Advances in Computer, Signals and Systems (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2023.070916
ISSN 2371-8838 Vol. 7 Num. 9

A study on vegetable commodity replenishment considering single item quantity limitations—based on gray prediction and linear regression modeling

Doudou Zhang*,#, Xuemei Xiang#, Jinyi Yang#

Bussiness College, Southwest University, Chongqing, 402460, China
**Corresponding author: zhangzhenruo1028@163.com*
#These authors contributed equally.

Keywords: GM(1,1) Model, Linear Programming, Pricing Strategy, Replenishment Strategy

Abstract: Data elements are beginning to play an important role in the selling of vegetable items in supermarkets. In order to get the best pricing and replenishment strategy under the limitation of the number of vegetable items, this paper firstly calculates the profit contribution rate of each item and sorts out the top 33 available items, and prioritizes the items with the top profit contribution rate for replenishment. Finally, a linear programming model is developed to solve the pricing strategy for maximizing the revenue of the superstore under the constraints of display quantity and weight of individual products.

1. Introduction

As a necessity of life, the pricing of vegetables has a significant impact on the protection of residents' lives and the improvement of hypermarket revenue. As a large-scale supply and marketing center radiating a region, the sales strategy of superstores for basic necessities such as vegetables directly affects the quality of livelihood protection. From the perspective of profitability, supermarkets can increase revenue by raising prices, but too high a price will lead to a decrease in sales volume; they can also increase revenue by supplying sufficient quantities, but too much backlog of vegetable goods will lead to losses. Therefore, rational pricing and replenishment decisions are important keys to maintaining a high revenue balance. In order to obtain the optimal commodity pricing for superstores, Yang Shuai (2022) obtained multi-commodity retail chain pricing as well as shelf allocation strategies based on the food decay model combined with the market demand function [1]; Gu Sihong (2023) proposed a multi-stage dynamic pricing model during the freshness decline period based on a linear programming model taking more into account the characteristics of consumable goods such as vegetables [2]. For the replenishment strategy, Bao Jufang (2019) introduced a strong adaptive three-parameter weibull distribution model to study the replenishment strategy of perishable goods [3]; Zhu Hongyu (2020) took into account the impact of shelf display volume on consumer demand in physical sales on the basis of this model to establish a dual-channel supply chain inventory model for fresh agricultural products that takes into account the characteristics of non-instantaneous deterioration and shelf display volume [4]. Throughout the previous research, it is found that for the short-term prediction of market prices, the commonly used methods are econometric analysis represented by regression analysis, mathematical and statistical prediction represented by time series, and intelligent analysis based on BP neural network and gray prediction [5]. These pricing prediction models are less explored for the replenishment prediction of fresh vegetables and other fresh products, while this paper indirectly predicts the price through the prediction of demand, using a small cross-domain prediction based on the unit of days, and in order to be closer to the reality of the situation, the rate of attrition and the limitations of the selling space and weight will be included in the study, and maximize to meet the market demand for various types of vegetable commodities.

2. Establishment and solution of vegetable pricing replenishment model

2.1 Data preparation

The data used in the model comes from the NCAA Mathematical Modeling 2023 C problem. In this paper, we use the pricing and replenishment data of previous vegetable food products of the superstore provided in the mathematical modeling competition to calculate the average contribution rate of each vegetable item from June 24 to 30, 2023, and select the top 33 items based on the contribution rate.

2.2 Establishment of GM(1,1) model for daily replenishment of optional items

2.2.1 Gray prediction model establishment

Predicting the replenishment quantity of an item on a future day based on the daily sales quantity of the item in the recent week belongs to the future prediction of small samples. Gray prediction is a model based on a small sample of data for prediction, it is dealing with less eigenvalue data, do not need the sample space of the data is large enough, it can solve the problem of less historical data, the completeness of the sequence as well as the low reliability of the original data can be generated to get the regularity of the original data to get the regularity of the stronger generation of the sequence[6].

Therefore, the gray prediction model is chosen to predict the daily replenishment of a single product on July 1st.

