Research on superstore vegetable replenishment decision-making based on time series and PSO particle swarm algorithm

Zijian Li [#, *], Chengyu Li [#], Yongcheng Zhang [#]

Xiamen University, Xiamen, China, 361005

* Corresponding Author Email: zijianli11@163.com

#These authors contributed equally.

Abstract. Because vegetables are easy to wear and tear and not easy to store, supermarkets will replenish vegetables every day. In order to make reasonable replenishment and pricing decisions, this paper firstly correlates the total sales data with other data, and finds that the total sales data is negatively correlated with the average purchase price and average selling price, and positively correlated with the number of discounts, and then adopts the Feedforward Neural Network (FNN) model, the Evenberg-Marquardt training algorithm, and the time series analysis prediction to predict the total daily replenishment and pricing policy of each vegetable in the coming week. Marquardt training algorithm, time series analysis and forecasting, predicting the total daily replenishment and pricing strategy of each vegetable in the coming week, and at the same time, adopting the method of dynamic planning, through the use of particle swarm algorithm of PSO to solve the total replenishment of each category and pricing strategy, and then iterated to get the total daily replenishment of each category of commodities and pricing strategy in the coming week.

Keywords: Neural Networks, Dynamic Planning, PSO Particle Swarm Algorithm.

1. Introduction

Vegetable products are an important category of commodities essential to the operation of fresh food superstores, which are easy to wear and tear and not easy to store, leading to the timeliness of their sales, and the increase of sales time will seriously affect the sales volume and sales price of vegetables. Seeking a balanced relationship between product freshness and product profitability has always been a concern for supermarkets [1-2]. The wide variety of vegetable varieties and origins, and the particular time of purchase and transaction, make replenishment and pricing decisions difficult for merchants. From the different perspectives of supply and demand, the correlation between sales volume and time, the seasonal timing of vegetable supply, and the spatial limitation of supermarket sales all have an impact on market demand analysis [3-4]. Therefore, accurate and reasonable market analyses of historical sales and demand for each product are important for supermarkets to make reasonable replenishment and pricing decisions. In order to give the total daily replenishment and pricing strategy of each vegetable category for one week, so as to maximise the income of supermarkets, this paper firstly carries out data preprocessing, deletes the missing values, and summarises the valid data through time and category division. Visual icons can be drawn to observe the relationship between the average selling price, the average purchase price, the number of discounts and the total sales volume of each type of goods, and to do correlation analysis between the total sales volume and the average selling price. In order to give the total daily replenishment and pricing strategy of each vegetable category in a week, we use the front-end neural network model for time series analysis and prediction, and select valid data and divide them into two categories according to a certain proportion: training set (80%) and prediction set (20%). At the same time, we use the PSO legacy algorithm to solve the dynamic programming method to solve the total daily replenishment and pricing strategy for one week [5].

2. Establishment and solution of replenishment model

2.1. Data processing and visualisation and analysis

Firstly, we carry out data preprocessing, delete the data before 5 July 2020, take seven days as a cycle, divide the data by time and category, aggregate the data, get the total amount of sales of each type of goods, the average selling price, the average price of goods [6-7], the number of times of discounts, and draw the correlation table matrix of these four types of data, and make correlation analysis. Here we take the foliage category as an example (Figure 1):

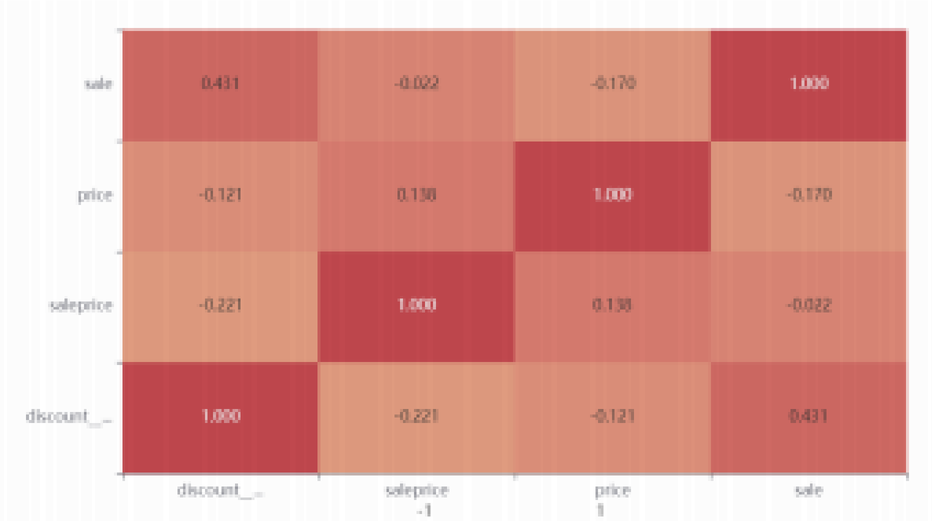

Figure 1. Correlation matrix of foliar data.

