WEP
Warwick
Evans
Publishing

Transactions on Economics, Business and Management Research
ISSN: 2960-1789, eISSN: 2960-2254 | Volume 11
EMFT 2024

Automatic Pricing and Replenishment Decision for Vegetable Commodities Based on Support Vector Regression

Yiqi Liu [1, *], Siyu Yi [2], Boxu Song [3]

[1] School of science, China Three Gorges University, Yichang, China, 443002

[2] School of Economics and Management, China Three Gorges University, Yichang, China, 443002

[3] School of Electrical Engineering and New Energy, China Three Gorges University, Yichang, China, 443002

* Corresponding Author Email: 19071170211@163.com

Abstract. Automated pricing and replenishment decisions for vegetable merchandise play a critical role in the retail industry. Accurate pricing and replenishment forecasts can effectively improve operational efficiency, reduce inventory costs, and increase customer satisfaction. In order to accurately realize the automatic pricing and replenishment decisions, paper use the recent 30-day data to forecast the future based on the cost-plus pricing method, taking into account the seasonality of vegetable commodities, and at the same time, taking into account the cyclicality of vegetable commodities, paper use the sales volume of the previous year in the same month to fit the sales volume of the coming week, and combine the weighting of ARIMA time-series forecasting results to get the final forecasted sales volume, and employ the SUPPORT VECTOR REGRESSION theory to calculate the pricing, using the loss rate to calculate the replenishment, and finally constructed the automatic pricing and replenishment decision model for vegetable commodities applicable to big data, and the results show that the choice of cost-plus ratio is relatively stable, and the standard deviation of the sample is 0.24, which can help hypermarkets to achieve better decision-making.

Keywords: Support Vector Regression, Prediction Model, Replenishment Decision.

1. Introduction

In order to increase the profitability of a superstore, it is particularly important to rationalise the replenishment and pricing strategy for the vegetable category. The relationship between the total amount of vegetables sold and the cost-plus pricing has a direct impact on the supermarket's profitability. The total amount of replenishment and pricing strategy for the coming week will determine whether the supermarket can meet the sales demand while reducing the inventory backlog and maximising revenue. By deeply analysing the relationship between sales and pricing of vegetable categories, supermarkets can more accurately formulate sales strategies and improve operational efficiency. At the same time, scientific replenishment and pricing strategies can not only reduce inventory backlog, but also improve customer satisfaction and ultimately enhance the market competitiveness of superstores. In order to cope with these problems, this paper adopts SUPPORT VECTOR REGRESSION (SVR) analysis in order to determine the relationship between total sales and total cost of each category of vegetables. The total daily replenishment and pricing strategy for each vegetable category for the coming week are given. Time forecasting and grey forecasting were used to predict the total replenishment amount by weighting the daily sales volume for the coming week in conjunction with the daily sales volume of previous years.

In past studies, many scholars have explored the importance of inventory management and pricing strategies in the retail industry.Chiru C G and Posea V V (2018) [1] suggested that by analysing historical sales data, future demand can be predicted more accurately, thus optimising the inventory management.Aalto H (2019) [2], on the other hand, emphasised the wide application of cost-plus pricing method in the retail industry , and states that a rational pricing strategy can significantly improve profitability. In addition, Fan G F (2020) [3] and others have shown that SUPPORT VECTOR REGRESSION is well applied in sales data forecasting and can effectively capture the

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

complex relationship between sales volume and multiple influencing factors. These studies provide theoretical support and methodological reference for this paper.

Through this study, it aims to address the replenishment and pricing strategies of superstores in vegetable category sales. By analysing the relationship between total sales volume and cost-plus pricing and combining it with a time prediction model, the total replenishment volume and pricing strategy for the coming week are formulated to maximise the hypermarket's revenue. This will help the superstore gain an edge in the competitive market and improve overall operational efficiency and customer satisfaction.

2. Data Source

Data from ChinaUndergraduate Mathematical Contest in Modelling 2023.

3. Support vector regression and arima prediction

3.1. Support Vector Regression

For a general regression problem, assume that given the training sample $D=\{(x_1,y_1),(x_2,y_2)\text{L }(x_n,y_n)\}$. Paper want a function $f(x)$, so that the predicted value $f(x)$. As close to the true value as possible y. In support vector regression, paper introduce a tolerance ε, permissible $f(x)$ and y have a maximum of misalignment of ε.