GM(1,1) gray prediction model is a single variable prediction of the first-order differential equation model, its discrete time response function is approximately exponential law, the establishment of GM(1,1) model [7]is:

Let $\mathrm{X}^{(0)} = \mathrm{X}^{(0)}(1), \mathrm{X}^{(0)}(2), \cdots, | \mathrm{X}^{(0)}(\mathrm{n})$ be the original nonnegative time series, $\mathrm{X}^{(0)}(\mathrm{t})$ be the cumulative generating series,i.e. $\mathrm{X}^{(1)}(\mathrm{t}) = \sum \mathrm{X}^{(0)}(\mathrm{m}), \mathrm{t} = 1,2. \ldots \mathrm{n}$

The whitened differential equation for the GM(1,1) model is:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \tag{1}$$

In the above equation, a is the parameter to be identified, also known as the development coefficient; u is the endogenous variable to be identified, also known as the gray role. Let the vector $\hat{a} = \begin{pmatrix} a \\ u \end{pmatrix}$, and find $\hat{a} = (B^T B)^{-1} B^T y$ by the least squares method, where

$$B=\begin{vmatrix} -\frac{1}{2}\left(X^{(1)}(1)+X^{(1)}(2)\right) & 1 \\ -\frac{1}{2}\left(X^{(1)}(2)+X^{(1)}(3)\right) & 1 \\ \dots\dots & \dots \\ -\frac{1}{2}\left(X^{(1)}(n-1)+X^{(1)}(n)\right) & 1 \end{vmatrix} \tag{2}$$

$$y=\begin{vmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \dots \\ X^{(0)}(n) \end{vmatrix} \tag{3}$$

The discrete time response function for gray prediction is obtained as:

$$\boldsymbol{X^{(1)}(t+1)=\left(X^{(0)}(1)-\frac{u}{a}\right)e^{-xt}+\frac{u}{a}} \tag{4}$$

$X^{(1)}(t+1)$ is the resulting cumulative predicted value, and reducing the predicted value is:

$$X^{(0)}(t+1)=X^{(1)}(t+1)-X^{(1)}(t),(t=1,2,3,\dots,n) \tag{5}$$

Before building the gray prediction model GM(1,1), the time series is subjected to a rank-ratio test. If the level ratio test, it means that the series is suitable for the construction of gray model, if not through the level ratio test, the series will be "shifted to transform", so that the new series to meet the level ratio test[8]. Gray prediction model to be tested to determine whether it is reasonable, only through the test model can be used for prediction, the system is mainly through the a posteriori difference than the C value of the gray prediction model to test.

2.2.2 Model solving

Gray prediction was performed using SPSSPRO and the results of the grade ratio test were obtained as shown in Table 1 below:

Table 1: Grade ratio test result table

Index term	Original value	Ratio value	Translated sequence value	Translated level ratio value
2023-06-24	9.387	-	26.387	-
2023-06-25	10.681	0.879	27.681	0.953
2023-06-26	8.083	1.321	25.083	1.104
2023-06-27	14.272	0.566	31.272	0.802
2023-06-28	13.401	1.065	30.401	1.029
2023-06-29	15.18	0.883	32.18	0.945
2023-06-30	16.9	0.898	33.9	0.949

The table shows the sequence values and the grade ratio values. If all the rank ratio values lie in the interval$\left(e^{-\frac{2}{n+1}},e^{\frac{2}{n+1}}\right)$, the data are suitable for modeling. If it does not pass the rank-ratio test, then the sequence is transformed into a "level shift", so that the sequence after the level shift meets the rank-ratio test. From the analysis of the above table, we can get that all the level ratio values of the shifted series are located in the interval(0.779,1.284), which indicates that the shifted series is suitable for constructing the gray prediction model.

Then the gray model is constructed to calculate the development coefficient, gray role quantity, and the value of the posterior difference ratio shown in Table 2.

Table 2: Summary of results

Development coefficient a	Gray role quantity b	posteriori difference ratio C value
-0.049	24.571	0.21

The gray prediction model can be constructed from the development coefficient and gray role quantity. The development coefficient indicates the development pattern and trend of the series, and the gray role quantity reflects the change relationship of the series. The a posteriori difference ratio can verify the accuracy of gray prediction, the smaller the a posteriori difference ratio, the higher the accuracy of gray prediction[9]. Generally, if the C value of the a posteriori difference ratio is less than 0.35, then the model accuracy is high, the C value is less than 0.5, which indicates that the model accuracy is qualified, the C value is less than 0.65, which indicates that the model accuracy is basically qualified, and if the C value is greater than 0.65, it indicates that the model accuracy is unqualified. Analyzing the table, we get that the a posteriori difference ratio value is 0.21 and the model precision is high[10].

