Financial Engineering and Risk Management (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/ferm.2023.061128
ISSN 2523-2576 Vol. 6 Num. 11

A Study on the Supermarket Replenishment and Pricing Strategies Based on SARIMA Model

Qifan Wang[1, *]

[1]*College of Science, China University of Petroleum (East China), Qingdao, Shandong, 266580, China*

[*]*Corresponding author: wangqifan239@foxmail.com*

Keywords: Vegetable Products, Merchandise Replenishment, Merchandise Pricing, SARIMA Forecasting, Cost Plus Pricing Method

Abstract: In the realm of fresh produce supermarkets, the expedited turnover of vegetable products necessitates routine replenishment and pricing adjustments by supermarkets. Nonetheless, formulating effective strategies for these aspects presents a substantial challenge to merchants. To address this issue, this study employs the SARIMA prediction model to anticipate future replenishment volumes and pricing for vegetable products in supermarkets. The cost-plus pricing method serves as the foundation for pricing in this predictive analysis. The predictive findings highlight a high accuracy in foreseeing both the upcoming month's replenishment volume and pricing. These forecasts reveal a fluctuating trend, displaying consistent periodicity akin to previous years, indicative of a robust predictive capacity. Consequently, this study culminates in devising a supermarket replenishment and pricing strategy grounded in the SARIMA model. This strategy aims to facilitate improved planning of future restocking and pricing by merchants, thereby fostering enhanced sales of vegetable products within supermarkets.

1. Introduction

Since its inception, the supermarket industry has evolved into a pivotal component of modern retail. The strategies governing replenishment and pricing of supermarket goods significantly impact profitability, market competitiveness, and consumer behavior. Rational approaches to replenishment and pricing can effectively bolster supermarket sales, profits, and market position. Hence, the exploration of these strategies holds critical significance and purpose.

The increasing importance of replenishment and pricing strategies in operations has garnered attention from numerous researchers and scholars. Lin Feng et al. proposed a basic method aiming to maximize total profit, presuming the demand curve's dependence on unit price, displayed quantity, and sale date, resulting in an equation for pricing, albeit with a limitation in precision[1]. Subsequently, more sophisticated algorithms, such as machine learning and deep learning methods, were employed by others to predict agricultural product pricing and replenishment[2-4], offering a notable advantage in accuracy. Notably, machine learning's predictive accuracy surpasses statistical models but lags behind deep learning[5-6]. For instance, Lobna Nassar et al. utilized the LSTM model within deep learning to address periodicity issues. Meanwhile, in machine learning, techniques like ridge regression, random forest, and support vector machines gained traction. Enhanced computational accuracy in machine learning, integrating factors like weather and humidity, is achievable, although this increases computational complexity[7]. Despite the advancements proposed by scholars, the intricate computational processes present challenges, limiting the effective application of models in addressing supermarket replenishment and pricing problems.

The SARIMA model, a statistical approach, accounts for periodicity while ensuring prediction accuracy, employing a simpler calculation process with fewer parameters. This renders it a suitable method for supermarkets to forecast goods replenishment and pricing. To address this challenge, this paper endeavors to formulate a supermarket replenishment and pricing strategy by constructing a SARIMA model[10]. The aim is to predict daily sales and pricing for the upcoming month, validating the model's accuracy by comparing it with historical transaction data from the same period, assuming the input price derived from the cost-plus pricing method[8-9].

2. SARIMA Principles and establishment of the model

2.1 The Principle of the SARIMA model

The SARIMA (Seasonal Autoregressive Integrated Moving Average) model is specifically designed for time series forecasting, offering accurate predictions by effectively modeling both trends and seasonal patterns within the data. Its parameters encompass autoregressive order, difference order, and moving average order, along with seasonal autoregressive order, seasonal difference order, seasonal moving average order, and seasonal period. Through a process of differencing and fitting the time series data, the SARIMA model effectively estimates these parameters, utilizing them to forecast future values.