In the SVR model, the loss is computed when, and only when, the absolute value of the difference with y is greater than ε[4]. That is, paper construct an interval band of width 2ε for the centre. This interval band represents the prediction we can tolerate the error range. If the true value y of the training sample falls within this interval band, the prediction is considered correct. The purpose of this is to treat the prediction error of the training samples asymmetrically in order to better accommodate the error sensitivity in different directions in the real problem. Noise and outliers in the regression problem can be dealt with to some extent in this way, as shown below in Figure 1:

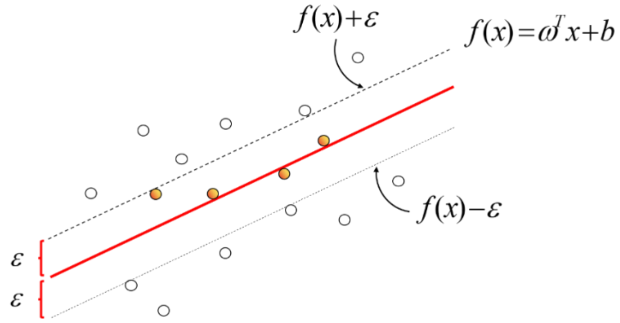

Figure 1. SUPPORT VECTOR MACHINE DIAGRAM

Support vector machine (SVR) has many advantages in solving small sample, nonlinear identification, its goal is to get the optimal solution under the available information for the small sample situation, in addition, the final solution obtained by the support vector machine is the global optimal solution, and its algorithmic complexity is independent of the number of sample dimensions, which is very suitable for dealing with nonlinear problems and has a very good generalisation ability. According to the data structure, there are 36 daily and 6 other major types of vegetables, and the two are nonlinear relationships, SVR has a strong theoretical basis and feasibility for solving the regression problem in this question.

From the theoretical analysis above, the SVR problem is transformed into the following objective function:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} l_c(f(x_i) - y_i) \tag{1}$$

After analysing the principle above, it can be found that the larger the isolation band, the greater the number of samples with a loss of zero.

So the penalty coefficient is introduced and the slack variable on its basis makes the expression more concise, let:

$$l_\epsilon(f(x_i) - y_i) = \begin{cases} 0 & |f(x_i) - y_i| \le \epsilon \\ |f(x_i) - y_i| - \epsilon & |f(x_i) - y_i| \ge \epsilon \end{cases} \quad \xi_i + \bar{\xi}_i = l_\epsilon(f(x_i) - y_i) \tag{2}$$

where ξ_i denotes the relaxation factor and represents $|f(x_i) - y_i| - \epsilon$ the error. Rewrite the objective function as:

$$\begin{gathered} \min_{w,b,\xi_i,\bar{\xi}_i} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}(\xi_i + \bar{\xi}_i) \\ s.t.\begin{cases} f(x_i) - y_i \le \epsilon + \xi_i \\ y_i - f(x_i) \le \epsilon + \bar{\xi}_i \\ \bar{\xi}_i \ge 0, \xi_i \ge 0, i = 1, 2 \cdots m \end{cases} \end{gathered} \tag{3}$$

The upper side of the process has to fulfil the KKT condition at the same time, i.e:

$$\begin{cases} a_i(f(x_i) - y_i - \epsilon - \xi_i) = 0 \\ \hat{a}_i(y_i - f(x_i) - \epsilon - \hat{\xi}_i) = 0 \\ a_i\hat{a}_i = 0, \xi_i\hat{\xi}_i = 0 \\ (C - a_i)\xi_i = 0, (C - \hat{a}_i)\hat{\xi}_i = 0) \end{cases} \tag{4}$$

Finally, the relationship between the two SVRs can be derived as:

$$\begin{gathered} f(x) = \sum_{i=1}^{m}(\hat{a}_i - a_i)x_i^T x + b \\ b = y_i + \epsilon - \sum_{i=1}^{m}(\hat{a}_i - a_i)x_i^T x \end{gathered} \tag{5}$$

Based on the SVR regression, the relationship between total sales and cost-plus pricing can be further solved.

3.2. ARIMA prediction

In time series analysis, ensuring the smoothness of the data is a necessary condition before proceeding to the subsequent steps of time series analysis. For non-smooth time series, the statistical analysis will be limited by methodological and theoretical limitations, such that the time series is not meaningful for the future. So the data is differenced to ensure smoothness.