The results of the model fit are shown in Table 3:

Table 3: Model fitting results

Index term	Original value	Predicted value	Residual	Relative error (%)
2023-06-24	9.387	9.387	0	0
2023-06-25	10.681	9.512	1.169	10.948
2023-06-26	8.083	10.846	-2.763	34.177
2023-06-27	14.272	12.247	2.025	14.192
2023-06-28	13.401	13.718	-0.317	2.366
2023-06-29	15.18	15.264	-0.084	0.55
2023-06-30	16.9	16.887	0.013	0.078

The average relative error of the model is 8.902%, which means that the model fits well. Finally, the daily replenishment volume of broccoli on July 1 was obtained as 20.489 kg. Similarly, the sales volume of other optional single items on July 1 is obtained, and according to the replenishment volume = sales volume / (1 - wastage rate), the replenishment volume of each single item on July 1 is obtained as shown in Table 4.

Table 4: Individual Product July 1 Sales and Replenishment

Category	July 1 Sales Volume	July 1 Restocking Volume
Broccoli	18.592	20.489
Zijiang Green Stem Scattered Flowers	13.971	15.426
Spinach (portion)	16.016	17.684
Sweet Potato Tips	6.192	6.761
Wood Ear Vegetable	5.591	6.052
Milk Cabbage	10.744	12.742
Shanghai green	8.236	9.625
Baby Cabbage	12.584	12.904
Amaranth	6.089	7.473
Small bok choy(1)	5.365	5.983
Yunnan Lettuce	0.671	0.000
Yunnan Lettuce (portion)	38.32	42.310
Yunnan Oil Wheat Lettuce(portion)	27.921	30.828
Bamboo Leaf Lettuce	11.674	13.515
Red Pepper(2)	2.031	0.000
Ginger, Garlic & Millet Pepper Combo Pack (small portion)	5.182	5.722
Screw Pepper	8.749	9.741
Screw Pepper (portion)	13.377	14.770
Colorful Peppers(2)	0.167	0.000
Wuhu Green Pepper(1)	16.564	17.565
Small Millet Pepper (portion)	28.659	31.643
Small Wrinkled Pepper(portion)	9.276	10.242
Green Eggplant(1)	0.192	0.000
Long Line Eggplant	7.968	8.559
Purple Eggplant(2)	6.175	6.574
Seafood Mushroom(Pack)	7.274	7.274
Enoki Mushroom(box)	22.729	22.832
Agaricus bisporus(box)	9.279	9.298
Xixia Mushroom(1)	5.884	6.596
Gourd(1)	4.044	5.716
Honghu Lotus Root	4.53	5.964
Lotus Root(1)	7.535	7.977
Rhododendron	3.87	4.281

2.3 Regression Modeling of Sales Volume and Price

2.3.1 Regression modeling

Analyzing the sales flow from 2020 to 2023, it was found that the correlation between the sales volume of each vegetable item is weak, and vegetables belong to the necessities of life, so it is possible to price each item separately, and when the revenue of each individual item is maximized, the total revenue is also maximized.

As an example, the demand price curve of millet peppers is analyzed. Plot the scatter plot of average sales price and sales volume of millet peppers and fit a linear relationship between the two as in Figure 1.