The analysis reveals a weak positive correlation between the number of discounts and the total number of sales, and a weak negative correlation between the average purchase price and the average selling price and the total number of sales, which leads to the preliminary conclusion that the relationship between the total number of sales and cost-plus pricing of foliage and flowers products is relatively stable [8-9].

Plotting the relationship between the average selling price and the total volume of sales for each type of product, taking the foliage product as an example (Figure 2):

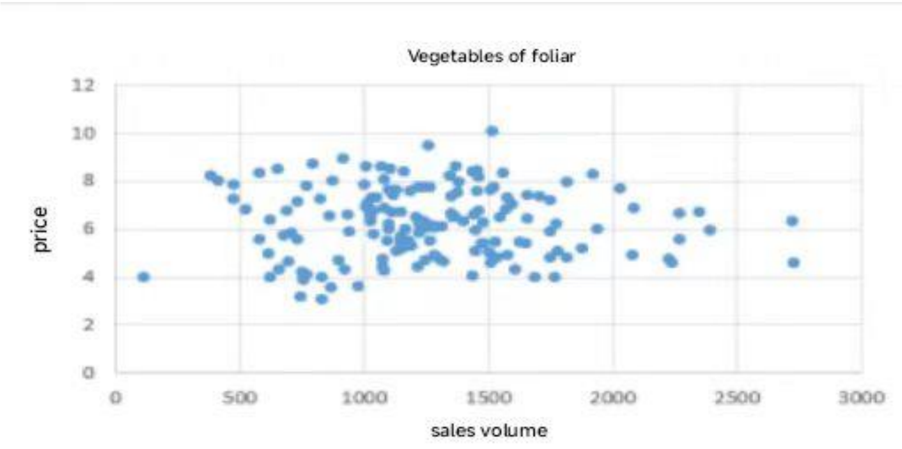

Figure 2. Pricing vs. sales for foliage category.

The analysis shows that the relationship between pricing and sales volume is relatively stable, with the overall sales volume decreasing as pricing increases, but there are also cases where sales volume decreases as pricing decreases, so we speculate that this type of change is related to the number of discounts [10].

In the following, we analyse the impact of the number of discounts on the total number of sales and draw a visualisation, again using the flowers and foliage category as an example (Figure 3):

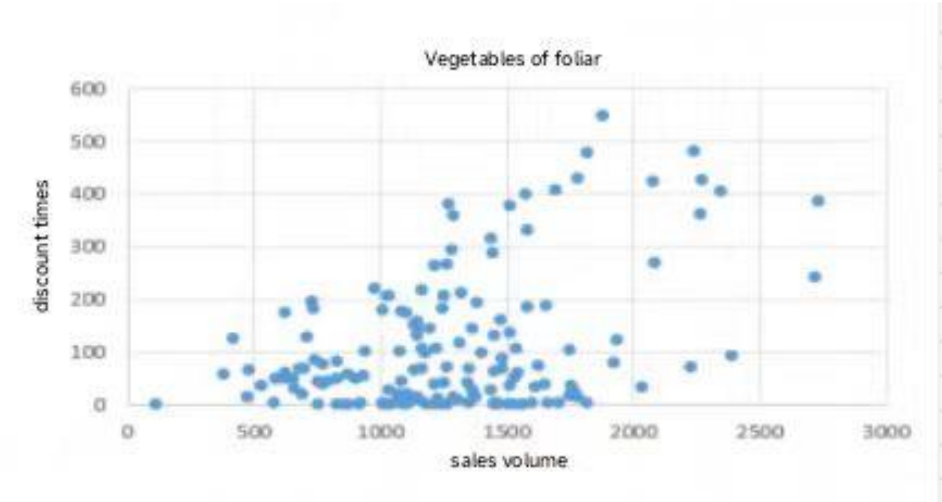

Figure 3. Relationship between Discounting and Sales in the Foliage Category.