$ARIMA$ is determined by three important parameters (p,d,q) [5]. p is the regression coefficient, which indicates that the sequence values are lagged by order p; d corresponds to "I" in the model $ARIMA$, which indicates that the time series data becomes a smooth sequence, which requires at least d times of differencing; and q is the sliding average coefficient, which indicates that the error term is lagged by order q. The autoregressive sliding average model formula is as follows:

$$y_t = \mu + \sum_{i=1}^{p} r_i y_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{6}$$

Using the autocorrelation function:

$$\rho_k = \frac{\mathrm{cov}(y_t, y_{t+k})}{\sqrt{\mathrm{var}(y_t)\,\mathrm{var}(y_{t+k})}} \tag{7}$$

To reflect the correlation of neighbouring views in time-series data, while using the partial autocorrelation function:

$$\varphi_{kk} = \begin{cases} \rho_1 & \text{k=1} \\ \dfrac{\rho_k - \sum_{j=1}^{k=1} \varphi_{k-1,j}\varphi_{k-j}}{1 - \sum_{j=1}^{k=1} \varphi_{k-1,j}\varphi_{k-j}} & \text{k>1} \end{cases} \tag{8}$$

The time series data was subjected to autocorrelation plot smoothing test to check the truncation of ACF and PACF plots respectively, both are either trailing or truncated, otherwise they are non-smooth series. The non-smooth series are tested for smoothness after differencing until the smoothness condition is met, then the d-value is determined. After the smoothness of the sequence is determined, the parameters p, q can be determined according to the following table.

Table 1: ACF and PACF tables

model	ACF	PACF
AR(p)	trailing tail	P-order post-tail truncation
MA(q)	q-order post-truncation	trailing tail
ARMA (p.q)	q-step rear drag tail	P-step rear drag tail

The above model only takes into account the data of each vegetable variety in the last thirty days when forecasting, in order to tap into the cyclical sellability data, this paper uses the data of those seven days of July 21 and 22 daily, respectively, to calculate the share of the six vegetable categories, the specific formula is as follows:

$$p_i = \frac{n_i}{m_i} (i = 1, 2 \ldots 6) \tag{9}$$

n_i is the number of i-variety vegetables on 1-7 July 21, m_i is the number of i-variety vegetables on 1-7 July 22, and p_i is the percentage of i-variety vegetables.

After finding out the percentage of vegetable varieties, then weighted with the corresponding predicted sales volume to get the sales volume in the coming week, the corresponding calculation formula is as follows:

$$TS_{ij} = 0.5 * A_{ij} + 0.5 * \frac{m_i}{7} * P_i (j = 1, \cdots 7) \tag{10}$$

After calculating the daily sales of different varieties of vegetables, the corresponding cost-plus pricing rate is calculated based on the SVR, and the profit of each category of vegetables is calculated based on the cost-plus pricing rate, accordingly, the pricing strategy is formulated, the corresponding total replenishment, and the total amount of replenishment can be found based on the sales, increasing the total loss rate in Appendix IV. For the sake of simplicity, the flow chart is as follows. As shown below in Figure 2:

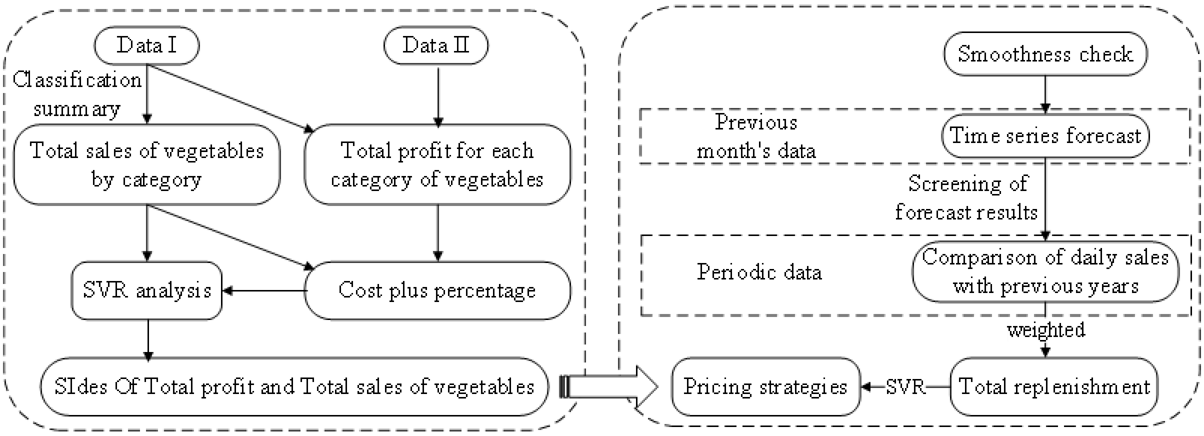

Figure 2. Idea flow chart

4. Results

4.1. Analysis of data

By analysing the data given in the question, it can be seen that there are no missing values in the data for this question and the features are already at similar scales[6]. Therefore, there is no need for missing value processing and data normalisation.

Due to the subsequent content involves correlation and time series and other aspects of the solution needs to retain all the data points, for the outlier data can not be directly eliminated, combined with the actual analysis of the outliers, if the outliers are consistent with the actual life of the accidental occurrence of the situation needs to be retained to participate in the subsequent calculations, the following list of abnormal outliers, analyse one by one whether it is reasonable.

