Optimization Research on Category Sales Strategy Based on Cost Profit Margin and Time Series Forecasting Models

Haoyu Wang [#], Xiaoyuan Wan [#], Yujie Ren [*], Wanying Lu

Academy of Reading, Nanjing University of Information Science & Technology, Nanjing, China, 210044

* Corresponding Author Email: 202183100045@nuist.edu.cn

[#]These authors contributed equally.

Abstract. Fresh food supermarkets often need to make multi-factor considerations when managing and selling vegetables, due to deterioration and out-of-stock problems of vegetables, stores also need to develop daily replenishment and pricing plans based on historical sales data and demand for each vegetable. In order to develop a reasonable replenishment and pricing strategy program, this paper first explores the relationship between total sales and cost-plus pricing on the basis of vegetable categories, considering the cost profit margin under the effect of the attrition rate, and portraying cost-plus pricing through this variable. Then, by selecting the time period that is relatively smooth and regular in data distribution, the sales volume and cost profit margin are fitted by MATLAB to obtain the functional relationship, and the inverse relationship and direct relationship are all found to be the utility curves, which are “small profit but sell more” and “more profit and sell more”. These are all product characteristics. Based on the respective functional relationships between sales volume and cost profit margin, this paper applies the time series model to predict the value of the attrition rate and wholesale price in the coming week as well as statistics and gives the daily replenishment strategy and the pricing strategy of the six categories in the coming week.

Keywords: Introduction

In fresh food supermarkets, the freshness period of vegetable commodities is not very long, and their quality deteriorates with the increase of sales time. Most varieties, if not sold on the same day, the freshness of the vegetables will be greatly reduced, and they cannot be sold again on the next day. Therefore, the supermarket will often be based on the historical sales data and demand for each commodity daily replenishment. Supermarkets sell a wide range of vegetables, different origins, so merchants have to be unsure of the specific single product and the purchase price of the case, the day of the replenishment of various types of vegetables to make decisions. Pricing of vegetables is generally based on the "cost-plus pricing" method, and supermarkets usually offer discounts for transportation damage and poor-quality goods. In addition, reliable market demand analysis is important for replenishment and pricing decisions. From the demand side, there is often a correlation between the sales volume of vegetables and the time of year; from the supply side, the supply of vegetables is more abundant from April to October, and the limitations of the supermarket's sales space make a reasonable sales mix extremely important.

Prior studies are usually divided into two camps. On the one hand, scholars have empirically analyzed the special properties of agricultural products such as fresh fruits and vegetables (including seasonality, climatic influences, and maturity) and their impact on inventory management strategies based on traditional theories such as EOQ and safety stock. On the other hand, the researchers used the deteriorating inventory model to analyze the replenishment strategy and carried out model construction and optimization for the replenishment decision of fresh produce under different storage capacity and life cycle constraints. Among them, Wang and Liu investigated the replenishment strategy of fresh products under online retailing by means of dynamic pricing and inventory control models, with a special focus on the potential impact of pre-ordering behavior on retailers' profits [1]. Lv Yang proposed double loss to formulate reasonable operational decisions when fresh food e-commerce companies face uncertain demand and external environmental influences [2]. The topic of this paper is quite related to the research of Shuyun Wang et al. in examining the cold chain integrated inventory decision making, which analyzed the impact of quality and quantity loss on inventory

