

# Research on Pricing and Replenishment Strategy for Vegetable Commodities based on Non-linear Programming

Tingting Chen <sup>#, \*</sup>, Lihong Zhou <sup>#</sup>, Senxin Guo <sup>#</sup>

College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610225, China

\* Corresponding author Email: yilantingting@qq.com

<sup>#</sup>These authors contributed equally to this work

**Abstract.** With the improvement of people's consumption level and the change in consumer mindset, how to develop pricing and replenishment strategies has become an urgent issue for fresh food supermarkets. This paper first utilizes an improved ARIMA-LSTM model to forecast the purchase price and sales volume of six types of vegetables for the next 7 days. It then conducts regression analysis on the sales volume and cost-plus pricing using GBoost algorithm. Furthermore, a Non-linear Programming model for maximizing profits is established, and this article solves it based on optimizing Genetic Algorithm and obtains a pricing and replenishment strategy finally. The strategy shows that the pricing of "Edible mushrooms" is the highest and the daily replenishment quantity of "Leafy greens" is the largest. The maximum profit over seven days is 6978 yuan. Overall, this study contributes to optimizing sales and supply chain management, achieving better profitability and customer satisfaction.

**Keywords:** Pricing and Replenishment Strategy; Non-linear Programming; ARIMA-LSTM; GBoost.

## 1. Introduction

As an emerging retail model, fresh food supermarkets have gradually become the primary shopping channel for consumers to purchase vegetable products.

To accurately formulate pricing and replenishment strategies for vegetable products, Fan (2019) applied a regression model to pricing of vegetables [1]. Zou et al. (2018) proposed a joint replenishment method for multiple perishable products under random demand [2].

Previous studies often focus on specific stages of vegetable products and rely on small sample data or simulated data. Therefore, this study utilizes real and comprehensive data from a certain supermarket chain in recent years. It includes not only detailed sales information but also considers wholesale prices and losses rates, among other factors.

This paper aims to provide guidance to operators in the fresh food supermarket industry or related industries by comprehensively applying mathematical modeling and algorithm optimization methods. It assists them in formulating optimal pricing and replenishment strategies for vegetable products, ultimately improving profitability and customer satisfaction.

## 2. Data Selection and Variable Meaning

### 2.1. Data Selection

In this study, a certain supermarket chain is selected as the research object. The data includes product information for the six vegetable categories distributed by the market, sales transaction details and wholesale price data from July 1, 2020, to June 30, 2023, and recent loss rate data for each product.

## 2.2. Variable Meaning

**Table 1.** Variable definitions

| Variable             | Definition                             | Unit |
|----------------------|----------------------------------------|------|
| <i>Waste</i>         | Loss rate                              | %    |
| <i>MaxPrediction</i> | Upper limit of sales forecast interval | -    |
| <i>MinPrediction</i> | Lower limit of sales forecast interval | -    |
| <i>Replen</i>        | Replenishment quantity                 | -    |
| <i>Salevalue</i>     | Cost-plus pricing                      | yuan |
| <i>Cost</i>          | Unit purchase price                    | yuan |
| <i>Dayincome</i>     | Daily income                           | yuan |
| <i>Daysalevalue</i>  | Daily total sales                      | yuan |
| <i>Daycost</i>       | Daily procurement cost                 | yuan |
| <i>Cpratio</i>       | Cost-profit rate                       | %    |
| <i>G(Cpratio)</i>    | Sale volume                            | -    |

(Cont'd)

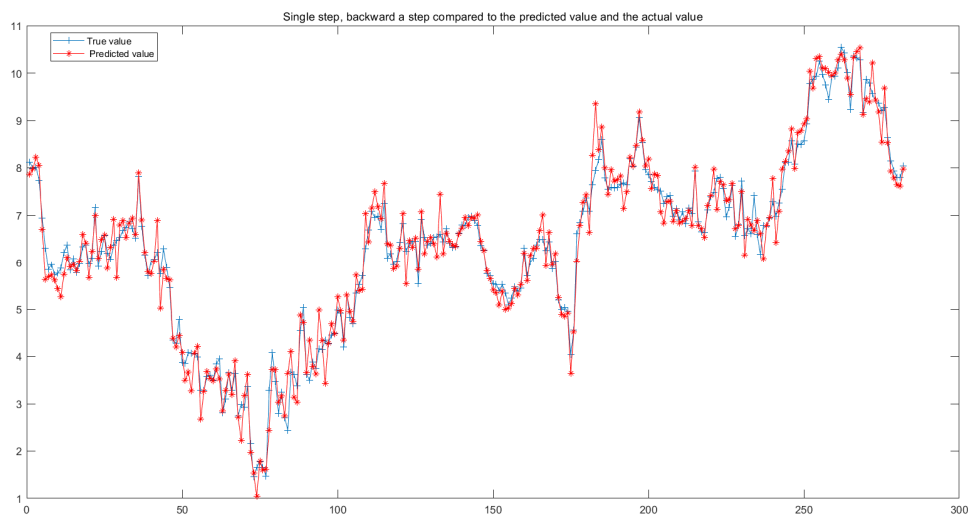
## 3. Prediction of Sales and Purchase Prices based on ARIMA-LSTM Model

### 3.1. Establishment of ARIMA Model

This paper preprocesses the time series data using the ARIMA model to analyze the trends and seasonal information in the time series.