Firstly, the outliers of sales volume are analysed. As shown below in Table 2:

Table 2: Enumeration of unusual outliers

Date of sale	product code	item name	sales (kg)
2022/6/9	102900011034354	Fresh Dumpling Leaves (bag)(1)	160
2023/6/16	106972776821582	Fresh Dumpling Leaves (bag)(3)	30
…			
2023/1/15	102900011021842	Honghu Lotus Root(Powdered Lotus Root)	17
2022/8/8	102900011001561	Lotus seedling (pcs)	17

Continued analysis of outlier anomalies, through the above chart, it can be seen that only some of the data in the existence of outliers, for outliers can not be directly removed. Part of the data in the above table as an example, in which the single sales of 160 kg and 30 kg of a single product to become fresh rice dumpling leaves, combined with the sales date of 9 June and 16 June. At this time coincides with the Dragon Boat Festival, the festival of fresh rice dumpling leaves use a large demand, in line with the actual demand. The rest of the sales are within 20 kg, although the outliers, due to the sample base is large, there is a high probability of special circumstances, so the above outliers are retained.

Secondly, the outliers of the sales unit price are analysed. As shown below in Table 3:

Table 3: Information table for anomalous outliers

Date of sale	product code	Item Name	Unit sales price
2023-03-13	102900005115199	Sichuan Red Toon	119.90yuan/kg
2023-03-13	102900005115199	Sichuan Red Toon	119.90yuan/kg
…			
2020-10-31	102900011012871	black-skinned macrolepiota	116.00yuan/kg
2020-11-12	102900011012871	black-skinned macrolepiota	116.00yuan/kg

The unit price of Sichuan red toon is RMB 119, which is much higher than the actual price of RMB 30, according to the information collected from the Internet. The selling price of black skinned chanterelle mushroom is 116 RMB which is similar to the selling price obtained from market statistics, so the selling price of black skinned chanterelle mushroom is retained and the selling price of Sichuan red toon is excluded.

After the removal of some outliers, it is found that there are differences in the number of missing values between different features, and this paper uses the KNNImputer class to fill in the missing values. This class uses machine learning techniques based on the KNN algorithm to impute missing values, and the similarity between "neighbours" is used to impute relevant missing values. We choose

Fill the missing samples by finding the K neighbours that are most similar to the missing samples. Use correlation between features to predict missing values, not just information about individual features. As shown below in Figure 3:

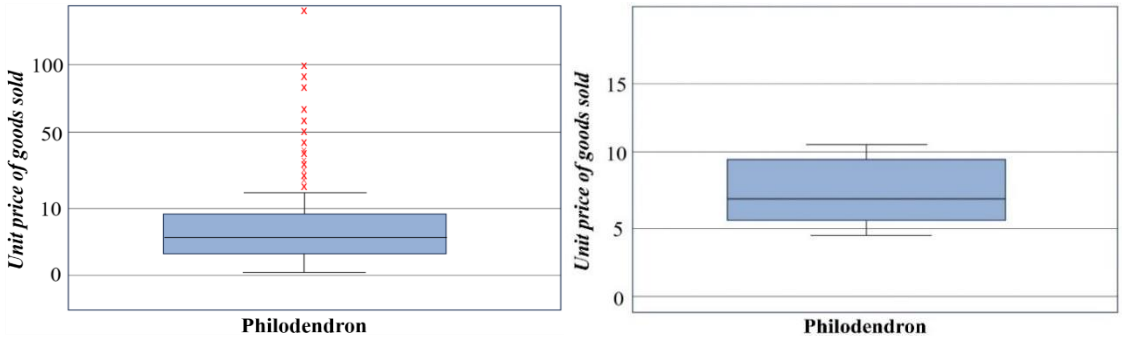

Figure 3: Box shape before and after outlier treatment

The outlier-processed box plots boxes are more centralised and the length of the whiskers is shortened, making the box plots more compact overall. The processed box plots may then better demonstrate the distribution of the data. Increasing the reliability of statistical inference may reduce the impact of outliers on statistical inference.

For the various types of vegetables in Annex II, the processing is done to get the total quantities of different categories of vegetables in different days, the total cost is obtained by multiplying the unit price of each type of vegetables in Annex III, the total turnover is calculated by using the selling price in Annex II, the total profit amount of each category of vegetables in different days is obtained by summing up the total profit amount of each category of vegetables in different days, and the corresponding cost-plus rate of each type of vegetables in each day is calculated by using the formula of cost-plus rate as The specific calculation formula is as follows:

$$\begin{gathered} S = \sum n_i P_{sale} \\ S_{profit} = \sum n_i (P_{sale} - P_{buy}) \\ a = \frac{S_{sale}}{S_{sale} - S_{profit}} - 1 \end{gathered} \tag{11}$$

Where n_i denotes the number of vegetables in day i, P_{sale} denotes the selling price of the corresponding vegetable, S_{sale} is the total sales, P_{buy} denotes the wholesale price of the corresponding vegetable, and S_{profit} denotes the total profit earned.