management of fresh products, taking into account the time factor and preservation efforts [3]. In addition, when Xu Hao improves the supermarket business strategy based on the results of forecasting, he establishes an inventory control model for deteriorating products to achieve the purpose of controlling the inventory of fresh products in supermarkets, taking into account the fact that the inventory of fresh products occupies a large amount of money, the sales cycle is continuous and some stock-outs are allowed [4], which is complemented by this paper's use of time-series modelling for forecasting the future demand and pricing. Notably, this paper also highlights the need to utilize historical sales data and market demand for decision making, which coincides with Feng's study [5], and considers seasonal demand, shelf life and price elasticity of perishable products, thus helping retailers to adjust their seasonal fruit and vegetable demand forecasting, inventory control and ordering strategies. By summarizing these prior studies and its own empirical analysis, this study provides a new perspective on inventory replenishment and pricing strategies for fresh produce, emphasizing the importance of close integration of market dynamics and product characteristics, while demonstrating the value of accurate forecasting and careful planning. It provides a scientific basis for fresh produce supermarket management to deal with the variable factors in vegetable sales, and helps supermarkets optimize their daily replenishment and pricing plans.

In this paper, total statistics of six major vegetable categories, namely, "Cauliflower", "Mosaic", "Capsicum", "Solanum", "Mushroom", and "Aquatic Roots and Tubers", were obtained and analyzed for the period of 2021-2023 [6]. In order to develop a rational replenishment and pricing strategy, this paper explores the relationship between total sales and cost-plus pricing in the context of the vegetable category. Cost profit margins are used to characterize cost-plus pricing. The problem is then transformed to investigate the relationship between total sales and cost profit margin. The relationship between total sales and cost profit margin is fitted by category, resulting in a relationship between total sales and cost profit margin that reflects the relationship between total sales and cost-plus pricing through the intermediate bridge of cost profit margin. Next, this paper addresses the issues of total daily replenishment and pricing strategy at maximum revenue. The total replenishment is used to portray the incoming quantity and the pricing strategy is used to portray the unit price of sales, which is modeled using a time series model to find the data in the coming week. The daily replenishment quantity and pricing strategy can be derived by solving the forecast results.

1. Model building and solving

1.1. Fitting study for the relationship between total sales and cost profit margins

The data of six main categories of vegetables, namely, "Cauliflower", "Mosaic", "Capsicum", "Solanum", "Mushroom", and "Aquatic Roots and Tubers", is mainly collected and analyzed [6]. For the cost-plus pricing of the vegetables, this paper uses the cost-profit margin calculation to measure and characterize the cost-plus pricing, namely profit divided by cost. Hence, set sales volume as a, unit price of sales as b, attrition rate as z, wholesale prices as c, cost profit margin as q, total replenishment as d and total profit as W. Set an assumption as "Vegetables not sold on one day at the supermarket will not be sold the next day.". Supermarkets are unable to sell vegetables that have deteriorated overnight, so this assumption is reasonable. Based on this assumption, it is considered that any vegetables not sold on the same day are considered as wastage. Define total replenishment as d and total replenishment is shown in Equation 1:

$$d = \frac{a}{1-z} \tag{1}$$

Based on the definition of the total replenishment as d, define the total profit as W. The total profit is shown in Equation 2:

$$W = a * b - c * d \tag{2}$$

In the Equation 2, set the sales volume as a, the unit price of sales as b, the wholesale prices as c, the total replenishment as d. The cost profit margin q is defined and shown in the Equation 3:

$$q = \frac{W}{c \bullet d} \tag{3}$$

Joint Equation 1 and Equation 3 and then Equation 4 **c**an be gotten and shown as follows:

$$\begin{cases} d = \frac{a}{1-z} \\ q = \frac{W}{c \bullet d} \end{cases} \tag{4}$$

In the Equation 4, a, b, c are all known, the total profit W can also be expressed in terms of known metrics from existing data, so in this paper the cost profit margin can be described using a known variable name to get an expression about the cost profit margin. Fitting is a method of getting as close as possible to the characteristics of the original data set by inscribing curves under a defined error function [7]. To get the characteristics of six main vegetable categories, based on this point, the sales volume that is, the total sales and cost profit margins into MATLAB for fitting, and according to the question by category were fitted. After importing all the data to be fitted, the visualization results of the characteristics of four vegetable categories based on fitting are all as shown in Figure 1 - 4:

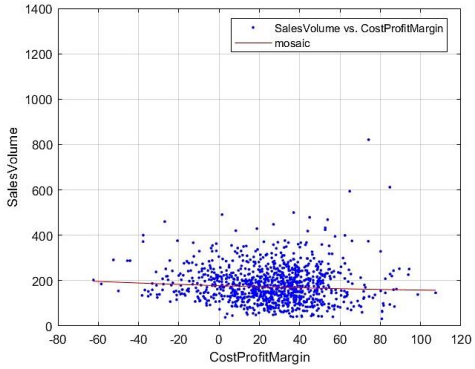


Figure 1: Fitted curves for sales and cost profit margins for the "Mosaic."

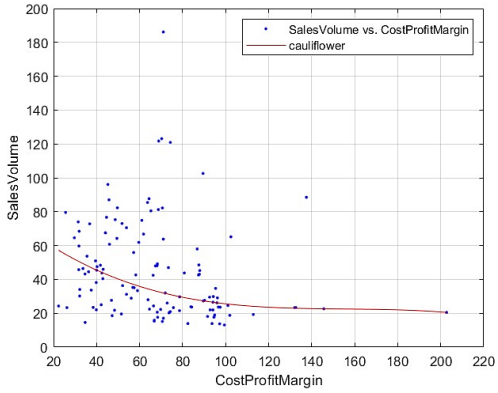


Figure 2: Fitted curves for sales and cost profit margin "Cauliflower."

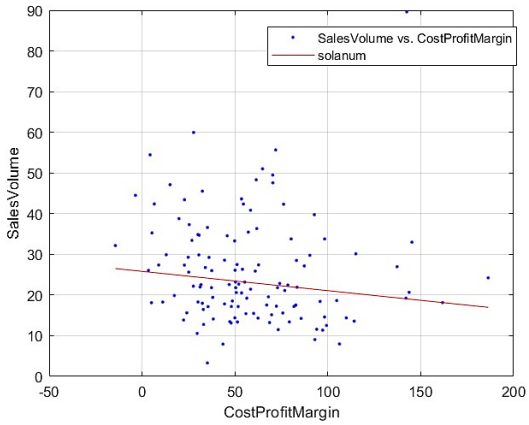


Figure 3: Fitted curves for sales and cost profit margins for the "Solanum."

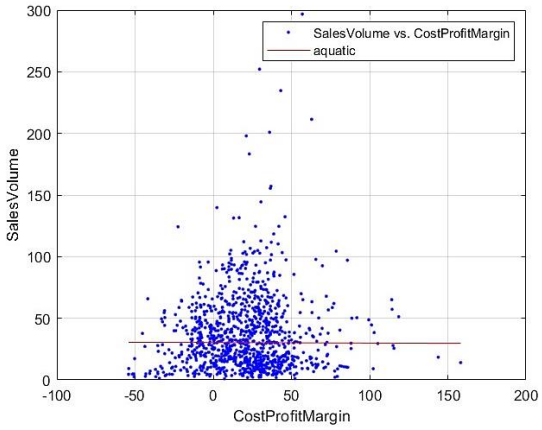


Figure 4: Fitted curves for sales and cost profit margins for the "Aquatic Root and Tuber."

Observing Figure 1 and Figure 4, it is found that the fitted curve shows a weak inverse proportional relationship, indicating that when the cost profit margin is increasing, the sales volume decreases, which is in line with the law of "the thinner the profit, the more sales". Taking "Cauliflower" and “Solanum” as an example, the curve is shown in Figure 2 and Figure 3, which shows that the sales volume and the cost profit margin are obviously inversely proportional to each other, which is also in line with the law of "the thinner the profit, the more sales".

