Wavelet Neural Network Based Study of Merchandise Replenishment and Pricing

Yousheng Lei*

School of international education, Neusoft Institute Guangdong, Foshan, China,528225

*Corresponding author: 18027348136@163.com

Abstract. Merchandise replenishment pricing strategy is the most critical step in a merchant's sales decision. The formulation of reasonable and accurate replenishment strategy is of great significance for controlling the cost of sales and improving economic efficiency. In order to reasonably formulate the replenishment strategy and price, this paper is based on the wavelet neural network theory, increased the forecast model for the nonlinear relationship between the commodity sales portfolio, combined with the Spearman correlation coefficient to fully explore the correlation between the commodity sales portfolio and its impact on the number of commodities for sale, the use of cost plus pricing method and the rate of loss of commodities with the loss of value of the commodities over time, to deal with, construct a dynamic pricing and replenishment model of goods. The model is more capable of handling the non-relationships in merchandising and is more generalizable than the traditional time series forecasting model. Compared with previous studies, this model better handles the relationship between merchandise sales combinations through wavelet transform, solves the limitation that time series forecasting cannot fully consider the correlation relationship between merchandise sales combinations, and provides a new solution for the treatment of sales combinations in merchandise sales forecasting.

Keywords: Wavelet neural network, Spearman's correlation coefficient, Predictive model.

1. Introduction

In daily life, the pricing and replenishment strategies of daily commodities are particularly important for vegetable-based superstores. Vegetable categories have a relatively short shelf life and their quality deteriorates with selling time [1], which can affect the selling price or make them Supermarkets usually restock based on historical sales and demand for each category. However, due to the many varieties of vegetables, different origins, and the time of purchase in the early hours of the morning, merchants must make replenishment decisions without a clear understanding of the specific individual products and purchase prices. Replenishment decisions and pricing decisions are particularly important for supermarkets, need to be analysed through reliable market demand, from the demand side of the analysis, the sales volume of vegetable products and the time of day there is a certain correlation; from the supply side of the analysis, the sales space of the superstore sales space limitations on the composition of a reasonable sales mix is extremely important [2].

In previous studies, most scholars use ARIMA model to forecast commodity prices [3-5], because ARIMA model can only deal with the linear relationship of the time series [6], and can not fully consider the nonlinear relationship of the commodity sales mix, so wavelet optimization can be introduced to optimize the forecasting method [7], such a method can be used in the nonlinear inventory holding cost makes the superstore sales of vegetable category Commodity profit maximization [8].

This paper is based on wavelet neural network for commodity data research and analysis, to discover the influence of commodity sales portfolio correlation on the future sales of commodities, adding the influence of this factor in the model to optimize the prediction model, so as to improve the formulation of the cost-plus pricing commodity strategy, so as to enable it to more comprehensively take into account the actual situation [9], and to make the final results more in line with the expectations [10].

2. Fundamentals of the model

2.1. Structure of wavelet neural networks

Wavelet neural network is a network developed on the basis of the topology of error back propagation neural network*, which has certain similarities with the structure of neural networks. In wavelet neural network, while the overall signal propagates forward, the error propagates backward, and its wavelet transform decomposes the original time series to obtain wavelet coefficients at multiple scales. These wavelet coefficients are then used as inputs for training and prediction by the neural network, which achieves multi-scale refinement through scaling and translation, highlighting the details of the problem to be dealt with and effectively extracting local information. Its conceptual diagram is shown in Figure 1:

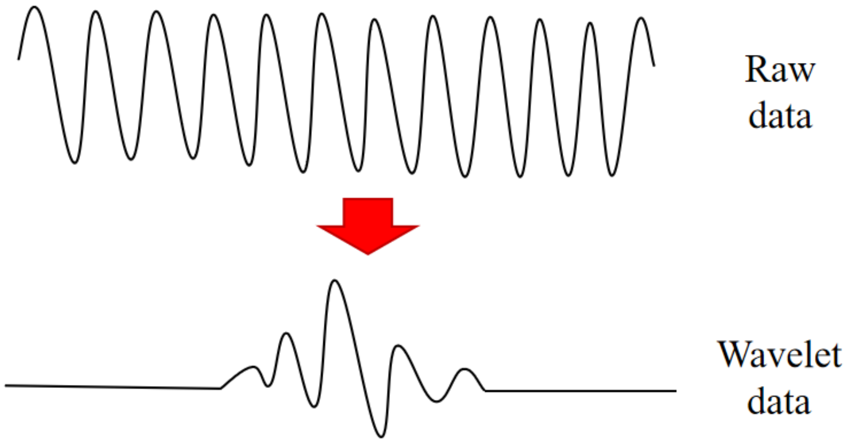

Figure 1. Conceptual diagram

The advantage of the wavelet neural network prediction method is that it can make full use of the multi-scale analysis ability of the wavelet transform and the nonlinear fitting ability of the neural network, so as to improve the accuracy and stability of the time series prediction, due to the development of error back propagation neural network to the network, the input layer, the intermediate layer, and the output layer are still the core of the constituent structure, and the topology of its network structure is shown in Figure 2:

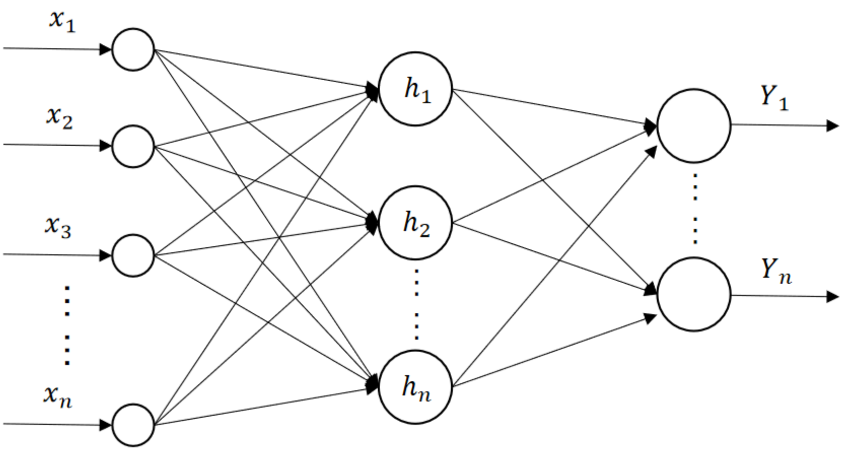

Figure 2. Neural network structure

The general model of wavelet neural network consists of three basic elements:

(1) Forward calculation formula:

$$h(j) = hj\left[\frac{\sum_{i=1}^{n} \omega_{ij} x - b_j}{a_j}\right] \qquad j = 1,2,...,\mathrm{m} \tag{1}$$

$$\mathrm{y}(k) = \sum_{\mathrm{i=1}}^{\mathrm{m}} \omega_{\mathrm{ik}} h(i) \qquad k = 1,2,...,l \tag{2}$$

(2) Error back propagation calculation formula:

$$e = \sum_{k=1}^{m} yn(k) - y(k) \tag{3}$$

where y(k) is the predicted output value, yn(k) is the expected output value for the network

$$\omega_{n,k}^{(i+1)} = \omega_{n,k}^{(i)} + \Delta\omega_{n,k}^{(i+1)} \tag{4}$$

$$a_k^{(i+1)} = a_k^{(i)} + \Delta a_k^{(i+1)} \tag{5}$$

$$b_k^{(i+1)} = b_k^{(i)} + \Delta b_k^{(i+1)} \tag{6}$$

where $\Delta\omega_{n,k}^{(i+1)}$, $\Delta a_k^{(i+1)}$, and $\Delta b_k^{(i+1)}$ are calculated by the networkk prediction error:

$$\Delta\omega_{n,k}^{(i+1)} = -\mu \frac{\partial e}{\partial \omega_{n,k}^{(i)}} \tag{7}$$

$$\Delta a_k^{(i+1)} = -\mu \frac{\partial e}{\partial a_k^{(i)}} \tag{8}$$

$$\Delta b_k^{(i+1)} = -\mu \frac{\partial e}{\partial b_k^{(i)}} \tag{9}$$

where η is the learning rate.