4.2. Descriptive statistics

The sales data of each category of vegetables, the same type of dishes are categorised, the total amount of sales volume in recent days are counted, and the distribution of sales volume of each category of vegetables is plotted as follows. As shown below in Figure 4:

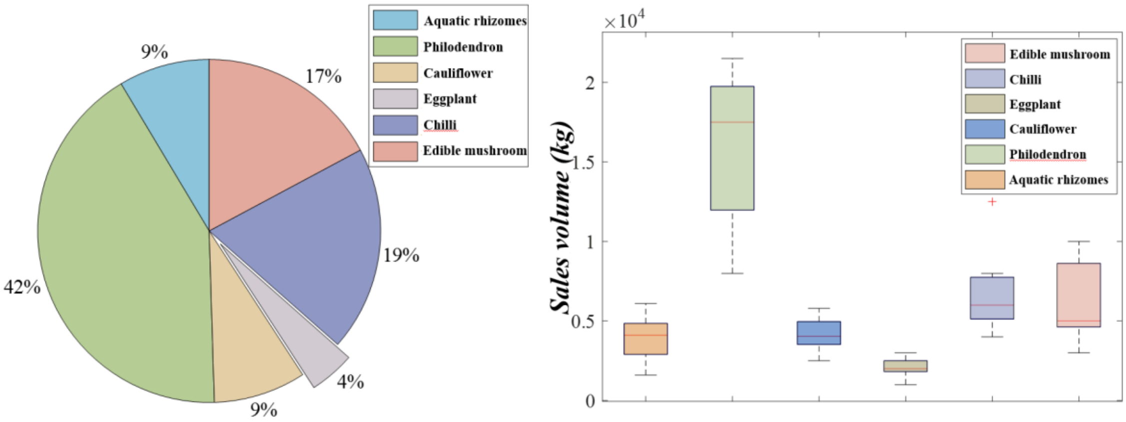

Figure 4: Distribution of Vegetable Sales by Category

Analysis of the graph shows that the sales volume of flowering and leafy vegetables is much higher than the sales volume of other vegetables, the sales volume of chilli and edible mushrooms are close to each other, and the sales volume of aquatic roots and cauliflowers are close to each other. The eggplant category has the least sales volume. The ratio of sales volume of each type of vegetables mentioned above can be met as much as possible in future purchases to ensure that the wholesale vegetables can satisfy the public's demand for vegetables and avoid waste as little as possible.

On the basis of the left chart on the annual sales volume of various types of vegetables, calculate the data of each statistical data to draw a box plot, by observing the box plot, the following two major pieces of information can be derived:

(1). The concentration trend of the data: the median is located in the middle of the box, indicating that the distribution of data is relatively symmetrical. The medians of most of the vegetable species in the chart on the right are located in the middle of the box. It can be judged that the distribution of the data is relatively symmetrical. There is a possibility to meet the normal distribution, the data can be further Shapiro-Wilk test to determine whether to meet the normal distribution.

(2). The distribution range of the data: the longest whiskers in the box plot of the foliage category indicate that the larger the range of the data, reflecting the characteristics of the seasonal supply of foliage vegetables. Different leafy and flowering vegetables have different growth cycles and suitable growing seasons. Therefore, the characteristics of seasonal supply would lead to a large variation in sales volume in different seasons. Seasonal changes in customer demand for leafy and flowering vegetables need to be taken into account in subsequent stocking to ensure maximum returns.

Secondly, the changes in the unit price of each category of vegetable sales were analysed, the changes in the average unit price of each category of vegetable sales over time were counted, and the trend chart of the changes in the unit price of vegetable sales was plotted as

follows. As shown below in Figure 5:

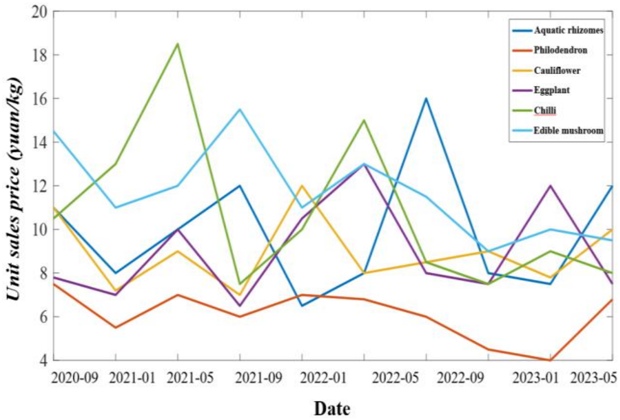

Figure 5: Trends in unit prices of vegetable sales

As can be seen from the line graph, the average selling price of cauliflower vegetables is maintained at a low level and does not fluctuate significantly, suggesting that cauliflower and leafy vegetables are not obviously seasonal and can be adequately supplied at all times of the year. Chilli class unit sales price by the seasonal impact of the amplitude of the ups and downs, indicating that these crops have seasonality, in a fixed season of high yield, the unit sales price will naturally decline, in the subsequent topic involves pricing and purchase of goods need special consideration.