The visualization results of the characteristics of remaining two main vegetable categories are obtained by the same method and are shown as Figure 5 - 6:

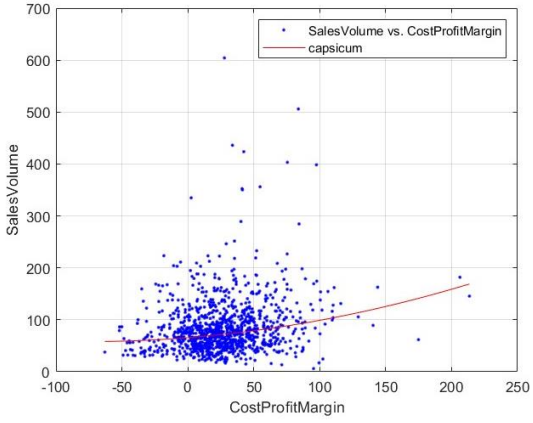


Figure 5: Fitted curves for sales and cost profit margins for the "Capsicum."

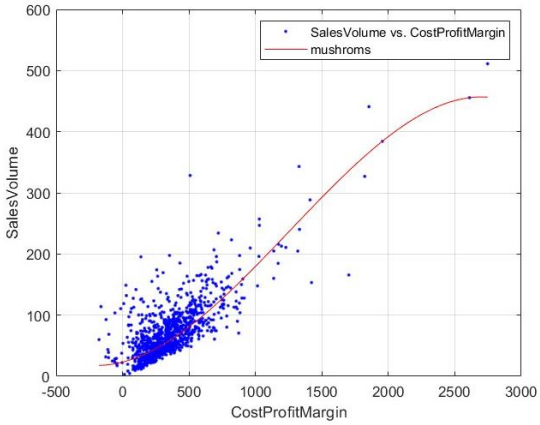


Figure 6: Fitted curves for sales and cost profit margins for the "Mushroom."

From the two visualization results, it is found that the fitted curve shows a clear direct proportional relationship, indicating that when the cost profit margin is increasing, the sales volume increases, which is in line with the law of "the more the profit, the more sales".

These are all the results of characteristics of six main vegetable categories, which are gotten through fitting. These characteristics are transformed into the form of the function and is used to making following strategies.

1.2. Time sequence forecast on obtaining replenishment and pricing strategy

The total daily replenishment and pricing strategy for the coming week is studied and this is aimed at maximizing the benefits to the supermarket. Based on the assumption that overnight vegetables will no longer be sold in the shop, the total daily replenishment volume is set as the daily purchase volume, and the pricing strategy is set as the established selling price per unit. In the prediction process, regarding the replenishment strategy, what is done is to study and predict the overall wastage rate of the vegetable category and carry out a series of mathematical calculations to achieve the prediction results of the replenishment quantity; regarding the pricing strategy, what is done is to study and predict the overall wholesale price of the vegetable category and carry out a series of mathematical calculations to achieve the prediction results of the pricing.

In SPSS, time series plots of wastage rates and wholesale prices of the six food categories are plotted against date, and the main trends of the six categories are observed. In this case, "Aquatic Roots and Tubers" is taken as an example, and take the data from 2021 to 2023, and plot the time series graphs of wastage rate and wholesale price change with date, and the results are shown in Figure 5 - 6 below:

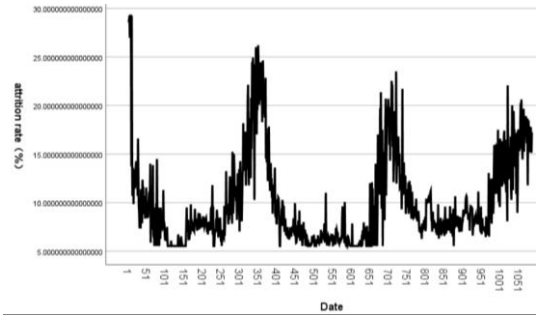


Figure 7: Time-series plot of "Aquatic Roots and Tubers" depletion rates against date



Figure 8: Time-series plot of changes in wholesale prices of "Aquatic Roots and Tubers" as a function of date