It is found that when the number of discounts goes up, the probability that the total sales volume becomes larger also goes up, which is reflected in the image: that is, when the number of discounts is high, the frequency level of high-sales volume data is higher.

Then the relationship between the purchase price of goods and the total sales volume is analysed and a visualisation image is drawn, as shown in Figure 4 (for example, for flowers and foliage).:

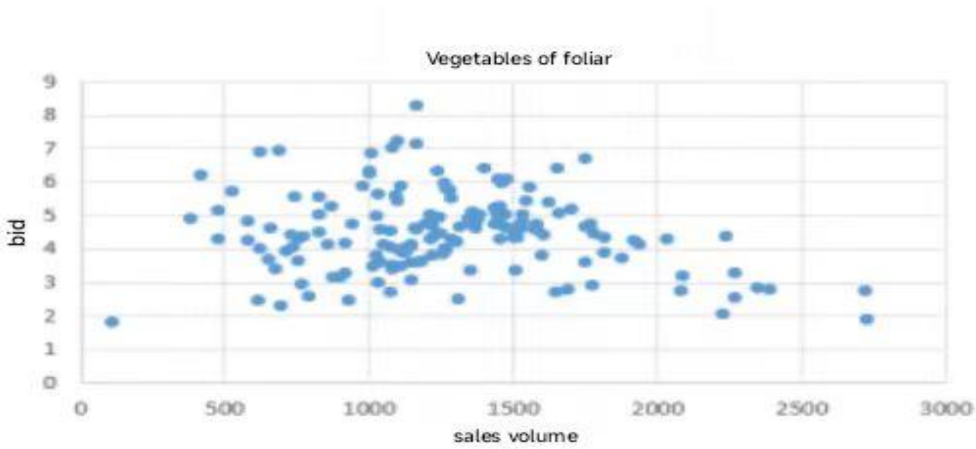

Figure 4. Relationship between purchase price and sales volume of foliage and flowers.

The analysis reveals that the highest sales volume of foliage and flowers occurs when the purchase price is low. In other words, the total sales volume is negatively correlated with the purchase price, but the total sales volume is stable in general.

In other words, the total sales volume is negatively correlated with the purchase price, but the total sales volume is stable. In summary, the relationship between the total sales volume and the cost-plus pricing of the six vegetable categories is as follows: in general, the total sales volume is negatively correlated with the average purchase price and the average selling price, and it is positively correlated with the number of discounts.

Total sales of foliage and flowers are stable and less affected by changes in cost-plus pricing.

Weak negative correlation between the number of discounts and total sales in the cauliflower category.

Total eggplant sales have a good negative correlation with average purchase price and average selling price, and a good positive correlation with the number of discounts.

Total sales of edible mushrooms category are strongly influenced by the purchase price of inputs Total sales of chilli category are affected to some extent by the number of discounted sales.

Total sales of aquatic roots and tubers category is affected to a greater extent by the average purchase price and the average selling price.

2.2. Time series analysis based on Matlab language

According to the data given in the title, preprocess the data, delete the missing values, select the sales records of each vegetable category in June 2021 and June 2022 as the training set, adopt the sales in June 2023 as the prediction set, normalise the total sales volume, and weight the wastage rate, average purchase price and average selling price. Firstly, the sales volume and average loss rate of each category of vegetables in the coming week were obtained, and the feedforward neural network model (FNN) was used to predict the sales volume of each category.

Feedforward Neural Network, Evenberg-Marquardt training algorithm, by constantly adjusting the number of training times and the ratio of the test set to the training set, to find the optimal parameter results that can predict the total daily replenishment in the coming week, and we carried out a number of adjustments to the parameter results, taking the flower and leaf category as an example (Figure 5):