4.3. The experimental results

Based on the definition of cost-plus pricing [7], the cost-plus pricing for each individual item sold is solved to obtain some of the results as shown in the table below, and the rest of the results are shown in the appendix:

Table 4: Partial cost-plus pricing table

Date of sale	single category	margins	costs	Cost-plus ratio
2020-07-01	Naga jolokia	2.95	4.649	63.47%
2020-07-01	Brassica pekinensis	0.50	2.701	18.45%
2020-07-01	Shanghai Qingdao	1.78	8.215	21.72%
2020-07-01	Chinese flowering cabbage	2.67	5.330	50.09%

Forecasting the daily sales for the coming week through the time series, weighted with the daily sales of previous years, and combining the total sales of each category of vegetables with the time of day and the quarter, the final forecasted total sales per day are obtained as shown in the table below:

Table 5: Projected total daily sales

date	philodendron	cauliflower	aquatic rhizomes	eggplant	chilli	edible mushroom
7.1	152.937	19.793	20.422	23.277	95.039	64.139
7.2	110.988	14.364	14.821	16.892	68.971	39.289
7.3	96.132	12.441	12.837	14.631	59.739	34.030
7.4	113.610	14.703	15.171	17.291	70.600	40.217
7.5	124.971	16.174	16.688	19.021	77.660	44.239
7.6	138.954	17.984	18.555	21.149	86.35	39.108
7.7	135.458	17.631	18.088	20.617	84.177	47.952

Combined with the relationship between time and cost-plus ratio, based on the support vector machine regression analysis obtained on the relationship equation between cost-plus ratio and cycle time, to find the predicted daily sales volume to set the pricing strategy as shown in the table below.

Table 6: Projected daily sales volume setting pricing strategy table

date	philodendron	cauliflower	aquatic rhizomes	eggplant	chilli	edible mushroom
7.1	0.641	0.488	0.491	0.573	0.644	0.584
7.2	0.646	0.512	0.524	0.573	0.601	0.598
7.3	0.659	0.522	0.537	0.576	0.596	0.598
7.4	0.645	0.510	0.522	0.572	0.602	0.598
7.5	0.640	0.503	0.512	0.591	0.609	0.596
7.6	0.638	0.494	0.501	0.581	0.620	0.591
7.7	0.638	0.496	0.503	0.591	0.616	0.592

The total profit obtained was $3,426.78, and the specific daily profit obtained is shown in the table below:

Table 7: Profit statement at forecast date

date	7.1	7.2	7.3	7.4	7.5	7.6	7.7
Daily profit ($)	489.54	467.43	498.52	492.72	487.63	489.65	492.73

5. Validation of results

5.1. Analysis of the degree of correlation between total sales and pricing cost-plus ratios

Based on the various types of vegetables found above, a line graph of the daily additive ratio data is plotted in Figure 6:

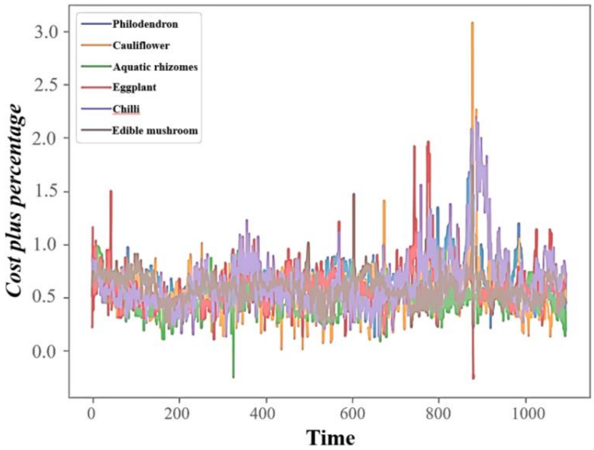

Figure 6: Line plotting of additive ratio data

Looking at the total category cost-plus ratios, despite the presence of individual outliers, the overall mean is 0.583 and the sample standard deviation is 0.204, showing a small sample variance, indicating a relatively stable choice of pricing cost-plus ratios[8]. Time variation has no significant effect on the cost mark-up ratio.