Taking the example of the time series analysis of the purchase unit price for "Brassica vegetables", it can be observed from the time series plot that there is a clear seasonal trend. After applying first-order differencing, the series becomes relatively stationary. Therefore, ARIMA (2,1,1) model is determined for constructing the time series forecasting model, and it passes the White Noise test.

### 3.2. Establishment of ARIMA-LSTM Model



**Fig 1.** ARIMA (2, 1, 1) real value and predicted value fluctuation curve

To use the preprocessed data as input for LSTM neural network model, the steps are as follows:[3]

Step 1 Define the LSTM neural network.

Step 2 Compile the network.

Step 3 Fit the network.

Step 4 Evaluate the network.

The fit of the prediction result is shown in Fig. 1.

It can be found that the model construction has achieved good results.

#### 4. Regression of Sales and Cost Margins based on GBoost Model

According to the concept of "supply and demand elasticity" in statistics, an increase in cost-profit ratio reflects an increase in selling price, which inevitably leads to a decrease in sales volume. Therefore, selecting the cost margins as an intermediate variable can reflect the supply-demand elasticity relationship between selling price and sales volume, as well as measure the relationship between purchase unit price and selling price. Based on GBoost, the paper performs regression fitting to model the relationship between the cost-profit rate and total sales volume [4].

The relationship between cost margin, selling price and pricing can be described as follows:

$$Cpratio = \frac{Salevalue}{Cost} - 1. \quad (1)$$

The regression equation for the relationship between vegetable category sales volume and cost margin is as follows:

$$G(Cpratio) = \sum \alpha_i Cpratio^i. \quad (2)$$

In the formula,  $G(Cpratio)$  represents the relationship between total sales volume and cost margin,  $\alpha$  is polynomial coefficient.

Based on the given sales and pricing data, the solution using GBoost in machine learning follows these steps:

Step 1 Input the training dataset and the chosen loss function.

Step 2 Output the regression trees.

Step 3 Generate gradient regression trees.

#### 5. Pricing and Replenishment Strategy based on Non-linear Programming

##### 5.1. Model Establishment

In order to determine the replenishment quantities and pricing strategies for each day over the next seven days, this study establishes a non-linear programming model based on the results of the research above [5].

Considering the loss rate between sales and replenishment quantities, for each of the six vegetable categories, the average loss rate of all individual items within the category is taken as the category's loss rate. The 95% confidence interval of the forecasted sales is taken as an approximation of the maximum probable value for sales. Taking into account the differences between replenishment quantities and sales due to transportation or time losses, the predicted replenishment quantity is obtained by adding the portion of lost quantity due to losses to the maximum probable value of sales.

Therefore, the mathematical relationship model between the predicted sales and replenishment quantities can be established as follows:

$$Waste = Mean(Waste_i). \quad (3)$$

$$ReplenMax = (MaxPrediction - MinPrediction) \times 95\% \times (1 + Waste/100). \quad (4)$$

The total revenue of a supermarket over seven days can be obtained by summing up the daily income, where the daily income is calculated as the difference between the daily total sales and the daily procurement cost. Therefore, the relationship can be expressed by the following formula:

$$Dayincome = Daysalevalue - Daycost. \quad (5)$$

Similarly, there are the following equations:

$$Salevalue = Cost \times (1 + Cpratio). \quad (6)$$

$$Daysalevalue = Salevalue \times G(Cpratio). \quad (7)$$

$$Daycost = Cost \times Replen. \quad (8)$$

$$Replen = \frac{G(Cpratio)}{1 - Waste}. \quad (9)$$

Replenishment constraint: Taking into account the need to minimize storage losses while meeting customer demands, a replenishment quantity constraint is set.

Price constraint: With the aim of profitability, the supermarket should set the daily selling price to be greater than the purchase price. Additionally, considering the impacts including the pandemic and natural disasters from the past three years are already included in the dataset, the daily selling price should also be less than the historical maximum selling price.