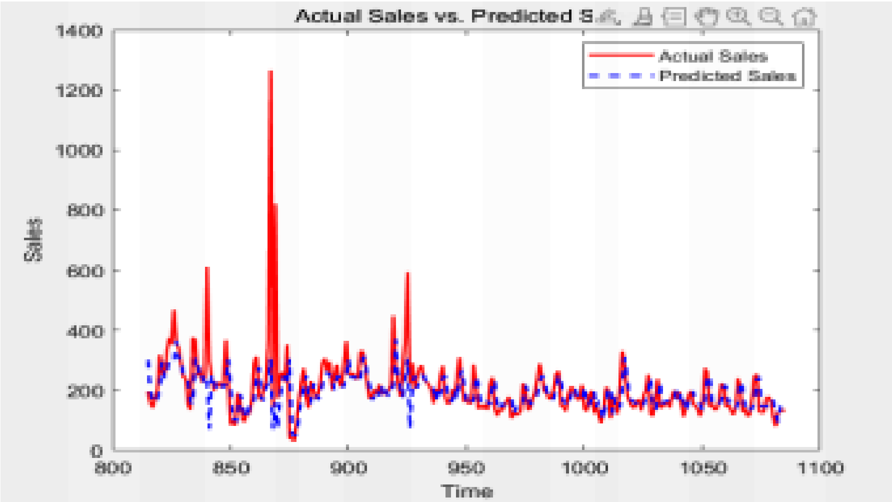

Figure 5. Predictions from foliar neural network fitting analysis.

As shown above, during the analysis process using the front-end neural network model, it was found that the most accurate model fitting results were obtained when the training set (80%), the prediction set (20%), the number of nodes in the hidden layer was 15, and the number of training times was 10,000, with a minimum mean square error of 0.005398.

Taking the attrition rate into account, it is easy to obtain the formula.

$$x = \frac{y}{1-a} \tag{1}$$

The daily replenishment forecast for the coming week for the Foliage category is given below as shown in Table 1:

Table 1. Forecast of daily replenishment for the coming week for the Foliage category.

Order	Prediction	Replenishment
1	153.028	174.847
2	163.246	186.522
3	165.849	189.522
4	166.569	189.548
5	166.756	190.319
6	166.809	190.593
7	166.824	190.611

Next, to solve for the pricing strategy for the week ahead, consider the following steps. 1:

1. Calculate the cost of each category: cost = wholesale price/1 - wastage rate, where the wholesale price is the purchase price and the wastage rate is all weighted. 2. Calculate the pricing: selling price = cost × markup.

2. Calculate the pricing: selling price = cost × markup rate, according to the forecast sales volume of each category b to determine the markup rate, such as sales volume to meet the 15 ≤ b ≤ 35 cases, we meet the 15 ≤ b ≤ 35 cases.

b ≤ 35, we develop a pricing strategy with a markup rate of 140%.

The final daily replenishment quantity and pricing strategy for each category are obtained as follows in Table 2:

Table 2. Replenishment and pricing results.

Foliar	Forecast	Cost price	Price	Solanacea	Forecast	Cost price	Price
2023/7/1	153.028	3.988	4.785	2023/7/1	13.347	6.137	9.206
2023/7/2	163.246	3.988	4.785	2023/7/2	15.259	6.137	8.592
2023/7/3	165.894	3.988	4.368	2023/7/3	17.755	6.137	8.592
2023/7/4	166.569	3.988	4.368	2023/7/4	20.547	6.137	8.592
2023/7/5	166.756	3.988	4.368	2023/7/5	22.151	6.137	8.592
2023/7/6	166.809	3.988	4.368	2023/7/6	22.641	6.137	8.592
2023/7/7	166.824	3.988	4.368	2023/7/7	22.774	6.137	8.592
Cauliflower	**Forecast**	**Cost price**	**Price**	**Capsicum**	**Forecast**	**Cost price**	**Price**
2023/7/1	30.133	6.878	9.629	2023/7/1	86.745	5.713	6.855
2023/7/2	34.809	6.878	9.629	2023/7/2	85.598	5.713	6.855
2023/7/3	36.941	6.878	8.941	2023/7/3	84.986	5.713	6.855
2023/7/4	38.276	6.878	8.941	2023/7/4	84.642	5.713	6.855
2023/7/5	39.352	6.878	8.941	2023/7/5	84.443	5.713	7.427
2023/7/6	40.358	6.878	8.941	2023/7/6	84.326	5.713	7.427
2023/7/7	41.387	6.878	8.941	2023/7/7	84.257	5.713	7.427
Aquatic rhizomes	**Forecast**	**Cost price**	**Price**	**Edible fungi**	**Forecast**	**Cost price**	**Price**
2023/7/1	28.754	6.744	9.442	2023/7/1	51.587	6.159	8.007
2023/7/2	34.874	6.744	9.442	2023/7/2	55.794	6.159	8.007
2023/7/3	39.828	6.744	8.767	2023/7/3	61.324	6.159	8.007
2023/7/4	42.663	6.744	8.767	2023/7/4	68.090	6.159	8.007
2023/7/5	43.888	6.744	8.767	2023/7/5	74.641	6.159	8.007
2023/7/6	44.354	6.744	8.767	2023/7/6	78.941	6.159	8.007
2023/7/7	44.537	6.744	8.767	2023/7/7	88.885	6.159	7.931

2.3. Dynamic Programming

Analysis of dynamic programming problems based on objective functions

$$Max(1-a)*(P_s-P_P)*S \tag{2}$$

Given the constraints

$$P_s > P_P, S > 0 \tag{3}$$

Solve this problem using PSO algorithm using python.