From the time series plot, "Aquatic Roots and Tubers" is not seasonal, and the data show that the wastage rate and wholesale price of this category are stable, and there is no trend. At the same time, the data also show the stochasticity. Therefore, it can be summarized that "Aquatic Roots and Tubers" has the characteristics of "non-seasonality" " stochasticity " and "smoothness". Similarly, the characteristics of the six vegetable categories were analyzed. Li Na stated that ARIMA is often used for forecasting time series data for dynamic stochastic processes [8]. Consequently, ARIMA is the main forecast model that is chosen. According to the characteristics of each of the six vegetable categories, the suitable time series forecasting models for the six categories were initially obtained, and the results are shown in Table 1:

Table 1: Time series forecasting models corresponding to the selection of the six categories.

	Cauliflower	Mosaic	Capsicum	Aquatic	Mushroom	Solanum
AR	ARIMA(0,0,0)	ARIMA(0,1,3)	ARIMA(0,1,3)	ARIMA(0,1,1)	ARIMA(0,1,2)	ARIMA(3,1,9)
WP	ARIMA(2,1,8)	ARIMA(0,1,1)	*Simple*	ARIMA(0,1,15)	ARIMA(0,1,1)	*Simple*

Note: “AR” is an abbreviation for attrition rate; “WP” is an abbreviation for wholesale price.

However, the method of finding a suitable time series model based on time series graphs for data trending is subjective [9]. In order to further verify whether the selected model can make a better prediction of the data in the coming week, based on Guan heshan’s thought, which is that the Box-Pierce Q test is an objective analytical method that can give a unique result for the smoothness test of time series data [10], the residual series of the predicted data are subjected to Box-Pierce Q-test. Set the original hypothesis as $H0$: the residual series is the white noise series, set the alternative hypothesis as $H1$: the residual series is not the white noise series, with 95% confidence level, calculate the p-value of the test, the results are as follows in Table 2:

Table 2: Box-Pierce white noise test for time series forecasting models.

	Cauliflower	Mosaic	Capsicum	Aquatic	Mushroom	Solanum
AR’s p-value	0.131	0.302	0.263	0.172	0.117	0.476
WP’s p-value	0.304	0.821	0.695	0.328	0.674	0.534

Note: “AR” is an abbreviation for attrition rate; “WP” is an abbreviation for wholesale price.

From the results, the p-values of the Box-Pierce Q-tests of the corresponding time series models of the six categories are all greater than 0.05, and it is considered that none of them can reject the original hypothesis, and the residuals are considered to be white noise sequences, therefore, the models corresponding to each category can identify the corresponding data well, and the models corresponding to the six categories can be used to predict the attrition rate and the wholesale price of the category. Taking "Aquatic Roots and Tubers" as an example, the visualization results are shown in Figure 7 - 8 below:

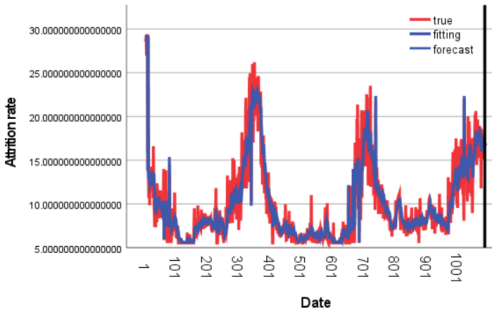


Figure 9: Attrition rate prediction chart

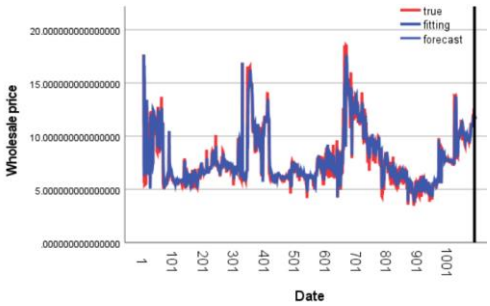


Figure 10: Wholesale price prediction chart

Set the cost profit margin as x. Based on the characteristics of six main vegetable categories obtained from the fitting of sales volume and cost profit margin, therefore let the sales volume fitted function be $f(x)$. At this point for the unit price of sales b, Equation 5 is shown below:

$$x = \frac{f(x)*b}{c*d} - 1 \tag{5}$$

The simplification leads to Equation 6:

$$b = \frac{(1+x)c}{1-z} \tag{6}$$

Then, the total replenishment d is shown in Equation 7 below:

$$d = \frac{f(x)}{1-z} \tag{7}$$

From the above, the total profit W is "sales minus cost" and get the Equation 8 shown below:

$$W = a*b - c*d = xf(x)\frac{c}{1-z} \tag{8}$$

2. Results

2.1. The forecast of replenishment based on time sequence model

Based on the already established time series model and related mathematical equations, the total daily replenishment and the unit sales price for the coming week are derived using MATLAB and collated in Table 3 below:

Table 3: Replenishment Forecast Table for July 1st to July 7th.

	Cauliflower	Mosaic	Capsicum	Solanum	Mushrooms	Aquatic
7.1	79.9866	200.4123	72.0828	27.4946	24.4621	36.4117
7.2	79.9866	200.3830	72.1056	27.4982	24.4621	36.4183
7.3	79.9866	200.3512	72.1192	27.5084	24.4621	36.4248
7.4	79.9866	200.3512	72.1192	27.5094	24.4621	36.4314
7.5	79.9866	200.3512	72.1192	27.5103	24.4621	36.4380
7.6	79.9866	200.3512	72.1192	27.5006	24.4621	36.4446
7.7	79.9866	200.3512	72.1192	27.5044	24.4621	36.4512

Of these, the total number of replenishments by category, for example, as of 1 July, is shown in Figures 9 below:

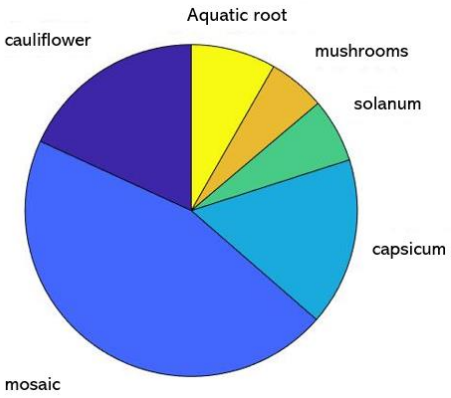


Figure 11: Pie chart of total replenishment of 1 July

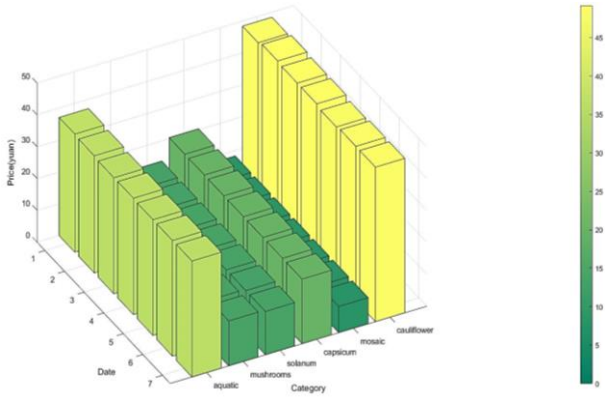


Figure 12: Seven-day pricing strategy bar chart of all the categories

In the pie chart, "Mosaic" accounted for the largest share of the total replenishment, while "Mushroom" accounted for the smallest proportion.