Here is a summary of the dynamic programming model established based on the analysis above:

$$\begin{aligned} & \text{Max} \sum_{i=1}^6 \sum_{j=1}^7 Dayincome_{ij} \\ & = \sum_{i=1}^6 \sum_{j=1}^7 Salevalue_{ij} \times G(Cpratio_{ij}) \times \left[ 1 - \frac{1}{(1 + Cpratio_{ij}) \times (1 - Waste_i)} \right]. \\ & \text{s.t.} \begin{cases} G(Cpratio_{ij}) \leq ReplenMax_i \times (1 - Waste_i), \\ Cost_{ij} \leq Salevalue_{ij}, \\ Salevalue_{ij} \leq Maxsalevalue_i, \\ 0 \leq Cpratio_{ij}, \\ 0 \leq Cost_{ij}, \\ 0 \leq Salevalue_{ij}. \end{cases} \end{aligned} \quad (10)$$

## 5.2. Model Solving

Using Genetic Algorithm to solve the dynamic non-linear programming model, the daily replenishment quantities and pricing data for the six vegetable categories are shown in Table 2, with the top row corresponding to July 1st and the bottom row corresponding to July 7th: [6]

**Table 2.** Replenishment and pricing strategy results

| Vegetable category     | Daily replenishment | Pricing | Vegetable category         | Daily replenishment | Pricing |
|------------------------|---------------------|---------|----------------------------|---------------------|---------|
| Brassica<br>vegetables | 20.3713             | 18.9929 | Solanaceous<br>vegetables  | 14.7963             | 12.9689 |
|                        | 20.4908             | 18.9340 |                            | 14.8316             | 12.9512 |
|                        | 20.6083             | 18.8760 |                            | 14.8683             | 12.9327 |
|                        | 20.7258             | 18.8181 |                            | 14.9001             | 12.9168 |
|                        | 20.8531             | 18.7553 |                            | 14.9309             | 12.9013 |
|                        | 20.9614             | 18.7019 |                            | 14.9700             | 12.8816 |
|                        | 21.0679             | 18.6494 |                            | 14.9974             | 12.8678 |
| Leafy<br>greens        | 122.4714            | 13.5026 | Edible<br>mushrooms        | 39.2617             | 23.6804 |
|                        | 122.9659            | 13.4629 |                            | 39.5440             | 23.5810 |
|                        | 122.1526            | 13.5282 |                            | 39.8528             | 23.4723 |
|                        | 122.0616            | 13.5355 |                            | 40.1656             | 23.3622 |
|                        | 122.8679            | 13.4708 |                            | 40.4931             | 23.2469 |
|                        | 121.7259            | 13.5624 |                            | 40.8445             | 23.1232 |
|                        | 121.5525            | 13.5763 |                            | 41.5906             | 22.8605 |
| Chili<br>peppers       | 51.1270             | 21.2321 | Aquatic root<br>vegetables | 15.8241             | 16.2991 |
|                        | 50.8869             | 21.3094 |                            | 16.5005             | 16.1763 |
|                        | 50.7547             | 21.3519 |                            | 17.1386             | 16.0604 |
|                        | 50.6009             | 21.4013 |                            | 17.8307             | 15.9347 |
|                        | 50.4136             | 21.4616 |                            | 18.4465             | 15.8228 |
|                        | 50.2585             | 21.5114 |                            | 18.9493             | 15.7315 |
|                        | 50.1756             | 21.5381 |                            | 19.4763             | 15.6358 |

(Cont'd)

From the replenishment strategy, it can be observed that "Leafy greens" have the highest daily replenishment quantity, followed by "Chili peppers" and "Edible mushrooms". "Brassica vegetables", "Aquatic root vegetables" and "Solanaceous vegetables" have relatively smaller replenishment quantities.

From the pricing strategy, it can be seen that there is a higher demand for "Leafy greens", indicating their essential role in daily life. Therefore, they should to be priced to be affordable and suitable. "Brassica vegetables" are complementary to "Aquatic root vegetables", and "Chili peppers" are complementary to "Edible mushrooms", so their pricing does not differ significantly. "Solanaceous vegetables" are relatively less popular, resulting in the lowest pricing. Finally the maximum profit obtained is 6978 yuan.