The relationship between the total sales volume of each category of vegetables and time has been calculated in Question 1, and it can be observed through the line graph that the sales volume of each category of vegetables is greatly affected by seasonality and presents cyclical changes, so the sales volume is more significantly affected by time. Further validation is required to verify the correlation between sales volume and cost-plus ratio for each category of vegetables.

In order to further determine whether there is a significant effect between the cost mark-up percentage and the total sales volume of vegetable categories, support vector machine regression analysis was used. Correlation coefficients were calculated for the sales volume and cost-plus percentage for each category and are presented in the table below:

Table 8: Table of correlation coefficients

philodendron	cauliflower	aquatic rhizomes	eggplant	chilli	edible mushroom
0.185	0.136	-0.198	0.057	0.272	-0.090

It is concluded that total sales are less linearly correlated with the pricing cost-plus ratio.

5.2. Cost-plus ratio SVR regression fitting analysis

The SVR regression fitting algorithm was performed on all categories of data to obtain the non-parametric relationship curves as shown below, As shown below in Figure 7:

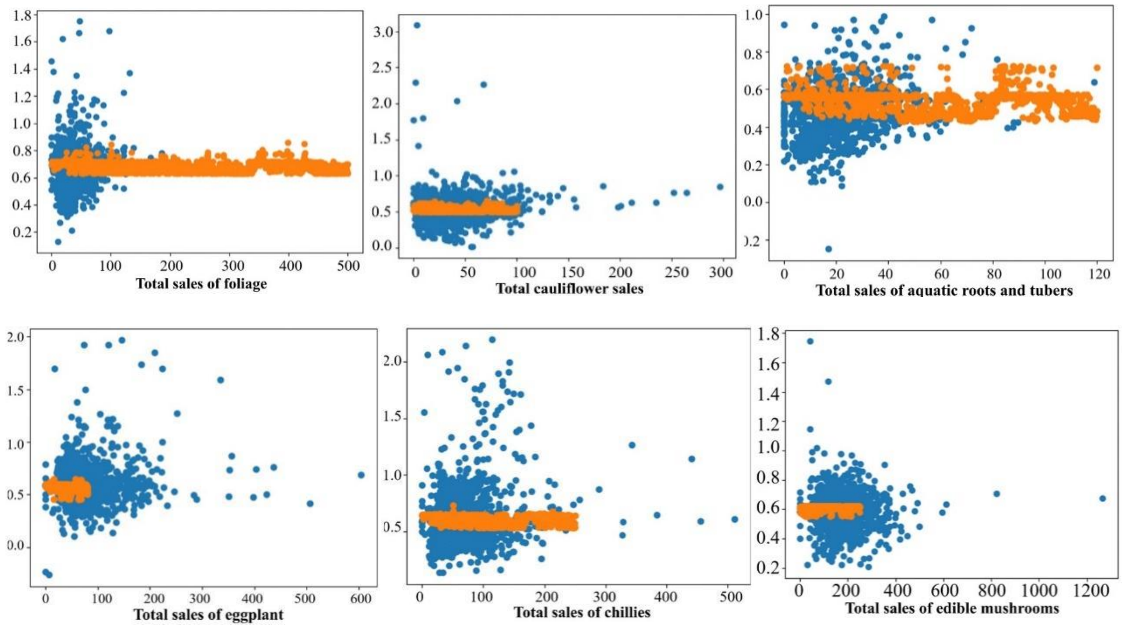

Figure 7: Plot of cost-plus ratio fitting results

Based on the observations in the above graph, the cost-plus ratio fitting results show that they basically fluctuate within the mean value range, which is significantly affected by time and other factors[9]. In order to predict the total daily replenishment and pricing strategy for each vegetable category in the coming week, we first need to examine the sales volume and cost-plus ratio during that time period.

5.3. Analysis of the reasonableness of forecast data

In this paper, using the most recent 30 days of data as a training set, time series forecasting of future sales is performed separately.

The SVR regression fitting analysis shows that the dependent variable in this question is indeed mainly affected by time variation[10].

The time series forecasting method is based on past time series data, and by building models and analysing the trend of time, seasonality, etc., it is able to better capture the pattern of data changes over time. The model selection is professionally matched.

