricing strategy for each vegetable category that maximizes the superstore's gains from July 1-7, 2023 is obtained by inputting the above equations and related data into MATLAB software.

Table 2. Average selling prices and replenishment data for each vegetable category are available from July 1-7, 2023

vegetable category	Sales price of each vegetable category (yuan/kg)	Replenishment (kg)
cauliflower	6.23	3032.6
flowers and leaves	2.77	6843.2
capsicum	2.75	2271.44
eggplant	1.94	1924.9
edible mushroom	4.02	2011.7
Aquatic rhizomes	3.31	293.26

3.2. Solving the gray prediction model

As shown in Table 3, Based on the constructed gray prediction model, the linear regression equations for the total daily sales volume of each individual product in June 2023 and the average price of each vegetable product are obtained by solving MATLAB software. This paper only presents the total daily sales volume of the first 8 individual products in June 2023 and the linear regression equation for the average price of each vegetable product.

Table 3. Linear regression equations for the sales volume and sales price of each vegetable item on July 1, 2023

Item Name	regression equation
White Mushroom(Bag)	Q(t)=0.1739P(t)-0.0011
broccoli	Q(t) = 0.1624P(t)
Spinach (portions)	Q(t)=0.7609P(t)+0.0925
Chinese flowering cabbage	Q(t)=0.4085P(t)+0.0956
Cordyceps flowers (portions)	Q(t)=0.7602P(t)
Kogua (1)	Q(t)=0.1327P(t)+0.0947
Takagua (2)	Q(t)=0.038P(t)-0.0014
Seafood Mushroom(Pack)	No price change, no relevance seen

As shown in Table 4, The fitted equations were input into the MATLAB software for solving, enabling a more intuitive visualization of the total replenishment amount and pricing strategy for each vegetable item that maximizes profitability at the superstore on July 1, 2023.

Table 4. Sales unit prices and replenishment for each vegetable item in 2023

July 1, 2023		
Item Name	Sales unit price (yuan/kg)	Replenishment (kg)
White Mushroom(Bag)	2.8902	0.3875
broccoli	15.3741	2.2746
Spinach (portions)	7.0258	6.0473
Chinese flowering cabbage	4.9558	0.8327
Cordyceps flowers (portions)	3.4449	2.1382
Kogua (1)	16	1.3408
Takagua (2)	17.5122	1.2853
Seafood Mushroom(Pack)	3	6.5601

3.3. Model Extension

In order for superstores to make more accurate replenishment and pricing decisions for vegetable items, this paper also needs to consider the following aspects, which are applicable to replenishment plans and pricing strategies for other categories of items in superstores as well.

1. Considering the challenge of integrating online and offline marketing in superstore operations, separate replenishment and pricing decisions for vegetable commodities are formulated by collecting sales data from both online and offline sources.

2. The overall level of service, the layout of the super environment, and other aspects of planning will affect customers' overall purchasing experience in the supermarket, thus impacting its total revenue. Therefore, a reasonable store layout should be designed for efficient operation and comprehensive business premises planning.

3. The daily sales flow data of the superstore on all kinds of products were analyzed for correlation to determine the relevance of each product, thereby inferring consumers' eating habits. Considering grouping certain products together can enhance customers' purchasing ability.

4. Considering that the appearance of certain vegetables is seasonally distributed and may not be frequent at certain times, the superstore should gather data on the production distribution of different categories of vegetables throughout the year. By combining this information with its own analysis, it can develop replenishment plans and pricing strategies tailored to different seasons.

5. Investigate the market pricing laws of various types of vegetables, analyze customers' consumer psychology, develop a reasonable pricing strategy, and purchase goods that cater to consumers' consumption habits as much as possible in order to enhance their repeat purchasing ability.

The replenishment plan and pricing strategy of various types of commodities in superstores can be further optimized by reviewing relevant literature and collecting pertinent data to enhance the overall operational efficiency of superstores.

4. Conclusion

To address the automatic pricing and replenishment decision-making problem for each vegetable category in a superstore, we analyzed the linear relationship between unit price and sales volume of different commodity categories using a linear regression model based on marketing data from 2020 to 2023. By maximizing revenue, we applied the linear regression model to predict commodity prices and used the gray prediction model to forecast replenishment volume. This paper focuses on studying the total daily replenishment quantity and pricing strategy for each vegetable category from July 1-7, 2023, based on detailed sales flow data from a superstore. Additionally, it provides replenishment quantity and pricing strategies for individual products on July 1, 2023, considering available varieties from June 24-30, 2023. The aim is to maximize revenue while meeting market demand for each vegetable category's commodities. This study realizes specific-time pricing and replenishment decision-making for vegetable commodities in the superstore, providing a theoretical basis for other commodities' pricing and replenishment strategies in order to significantly improve overall operational performance.

References

[1] Zhang X,Yang J,Zhai R, et al. Analysis of Product sales Forecast based on time Series Model[P]. 2021 3rd International Conference on Information Science and Electronic Technology,2021.

[2] Ruihai L ,Jinn-Tsair T . Pricing and lot-sizing decisions for perishable goods when demand depends on selling price, reference price, product freshness, and displayed stocks [J]. European Journal of Operational Research,2018,270(3).

[3] Liu L ,Zhao L ,Ren X . Optimal preservation technology investment and pricing policy for fresh food[J]. Computers & Industrial Engineering,2019,135.

[4] Qu T ,Zhang J ,Chan T F , et al. Demand prediction and price optimization for semi-luxury supermarket segment[J]. Computers & Industrial Engineering,2017,113.

[5] Xu ,Guangshu ,Song , et al. Analysis Research on the Lateral Replenishment Strategy of the Chain Supermarket Fresh Agricultural Products *[J]. Journal of Service Science and Management,2015,8(4).

[6] Tang L. Dynamic Pricing and Inventory Temperature Decisions for Fresh Foods with Variable Demands[J]. Applied Mechanics and Materials,2014,3335(596-596).

[7] Junxuan L,Alejandro T,He W, et al. Dynamic Inventory Allocation for Seasonal Merchandise at Dillard's[J]. INFORMS Journal on Applied Analytics,2021,51(4).

[8] Deniz B,Karaesmen I,Scheller-Wolf A. A comparison of inventory policies for perishable goods[J]. Operations Research Letters,2020,48(6).

[9] Liang D,Sun G,Wu S. Desserts sales forecast based on the new information improved gray model[P]. Proceedings of the 2015 International conference on Applied Science and Engineering Innovation,2015.

[10] Si Shoukui,Sun Xijing. Algorithms and Applications of Mathematical Modeling. Place of Publication. Publisher:National Defense Industry Press, 2022, Start Page Code-End Page Code:450-453