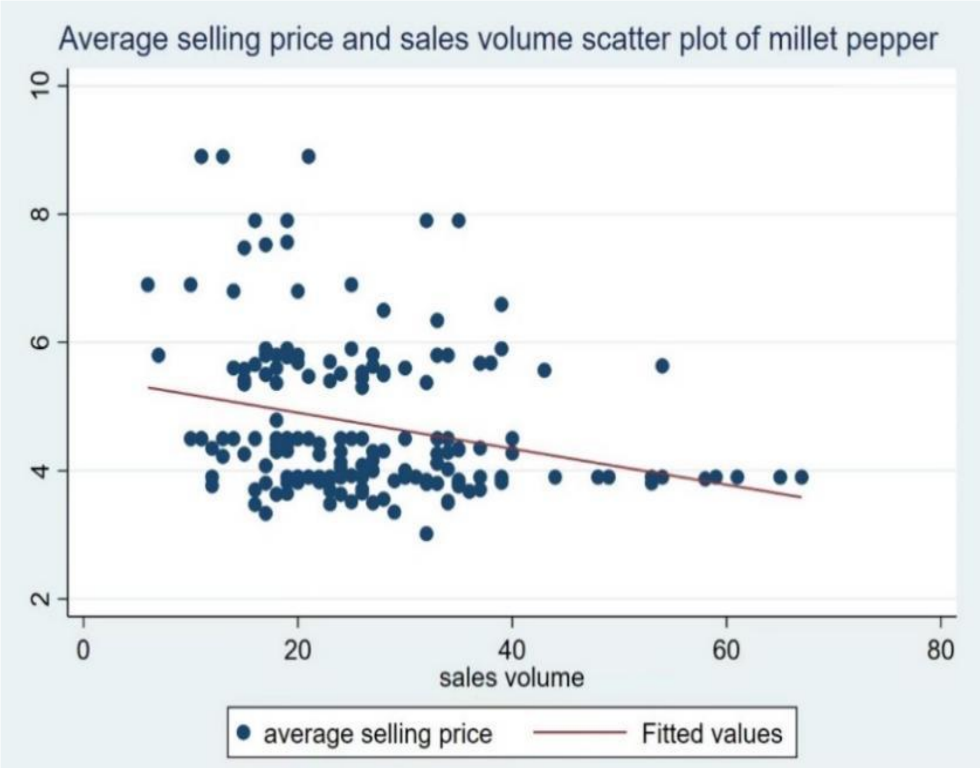

Figure 1: Average selling price and sales scatter plot of millet pepper

The linear back-analytic model is established as:

$$\begin{cases} y = \beta_0 + \beta_1 x + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \tag{6}$$

where β0, β1, σ2 are all unknown parameters independent of x, where β0, β1, are called regression coefficients. Now we get n independent observations (yi, xi),i=1,...,n, from (1) we get

$$\begin{cases} y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2), i = 1, \dots, n \end{cases} \tag{7}$$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \mathrm{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{8}$$

$$\varepsilon = [\varepsilon_0 \cdots \varepsilon_n]^T, \beta = [\beta_0 \ \beta_1]^T \tag{9}$$

$$y = \beta x + \varepsilon \tag{10}$$

The regression equation is obtained according to equation (7)(8)(9):

$$\mathrm{y} = 37.88 - 2.44\mathrm{x} \tag{11}$$

2.3.2 Regression coefficient hypothesis testing

When H0 above is rejected, βj is not all zero, but it is not excluded that several of them are equal to zero. So further m+1 tests are made as follows (j=0,1,…,m):

$$H_0^{(j)}: \beta_j = 0 \tag{12}$$

$$\hat{\beta} \sim \mathrm{N}(\beta, \sigma^2(X^T X)^{-1}) \tag{13}$$

$$\hat{\beta}_j \sim N(\beta_j, \sigma^2 c_{jj}) \tag{14}$$

C_{jj} is the(j,j)element in$(X^T X)^{-1}$ with s2 instead of σ2,when $H_0^{(j)}$ holds

$$t_j = \frac{\hat{\beta}_j/\sqrt{c_{jj}}}{\sqrt{Q/(n-m-1)}} \sim t(n-m-1) \tag{15}$$

For a given α, accept $H_0^{(j)}$ if $|t_j| < t_{\frac{\alpha}{2}}(n-m-1)$, otherwise reject.

The t-value for the regression coefficient test was calculated for y=37.88-2.44x and obtained as shown in Table 5.

Table 5: Regression coefficient test table

t_0	t_1
11.39	-3.58

The mean value of regression coefficient is greater than $t_{\frac{\alpha}{2}}$, then the original hypothesis is rejected and the regression coefficient passes the test.

2.4 Linear Programming Pricing Modeling and Solution

2.4.1 Linear Programming Pricing Model Establishment

Take the average wholesale price of each optional single product from June 24 to 30, 2023 as the wholesale price of single product on July 1, and record the attrition rate of each single product. Assuming Xi is the price of individual item i,Ci is the cost of individual item i,Qi,is the replenishment quantity of individual item i, Si is the attrition rate of individual item i,and n is the number of selected individual items, the total revenue is:

$$W = \sum_{i=1}^{33}(X_i - C_i)Q_i(1-S_i) - Q_i S_i C_i \ (i = 1,2,\dots,n) \tag{16}$$

The constraints are as follows:

(1) The total number of saleable individual items is controlled to be 27-33, i.e., 27≤n≤33.

(2) The order quantity of each single item satisfies the minimum display quantity of 2.5 kg, i.e., Q_i≥2.5.

(3) To try to meet the premise of the market demand for various types of vegetable goods, then the selection of single products should be divided into 6 types of vegetables.

Combining the above analysis, the mathematical model is established as follows:

$$\max\ W = \sum_{i=1}^{33}(X_i - C_i)Q_i(1-S_i) - Q_i S_i C_i \tag{17}$$

2.4.2 Model solving

The model was solved using Lingo software to determine what the selling price of the millet peppers could be when the revenue could be maximized. The result was obtained that when the price of millet peppers is 8.74, the revenue of millet peppers on that day can reach the maximum value of 105.83 ￥.