Considering that some vegetable categories have seasonal characteristics, in order to better take into account the current period, this study adopts the prediction of daily sales volume in the coming week and takes into account the sales volume of the past years in order to arrive at a comprehensive and reasonable total amount of replenishment. For the prediction results a line graph is plotted as follows. As shown below in Figure 8:

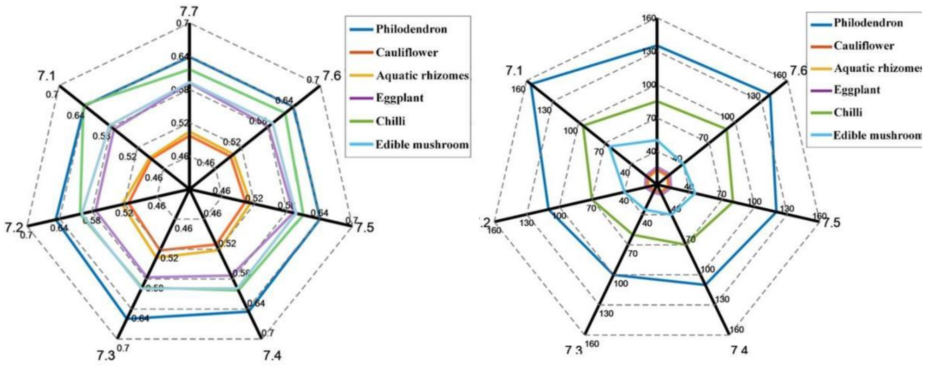

Figure 8: Line graph of cost-plus ratios

For the predicted value of the cost plus rate, it can be clearly seen that the line graph of each category fluctuates above and below a fixed value, and the average value is similar to the results obtained from the SVR regression fitting analysis above, which indicates that the results are accurate.

Secondly, these graphs of the predicted sales volume are analysed, in which the sales volume of flower and leaf products is significantly higher than that of other products, which is in line with the conclusion that the sales distribution of flower and leaf category ranks first in the sales volume statistics and far exceeds that of other categories, indicating the accuracy of the results[11].

6. Conclusions

For the problem of automatic pricing and replenishment decision making for vegetable commodities, firstly, the profit and cost of each category of vegetables are collated and used to calculate the cost-plus ratio, correlation analysis is carried out with the total number of sales, and it is found that the correlation is not strong, and the support vector machine regression is used to carry out the non-linear fitting. Considering seasonality, the most recent 30 days of data were used to make predictions for the future, considering cyclicality, the sales volume of different years in the same month were used to fit the sales volume for the week ahead, the two were weighted to get the final predicted sales volume, and then the pricing was calculated using the Support Vector Machine fitted curve to calculate the final total revenue. The replenishment is calculated using the attrition rate, and the experimental results show that the support vector machine regression has good predictability and robustness, and has some practical application value.

Acknowledgements

This work was supported by the Innovation and Entrepreneurship Training Programme for University Students (Walking with "Twilight" - A New Platform for Senior Dating with Emotional Connections)

References

[1] Chiru C G, Posea V V. Time series analysis for sales prediction [C]//Artificial Intelligence: Methodology, Systems, and Applications: 18th International Conference, AIMSA 2018, Varna, Bulgaria, September 12–14, 2018, Proceedings 18. Springer International Publishing, 2018: 163-172.

[2] Aalto H. Competition-Based Dynamic Pricing in E-Commerce [D]., 2019.

[3] Fan G F, Guo Y H, Zheng J M, et al. A generalized regression model based on hybrid empirical mode decomposition and support vector regression with back‐propagation neural network for mid‐short‐term load forecasting [J]. Journal of Forecasting, 2020, 39(5): 737-756.

[4] Zhang F, O'Donnell L J. Support vector regression [M]//Machine learning. Academic Press, 2020: 123-140.

[5] Benvenuto D, Giovanetti M, Vassallo L, et al. Application of the ARIMA model on the COVID-2019 epidemic dataset[J]. Data in brief, 2020, 29: 105340.

[6] Peck R, Short T, Olsen C. Introduction to statistics and data analysis [M]. Cengage Learning, 2020.

[7] Barusman A R P, Yuliana T M, Mirfazli E. Analysis of implementation cost plus pricing method in the decision on the determination of product sales Prices [J]. Intern J Adv Sci Techn, 2020, 29(06): 1832-1838.

[8] Tiara T, Maulidia Y, Kantun S. The Pricing Evaluation Using the Cost-Plus Pricing Method at CV Macarindo Berkah Group in Jember District [J]. RIGGS: Journal of Artificial Intelligence and Digital Business, 2023, 2(1): 6-11.

[9] Wang Y G, Wu J, Hu Z H, et al. A new algorithm for support vector regression with automatic selection of hyperparameters [J]. Pattern Recognition, 2023, 133: 108989.

[10] Zhang T, Zhang Z, Chi Y. Research on Vegetable Replenishment Strategy Based on Time Series Analysis [J]. Information Systems and Economics, 2024, 5(1): 74-81.

[11] Wang W, Luo Y, Yang X. Automated pricing and replenishment strategies for vegetable products based on SVR and simulated annealing [J]. International Journal of Global Economics and Management, 2024, 2(3): 1-10.