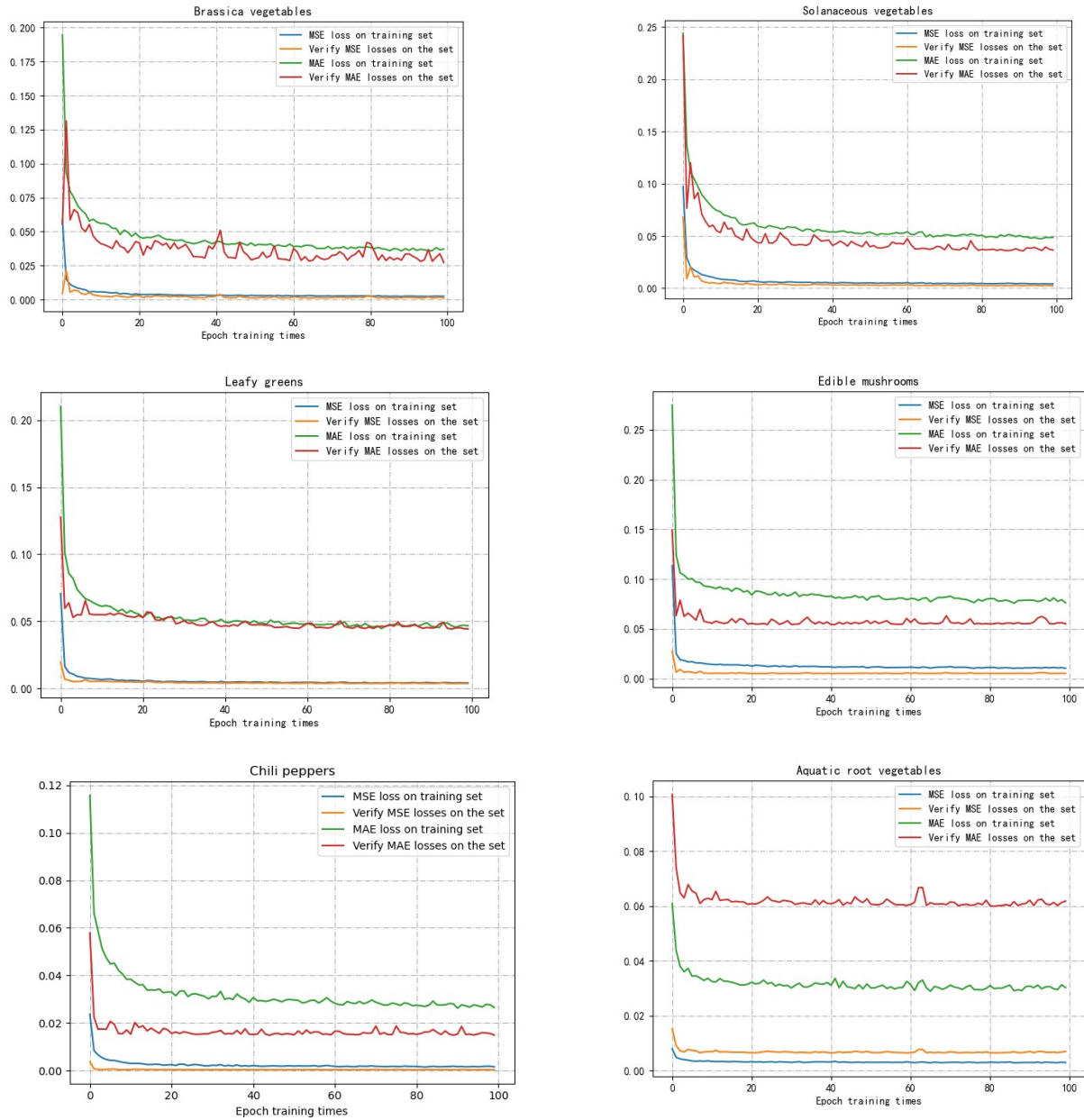
## 6. Model Evaluation

The model's parameter tuning optimization process performance chart is shown in Fig. 2.

## 7. Summary

This paper combines non-linear programming models, improved ARIMA-LSTM algorithm, GBoost algorithm, and Genetic Algorithm to optimize the pricing and replenishment strategy of fresh food supermarkets. Through experimental results, it identifies pricing and replenishment characteristics of different vegetable categories, providing guidance for sales and supply chain management.

Future studies can conduct more field investigations and empirical analyses to validate the feasibility and effectiveness of the strategies in real markets.



**Fig 2.** Algorithm iteration error convergence chart

Overall, the models have a small loss error, indicating a good construction effect.

## References

- [1] Fan Fengyan: Research on Supermarket Fresh Food Pricing Model (Master, Fujian Normal University, China 2019) [D]. p.9.
- [2] Zou Guangjun, Fu Shaochuan. Joint Replenishment Optimization of Multiple Perishable Products under Random Demand [J]. Management Science and Engineering. Vol. 7 (2018) No.1, p. 9-19.
- [3] Xu Lumeng. Prediction of Public Opinion Popularity for Emergent Events Based on ARIMA and Improved LSTM [J]. Cyber Security and Data Governance. Vol. 42 (2023) No. 3, p. 53-57.
- [4] Ke, G., Meng, Q., Finley, T., et al. A Highly Efficient Gradient Boosting Decision Tree [J]. Advances in Neural Information Processing Systems. Vol. 30 (2017) No. 2, p. 367-378.
- [5] Dimitri P.Bertsekas. Nonlinear programming [M]. Tsinghua University Press, 2013, p. 40-48.
- [6] Xuan Guangnan, Cheng Runwei. Genetic Algorithms and Engineering Optimization [M]. Tsinghua University Press, 2004, p. 117-125.