(3) Wavelet basis function selection formula:

$$\Psi(t) = \cos(1.75t)\exp(-\frac{t^2}{2}) \tag{10}$$

2.2. Spearman correlation analysis

Spearman's correlation coefficient is a nonparametric statistical measure used to assess the monotonic relationship between two variables. It does this by performing a rank transformation on two variables and then calculating their Pearson correlation coefficients. This is done by examining the strength of the monotonic relationship, that is, the extent to which the two stay in step, if not in proportion, in terms of their tendency to become larger or smaller. Since merchandising data characteristics do not in practice conform to a normal distribution, the use of Spearman's correlation coefficient allows for better anecdotal correlation analysis of the data. The formula is shown below:

$$p=\frac{\frac{1}{n}\sum_{i=1}^{n}(R(x_i)-\overline{R(x)})(R(y_i)-\overline{R(y)})}{\sqrt{(\frac{1}{n}\sum_{i=1}^{n}(R(x_i)-\overline{R(x)})^2)(\frac{1}{n}\sum_{i=1}^{n}(R(y_1)-\overline{R(y)})^2)}} \tag{11}$$

where $R(x)$ and $R(y)$ are the bits of x and y respectively

2.3. Pricing and stocking models for profit maximization

Good pricing is usually well integrated with historical performance, and this is the case with the cost-plus pricing method, which, by combining cost and markup, ensures that the category's acquisition costs are covered, thus ensuring the profitability of the goods, and that cost-plus pricing has a certain degree of stability because it is based on calculations of its own costs, as shown in the following pricing model:

$$z_{ib}=\hat{c}_{ib}\left(1+\omega_i\right) \tag{12}$$

where z_{ib} represents the pricing on day *b* for category *i*, $\hat{c}_{ib}$ is the predicted average cost on day *b* for category *i*, ω_i is the cost-plus rate, which is derived from historical selling prices and historical costs and is calculated as follows:

$$\overline{\omega}_i=\frac{\sum_{t=1}^{30}\omega_{it}}{30}=\frac{x_{it}-c_{it}}{c_{it}} \tag{13}$$

When the supermarket wholesale goods, the wholesale price of different single products are not the same, this paper will be attributed to them as procurement costs, while according to a priori experience we can know that each single product has its corresponding loss rate, the loss rate can be understood as in the process of transportation, compared to the overall degree of loss of the purchase, so the average cost of a single commodity for the calculation of the formula is as follows:

$$c_{it}=N_{it}\overline{p}_{it}+N_{it}\overline{\eta}_i\overline{p}_{it} \tag{14}$$

where $\overline{\eta_i}$ is the average attrition rate for each category, N_{it} is the number of purchases per category per day, the formula is as follows:

$$N_{it}=\frac{x_{it}}{\left(1-\overline{\eta}_i\right)} \tag{15}$$

Considering that the merchants will be limited by the cost when purchasing the incoming goods, the incoming goods should be selected as much as possible to sell more goods. After obtaining the profit, as the perspective of the super, need to ensure that the profit maximization of commodity sales, so the formula will be found on all the goods into the single product for the sum, and the total profit for the ranking, the super preferred selection from the revenue from the product sales, in order to ensure that the demand for the premise that the previous day's sales to obtain the maximum profit. Therefore, the formula for calculating the profit of the superstore is as follows:

$$Max\ \ U_2=\sum_{i=1}^{n}\left(\hat{x}_{it}z_{it}-\hat{c}_{it}\right) \tag{16}$$

Based on the cost-plus pricing method combined with the forecast data the pricing scheme can be derived, as well as the replenishment scheme on that date, as shown in the following equation:

$$I_{it} = \hat{x}_{it} + \frac{\hat{x}_{it}}{(1-\overline{\eta}_i)}\overline{\eta}_i \tag{17}$$