Firstly, we import the data weighted by sales volume such as wholesale price and wastage rate, that is to say, we calculate the weighted average wholesale price and weighted average wastage rate of each category weighted by sales volume, and fit the relationship between the predicted replenishment volume, pricing, and wholesale price of each category of goods on 1 July using XGBoost, and the accuracy of the regression is very high, with a score of 0.9144, which makes the model highly credible. Then, based on this model, the PSO algorithm is used to solve the replenishment quantity and pricing from 2 to 7 July 2023 by multiple iterations, taking the second iteration of the aquatic rhizome category as an example:

The PSO genetic algorithm is used for optimisation, and the convergence of the algorithm can be seen in Fig. 6, and the model converges after 25 iterations, and better fitting results are obtained, (Final results obtained.

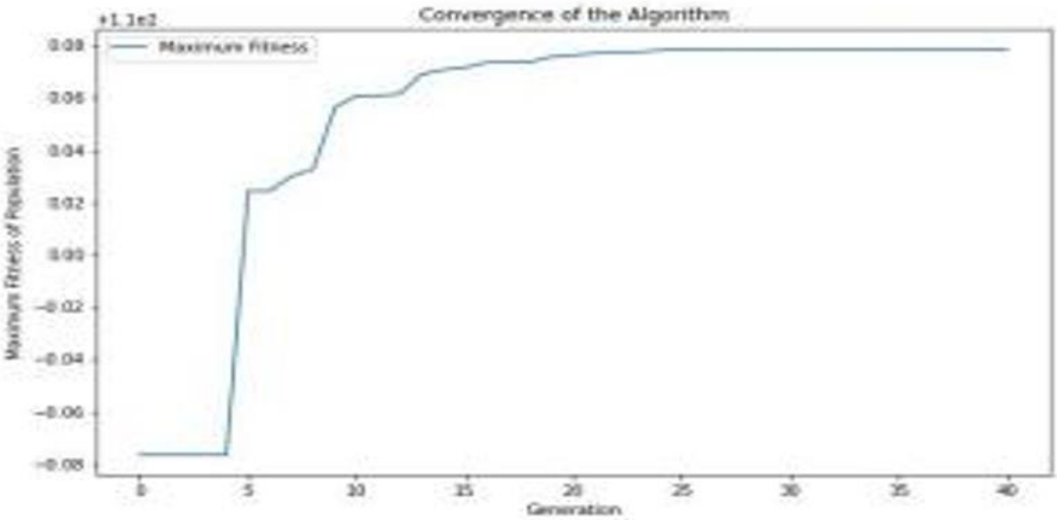

Figure 6. Convergence diagram of the aquatic rhizome class algorithm.

The calculated final replenishment, pricing and revenue results are shown in Table 3.

Table 3. Final replenishment, pricing and revenue results.