2.2. The forecast of pricing based on time sequence model.

Statistics on pricing strategies were also carried out, as shown in Table 4:

Table 4: Pricing Strategy Forecast Table for July 1st to July 7th

Date	Cauliflower	Mosaic	Capsicum	Solanum	Mushrooms	Aquatic
7.1	48.3830	8.2903	20.3176	13.6387	14.9168	37.0948
7.2	49.1891	8.2808	20.3240	13.6405	14.1379	36.8226
7.3	48.7539	8.2715	20.3279	13.6455	14.1379	36.4574
7.4	48.7885	8.2638	20.3279	13.6460	14.1379	36.5594
7.5	48.0739	8.2565	20.3279	13.6465	14.1379	36.3583
7.6	48.5139	8.2494	20.3279	13.6416	14.1379	36.3588
7.7	48.5105	8.2427	20.3279	13.6435	14.1379	36.4469

The result of visualizing the pricing strategy is shown in Figure 10 above. From the figure, what can be known is that “Cauliflower” has the highest pricing strategy and “Aquatic Root and Tubers” has the second highest pricing strategy, while “Mosaic” has the lowest pricing strategy.

3. Conclusions

This paper examines the data for six common vegetables, namely "Cauliflower", "Mosaic", "Capsicum", "Solanum", "Mushrooms" and "Aquatic Roots and Tuber", for the period 2021-2023. "Mushrooms" and "Aquatic Roots and Tubers" for the period 2021-2023. By portraying the cost profit margins and using MATLAB to implement the fitting technique, the inverse proportional relationship between the cost profit margins and the sales volume was found, presenting the economic law of "the

thinner the profit, the more the sales", and basically, all the categories of the vegetables present this economic law between the cost profit margins and the sales volume.

Then, based on the laws obtained from the fitting, this paper uses the time series plot to portray and summaries the characteristics of "whether there is seasonality" and "whether it is smooth" of the data of the six food categories, and based on these characteristics, combined with the expert modeler in SPSS, initially found a suitable time series model for each data. Based on these characteristics, combined with the expert modeler in SPSS, the time series model suitable for each data was found. Then, the residuals were subjected to Box-Pierce Q-test to check whether the residual series of the predicted data corresponding to the six categories are white noise series, so as to verify whether the selected time series model can well identify the data among the categories. The results show that the selected models are able to identify the data between the categories well enough to complete the prediction task. Based on this, the replenishment strategy and pricing strategy for the coming week is completed.

References

[1] Wang X, Liu S, Yang T. Dynamic pricing and inventory control of online retail of fresh agricultural products with strategic consumers[J]. Economic Research-EkonomskaIstraživanja,2023,36(1):1-20.

[2] Lv Yang. Research on fresh food e-commerce supply chain decision-making considering double loss[D]. Harbin University of Commerce,2022.

[3] WANG Shu-yun, JIANG Ying-mei, MOU Jin-jin.Inventory and Pricing Decision of an Integrated Cold Chain Based on Freshness[J]. China Management Science,2018,26(7):10.

[4] Xu H. Research on inventory control of supermarket fresh goods based on demand forecasting[D]. Henan University of Finance and Economics and Law,2022.

[5] Feng Y, Hu Y, He L. Research on coordination of fresh agricultural product supply chain considering fresh-keeping effort level under retailer risk avoidance[J]. Discrete Dynamics in Nature and Society,2021,2021:1-15.

[6] http://www.mcm.edu.cn/html_cn/block/8579f5fce999cdc896f78bca5d4f8237.html

[7] Wenjie Qin. Anderson acceleration for B-spline curve fitting[D]. University of Science and Technology of China,2021.

[8] Li Na. Lanzhou City based on EEMD-LSTMARIMA Air Quality Prediction Study[D]. Lanzhou University of Finance and Economics,2023.

[9] Xiaoqing Ma. Research on Price Forecast of Shandong Green Onion Based on Time Series Analysis[D]. University of Yantai,2021.

[10] Guan heshan, Wang Qian. Improved Method of Box-Pierce Q Test[J]. Statistics & Decision,2017(17):28-32.