Considering that in the use of actual sales, due to the lagging of cold goods, supermarkets are likely to select only some of the goods for replenishment in a stocking, then based on the analysis of the perspective of the superstore, it is possible that they need to look for sorting based on the categories they can sell in that period of time, so as to find out the profit-maximizing several categories of goods, replenishment and pricing design of the goods, and therefore this paper provides an algorithm to establish a screening model:

$$\hat{x}_{it} z_{it} - \hat{c}_{it} \begin{cases} > 0, & t_i = 1 \\ \leq 0, & t_i = 0 \end{cases} \tag{18}$$

where t_i "0-1" variables, when a single item can be filtered, a is set to1 and vice versa to 0

Eventually will be selected for the time period of sales of goods categories in its net profit is less than 0 part of the judgment, less than 0 can be directly removed from the screening queue, followed by different products with different profit size sorting, in the process of calculating the queue to select the largest profit until the selection of the number of replenishment categories need to be made up of the number of all the quota after the exclusion of the part of the loss may cause, as follows shown:

$$Max \;\; U_d = t_i\left(\hat{x}_{it} z_{it} - \hat{c}_{it}\right) \tag{19}$$

3. Results

3.1. Product relevance analysis by category

According to Spearman correlation coefficient commodity correlation model, *SPSS* software was used to process the data and calculate the correlation of each category its sales volume, and the correlation coefficient matrix was obtained by Spearman correlation coefficient calculation as follows:

$$M_{i(i+1)} = \begin{bmatrix} 0.19 & 0.25 & 0.10 & -0.12 & -0.21 & 1 \\ 0.40 & 0.44 & 0.33 & 0.60 & 1 & -0.21 \\ 0.46 & 0.60 & 0.54 & 1 & 0.60 & -0.12 \\ 0.43 & 0.59 & 1 & 0.54 & 0.33 & 0.10 \\ 0.63 & 1 & 0.60 & 0.60 & 0.44 & 0.25 \\ 1 & 0.63 & 0.43 & 0.46 & 0.40 & 0.19 \end{bmatrix}$$

It is difficult to comprehensively analyze the correlation between the categories through the matrix alone, so this paper visualizes the above matrix as a heat map of confusion matrix in order to more intuitively analyze the results is shown in Figure 3 below:

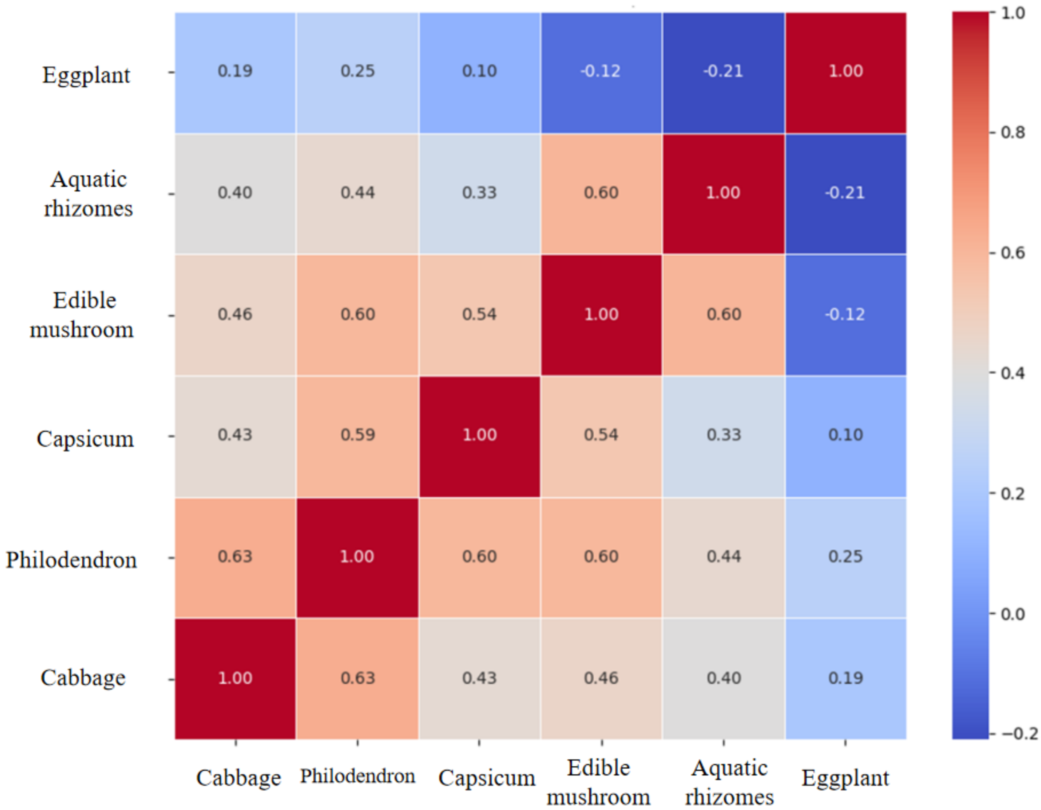

Figure 3. Heat map of correlation coefficient matrix

According to the above figure, the correlation coefficient is greater than 0.5 can be regarded as correlation between two variables, through the analysis of the above figure, to give one of the examples to illustrate, the correlation between flowers and leaves and cauliflower category is very strong, the reason may be divided into two kinds of factors, one that is, seasonal factors, flowers and leaves and cauliflower category of planting and harvesting the existence of the seasonal overlap, and then the same season or time period of growth and supply; and secondly, there may be strong correlation in the recipes and culinary requirements, so there is a strong correlation between other categories. There may be a strong correlation in terms of recipes and culinary requirements, hence a strong correlation between foliage and cauliflower, and the same can be said for the correlation between the other categories. This correlation analysis will inform the prediction when using wavelet neural network forecasting, i. e. the final predicted sales results should demonstrate similar category correlations.

3.2. Cost-plus ratio by category

In order to reduce the part of the markup rate due to the extreme value is too high, here take the average as the analysis, taking into account the commodity ordering price changes over time, through the time series analysis of the ordering price historical data, combined with the markup rate formula for the cost of the six categories of the markup rate were calculated by the value of the cost of the markup rate, so as to arrive at the cost of each of the six categories of the cost of the markup rate, and the results of the data are shown in the Table 1:

Table 1: Cost-plus ratio by category

Category	Cabbage	Philodendron	Capsicum	Eggplant	Edible mushroom	Aquatic rhizomes
Cost-plus ratio	89.33%	64.09%	30.28%	80.06%	141.49%	67.14%

Using the above model, the sales volume from July 1-7, 2023 is forecasted and combined with the merchandise attrition rate and the forecasted sales volume is calculated so as to arrive at the replenishment plan for each category on a daily basis during the period of 1-7, 2023, and the resultant data is shown in the Table 2:

Table 2: Daily replenishment plan per category, July 1-7,2023 (kg)

Date	Cabbage	Philodendron	Capsicum	Eggplant	Edible mushroom	Aquatic rhizomes
7.1	28.86	203.66	97.63	29.85	70.34	24.92
7.2	24.91	190.15	95.82	30.32	62.92	22.12
7.3	17.47	142.15	94.93	20.89	4591	16.97
7.4	17.47	156.82	94.70	18.74	45.88	17.20
7.5	17.69	152.93	94.70	19.02	50.94	18.49
7.6	19.06	156.35	94.73	18.93	49.38	20.40
7.7	20.53	169.54	94.75	21.74	55.27	21.16

The forecasting model predicts the daily wholesale price per category for the vegetable category of goods, using the pricing method of cost-plus pricing, combined with the cost-plus rate per category obtained from the previous calculations to arrive at the average daily pricing per category for July 1-7, 2023, with the resultant data shown in the Table 3:

Table 3: Average pricing per category per day, July 1-7, 2023 (￥/kg)