Similarly, under the constraints, the July 1 single-item replenishment and pricing are obtained as shown in Table 6.

When the superstore's replenishment and pricin on July 1 satisfy the above table, it can make the maximum revenue of 1,848.90￥yuan from vegetable items on that day.

Table 6: Pricing for Available Individual Items

Catelogy	July 1 Restocking Volume	Pricing
Broccoli	20.489	18.09
Zijiang green stalks loose flowers	15.426	16.61
Spinach (portion)	17.684	13.6
Sweet Potato Tips	6.761	8.82
Mullein	6.052	6.67
Milk Cabbage	12.742	10.18
Shanghai green	9.625	13.7
Baby Cabbage	12.904	7.43
Amaranth	7.473	11.4
Baby bok choy(1)	5.983	12.87
Yunnan Lettuce(portion)	42.310	9.5
Yunnan oilseed rape(portion)	30.828	5.47
Bamboo Leaf Lettuce	13.515	11.74
Ginger, Garlic & Millet Pepper Combo (small portion)	5.722	3.31
Screw Pepper	9.741	10.42
Screw Pepper(portion)	14.770	8.23
Wuhu Green Pepper(1)	17.565	8.9
Millet Pepper(portion)	31.643	8.74
Small Wrinkled Pepper(portion)	10.242	5.98
Longline Eggplant	8.559	10.44
Purple Eggplant(2)	6.574	12.13
Seafood Mushroom(Pack)	7.274	4.85
Golden Needle Mushroom(box)	22.832	2.62
Agaricus bisporus(box)	9.298	5.83
Xixia Mushroom(1)	6.596	21.83
Gourd(1)	5.716	15.5
Honghu Lotus Root	5.964	28.29
Net Lotus Root(1)	7.977	28.83

3. Conclusion

In this paper, for the problem of pricing and replenishment strategy under the quantity restriction of vegetable items in supermarkets, a method based on GM(1,1) model and linear programming is proposed based on gray prediction and linear regression modeling. By calculating the profit contribution rate, the top 33 sellable items are organized, and the items with the highest profit contribution rate are prioritized for replenishment. The established GM(1,1) model and attrition rate are utilized to predict the sales volume of individual items, and a replenishment plan is developed based on the demand. Finally, the pricing strategy for revenue maximization is solved under the constraints of single-item display volume and weight. The experimental results show that under the total number of saleable single items in a single day is 27-33 and the order quantity of each single item meets the minimum display quantity of 2.5 kg, this paper derives the maximum revenue of the superstore to be 1890.90¥. The method proposed in this paper can significantly improve the revenue and efficiency of supermarkets, and has practical application value and promotion significance.

References

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

[2] Yang Shuai, Huang Xiangmeng, Wang Junbin. Research on joint optimization strategy of shelf allocation and pricing for fresh food[J]. Supply Chain Management, 2022, 3(08): 49-59.

[3] Jufang Bao, Chengqian Xu. Research on replenishment strategy of perishables based on Weibull distribution[J]. Journal of Nanyang Institute of Technology, 2019, 11(02): 13-17+30.

[4] Hongyu Zhu. Research on inventory control strategy of dual-channel supply chain for fresh agricultural products [D]. Chongqing University of Posts and Telecommunications, 2021.

[5] Xiong Wei, Qi Qunquan, Gao Yu et al. Research on short-term prediction of agricultural market price based on combination modeling--Taking red Fuji apples, bananas and oranges as examples[J]. Agricultural Technology Economics, 2015(06): 57-65.

[6] Hu YC. A genetic-algorithm-based remnant grey prediction model for energy demand forecasting [J].PLoS ONE. 2017, 12(10): e0185478.

[7] Suo Di. Gray prediction expansion model based on time series feature analysis and its application in energy sector [D]. Jiangnan University, 2023.

[8] Liu Y, Wang Y, Yu FY et al. Forecasting analysis of power demand in Jilin Province in the 14th Five-Year Plan based on GM(1, 1) method[J]. Green Technology, 2022, 24(18): 232-236.

[9] Cen Wang. Evaluation of regional development of healthy China with GM(1, 1) gray prediction analysis[D]. Wuhan University, 2023.

[10] Kaiwusha Tayir, Lai Hua, Gulimige Erken et al.Spatio-temporal evolution and prediction of carbon emission in Urumqi region supported by FLUS and gray prediction model[J]. Journal of Soil and Water Conservation, 2023, 37(04): 214-226.