Days	Price	Sales	Earnings	Category	Days	Price	Sales	Earnings	Category
Day1	8.31	478.37	2016.187	Foliar	Day1	7.93	56.12	183.519	Solanacea
Day2	8.32	513.05	2254.900	Foliar	Day2	7.93	44.59	134.775	Solanacea
Day3	4.85	417.68	488.916	Foliar	Day3	7.93	41.48	127.995	Solanacea
Day4	4.85	417.68	489.181	Foliar	Day4	9.82	25.15	129.454	Solanacea
Day5	4.85	326.82	373.542	Foliar	Day5	7.93	49.40	166.875	Solanacea
Day6	8.32	458.90	1882.276	Foliar	Day6	26.52	51.89	1079.022	Solanacea
Day7	9.72	396.55	2223.001	Foliar	Day7	7.93	41.48	127.148	Solanacea
Day1	13.52	61.61	298.004	Cauliflowers	Day1	8.39	116.07	276.587	Capsicum
Day2	13.55	63.39	319.030	Cauliflowers	Day2	13.82	206.43	1593.675	Capsicum
Day3	13.56	41.30	235.661	Cauliflowers	Day3	11.74	251.75	1599.232	Capsicum
Day4	11.23	59.70	212.052	Cauliflowers	Day4	17.13	209.09	2359.611	Capsicum
Day5	13.56	49.41	268.161	Cauliflowers	Day5	12.12	193.31	1320.436	Capsicum
Day6	11.84	42.88	196.283	Cauliflowers	Day6	20.60	413.24	5955.449	Capsicum
Day7	13.56	48.33	295.635	Cauliflowers	Day7	7.84	157.08	388.854	Capsicum
Day1	13.13	111.99	162.259	Aquatic rhizomes	Day1	13.57	80.83	380.424	Edible fungi
Day2	17.25	35.39	205.891	Aquatic rhizomes	Day2	12.42	99.32	313.918	Edible fungi
Day3	13.11	109.65	259.985	Aquatic rhizomes	Day3	13.09	217.12	766.870	Edible fungi
Day4	13.13	119.59	421.804	Aquatic rhizomes	Day4	11.28	211.10	445.057	Edible fungi
Day5	13.10	82.20	255.052	Aquatic rhizomes	Day5	11.93	142.83	527.954	Edible fungi
Day6	12.75	110.09	363.701	Aquatic rhizomes	Day6	11.29	231.83	816.774	Edible fungi
Day7	11.26	90.86	264.095	Aquatic rhizomes	Day7	11.21	165.21	628.746	Edible fungi

3. Conclusions

In this paper, they first pre-processed the data, deleted the data before 5 July 2020, and aggregated the data by time and category to obtain indicators such as the total sales volume, the average selling price, the average purchase price, and the number of discounts for each type of goods. They then plotted a graphical matrix of correlations between the different indicators and analysed the data specifically for the flower and foliage category of goods.

In the analyses of the flower and foliage category of goods, a weak positive correlation was found between the number of discounts and the total number of sales, while a weak negative correlation was found between the average purchase price, the average selling price and the total number of sales. They also plotted the relationship between average selling price and total sales and the relationship between number of discounts and total sales and found that the relationship between pricing and total sales is more stable, while the probability of total sales becoming larger when the number of discounts goes up also goes up. Then, time series analysis based on Matlab language was carried out, the data were preprocessed and normalised, and the feed-forward neural network model and Levenberg-Marquardt training algorithm were used for model training and prediction. Taking floral and foliage goods as an example, they obtained the most accurate model fitting results and predictions by adjusting the parameters such as the number of training times and the ratio of the test set to the training set several times. In addition, they used dynamic programming and PSO algorithms for model solving. They used the XGBoost model to fit the relationship between the predicted replenishment quantities, pricing and wholesale price of each category of goods, and solved it using the PSO algorithm for several iterations to obtain the replenishment quantity and pricing strategy for July 2023.

References

[1] Zhao Wendan, Zhang Meichen. Optimisation study of cross replenishment control in dual channel supply chain [J]. Manufacturing Automation, 2023, 45(09):44-48.

[2] Cai Jianhu, Cao Weizang, Zhou Qian et al. Research on VMI inventory decision model for assembly system under supplier output uncertainty environment [J/OL]. China Management Science: 1-14[2023-10-20].

[3] Junxia Wang. A review of research on joint replenishment based on perishables grouping strategy [J]. China Storage and Transportation, 2023(08):97-98.

[4] Zheng Cou. Research on dynamic bulk inventory replenishment problem of community group-buying retailers [D]. Zhongnan University of Economics and Law, 2022.

[5] Wu, Yannan. Increase the frequency of replenishment to ensure personnel input [N]. Shenzhen Special Zone News, 2022-03-15(A03).

[6] Shuting Liang. Research on inventory strategy of community fresh food retail terminal where time and stock quantity jointly affect spoilage rate [D]. Beijing Jiaotong University, 2021.

[7] Qiao Yannan. Research on Q supermarket outlet layout optimisation scheme based on improved SLP [D]. Tianjin University of Finance and Economics, 2020.

[8] Epidemic now, supermarket fresh C position "leading the way", these fresh basic skills must be strengthened [J]. Corporate Observer, 2020(03):111-113.

[9] Zhang Jingfang, Xiang Wanli. Application of dynamic buffer inventory model in agricultural products inventory management [J]. Logistics Technology, 2018, 37(12):83-88.

[10] Han Wenguang. Research on inventory management programme of frozen products category in community vegetable direct shops in Urumqi [D]. Xinjiang Agricultural University, 2020.