Date	Cabbage	Philodendron	Capsicum	Eggplant	Edible mushroom	Aquatic rhizomes
7.1	14.01	6.74	8.13	6.93	12.67	19.64
7.2	14.74	6.87	8.37	7.38	12.88	19.77
7.3	15.03	6.96	8.52	7.57	13.19	20.15
7.4	13.94	6.63	8.03	6.87	12.51	19.47
7.5	14.13	6.85	8.23	7.36	12.83	19.73
7.6	14.22	6.91	8.30	7.32	12.72	19.69
7.7	13.97	6.69	8.09	7.16	12.59	19.62

4. Conclusions and Outlooks

Vegetable goods have a short shelf life, and most varieties cannot be sold again on the next day. With the help of data modeling to develop commodity pricing replenishment strategy can provide timely and reliable reference for merchants. This paper uses the cost-plus pricing method and analyzes the sales data and ordering data of the same period of time in the historical time series to obtain the cost-plus rate of each category, and based on the historical data, uses the time series forecast to obtain the sales volume prediction data from July 1-7, and at the same time, considers the vegetable loss rate to establish the replenishment model, and in the replenishment, considers the sales volume as well as the loss rate of the replenishment accordingly, and establishes a model of profit maximization for the hypermarkets. Profit maximization model. At the same time, the wavelet neural model is used to consider the interrelationship between the categories to predict the future sales, the experimental results show that the wavelet neural network model has good predictability and robustness, and the prediction results are applied to the replenishment pricing model to calculate the results in line with the actual demand, which has a certain value of practical application.

The model for sales trend, demand analysis and replenishment decision optimisation for vegetables can be extended to other similar fresh food superstores or retail industries, such as fresh food superstores for fruits and seafood, and retail industries such as supermarkets and convenience stores, in order to help real-time monitoring of sales, forecasting of demand, and adjusting of replenishment strategies, etc. , and to improve the efficiency and accuracy of superstore stocking. This model needs to be adjusted and adapted to the specific industry, and applied in conjunction with actual data and business requirements.

References

[1] Kayikci Y, Demir S, Mangla S K, et al. Data-driven optimal dynamic pricing strategy for reducing perishable food waste at retailers[J]. Journal of cleaner production, 2022, 344: 131068.

[2] Isa F M, Munirah Ariffin W N, Jusoh M S. Shelf space allocation problem (SSAP) in the retail industry: a systematic literature review[J]. International Journal of Business & Society, 2024, 25(3).

[3] Li H, Liu J, Qiu J, et al. ARIMA-driven vegetable pricing and restocking strategy for dual optimization of freshness and profitability in supermarket perishables[J]. Sustainability, 2024, 16(10): 4071.

[4] Liu Y W, Li M R, Pu Y P. Replenishment and Pricing Strategies for Vegetable Commodities Based on Optimization Class Models[J]. Academic Journal of Business & Management, 2023, 5(26): 164-171.

[5] Selvakumar S, Renugadevi G, Vinishah N, et al. Sales Forecasting Based on Time Series Analysis[C]// 2024 International Conference on Science Technology Engineering and Management (ICSTEM). IEEE, 2024 : 1-7.

[6] Liu J, Li P, Tang X, et al. Research on improved convolutional wavelet neural network[J]. Scientific Reports, 2021, 11(1): 17941.

[7] Qian J, Wu A, Ji J, et al. Wavelet-optimized LSTM-ARMA model for nonlinear time series prediction in geotechnical engineering[J]. Journal of Tongji University (Natural Science), 2021, 49(8): 1107-1115.

[8] Cai X. Research on profit maximization of community fresh food stores under nonlinear inventory holding costs[D]. Hefei: Hefei University of Technology, 2020.

[9] Paul R K, Garai S. Performance comparison of wavelets-based machine learning technique for forecasting agricultural commodity prices[J]. Soft Computing, 2021, 25(20): 12857-12873.

[10] Liu K, Cheng J, Yi J. Copper price forecasted by hybrid neural network with Bayesian Optimization and wavelet transform[J]. Resources Policy, 2022, 75: 102520.