The autoregressive component of the model leverages past time point values to predict current values, while the moving average component utilizes past time point errors for current value predictions. Simultaneously, the differencing element aims to eliminate trends and seasonal variations inherent in the data.

2.2 SARIMA Model building

(1) cost-plus pricing model

References[8-9] present the modeling for the cost-plus pricing method as follows:

$$X = C(1 + w) \tag{1}$$

Where X represents the price, C denotes the average cost, and w signifies the cost markup rate. This model is particularly applicable for pricing products that have comprehensive historical cost information.

$$X = C + Tr \tag{2}$$

Here, T represents the total capital investment and r denotes the return on investment. This model is primarily employed in pricing strategies for new products lacking historical cost information, relying solely on project decision data.

$$X = (V + F + L)/Y = b + a + L/Y \tag{3}$$

Where V stands for variable cost, b indicates unit variable cost, F denotes fixed cost, a indicates unit fixed cost, L represents target profit, and Y refers to business volume (i.e. sales volume, which constitutes 80% of production capacity).The cost-plus pricing method serves as the fundamental basis for pricing vegetable commodities, aiding in a comprehensive understanding and establishment of costs, pricing, and profit.

(2) SARIMA model

The SARIMA model was developed using Python with the fundamental structure:

$$ARIMA(p,d,q)(P,D,Q)s \tag{4}$$

In this equation, d represents the number of non-seasonal differences, p and q denote autoregressive and moving average orders, respectively, while P and Q signify seasonal autoregressive and moving average orders. D stands for the number of seasonal differences, and 's' represents the seasonal difference interval. The SARIMA model is constructed based on the principles of the cost-plus pricing method.

Drawing upon the introduction and application of the SARIMA model in literature[10], we employ the first pricing model (2) outlined in the following formula:

$$X = t / \sum_{i=1}^{n} a_i \tag{5}$$

$$C = \frac{\sum_{i=1}^{n} \frac{a_i}{1-b_i} c_i}{\sum_{i=1}^{n} a_i} \tag{6}$$

By integrating equations (5) and (6), the solution is obtained as follows:

$$t = \sum_{i=1}^{n} \frac{a_i}{1-b_i} c_i (1+w) \tag{7}$$

3. Results

3.1 SARIMA Solution of the model

The data source comprises multiple Excel files obtained from the Chinese Bureau of Statistics, encompassing data on total sales of vegetable products spanning from 2020 to 2023. The preprocessed data was incorporated into the cost-plus pricing model. Through Python programming, the partial results for each individual product were generated and visualized in Table 1:

Table 1: Mark-up rates for selected vegetable commodities

Item Name	Mark-on percentage	Item Name	Mark-on percentage
Shanghai blue	44.67%	Sea Mushrooms	53.02%
Chinese mustard	60.72%	Luffa tip	11.06%
Shepherd's purse	17.17%	Little cabbage	21.22%
Purple eggplant	42.50%	Wuhu green pepper	27.20%
Fresh black fungus	72.26%	Day lily	-5.43%
Big arabidopsis	-18.92%	Round eggplant	-23.92%

Upon acquiring the markup rate for each product, a sequence of data predictions was initiated to forecast the pricing and replenishment of vegetable products.

For an accurate projection of vegetable product sales data in the upcoming month, an initial analysis of its periodicity was imperative. Leveraging the periodicity analysis tool within the SARIMA model, an examination of the periodic trends in total vegetable sales spanning from 2020 to 2023 was conducted. The findings from this analysis are visually presented in Figure 1:

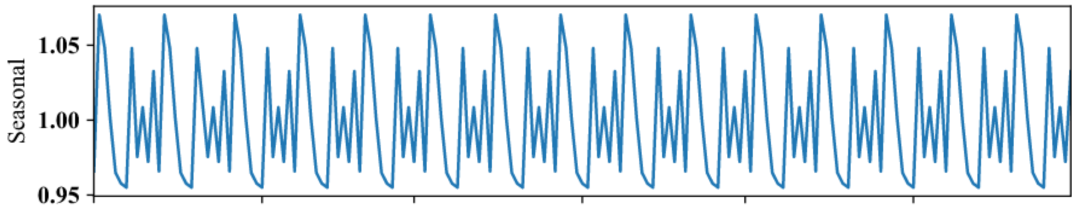

Figure 1: Monthly cyclical analysis for 2023

Observing Figure 1, it becomes apparent that gross sales exhibit a pronounced cyclic pattern, displaying consistent fluctuations from January 2023 to June 2023. These cycles showcase recurring shifts from peaks to troughs within each month. Consequently, the SARIMA model's periodicity is well-suited for sales forecasting, allowing for predictions based on the periodicity parameter.

Upon confirming the conformity of these periodicities with anticipated patterns, we proceeded to employ the SARIMA model for total sales volume forecasting. The training data spanned from 2020 to June 2023, employing a 12-month period. The resulting forecasts are outlined below:

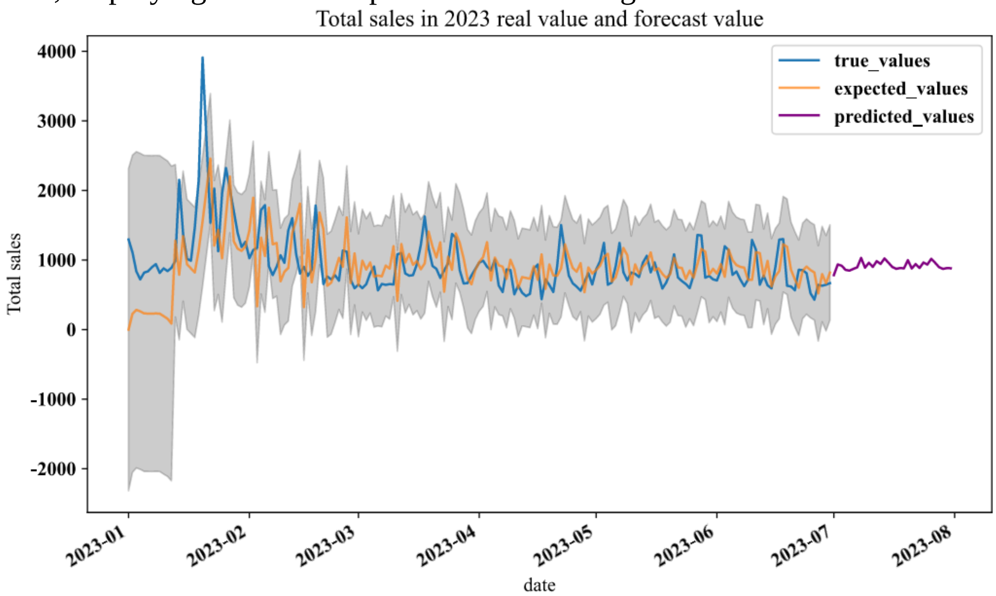

Figure 2: Real and forecast curves for total sales, July 2023

Referring to Figure 2, we can derive the true value, the expected value, and the predicted value for July 2023 by observing data trends from January 2023 to June 2023. In January, the fitting effect between the expected value and the true value is notably poor, with a substantial error interval. However, as subsequent fitting progresses, the expected value gradually converges toward the true value starting from February, accompanied by a reduction in the error interval. By March and April, the expected value aligns closer with the true value, although periodicity identification remains challenging. Notably, from April to June, the expected value demonstrates the most accurate fitting with the true value, reflecting a consistent cyclic pattern evident in the alignment of peaks and troughs within the wave. Consequently, the accuracy of predicted values for July exhibits improvement.

Following a similar forecasting methodology employed for total sales, based on sales data spanning from 2020 to June 2023, I utilized the SARIMA model with identical model parameters to forecast sales data for July 2023. The outcomes are depicted in Figure 3:

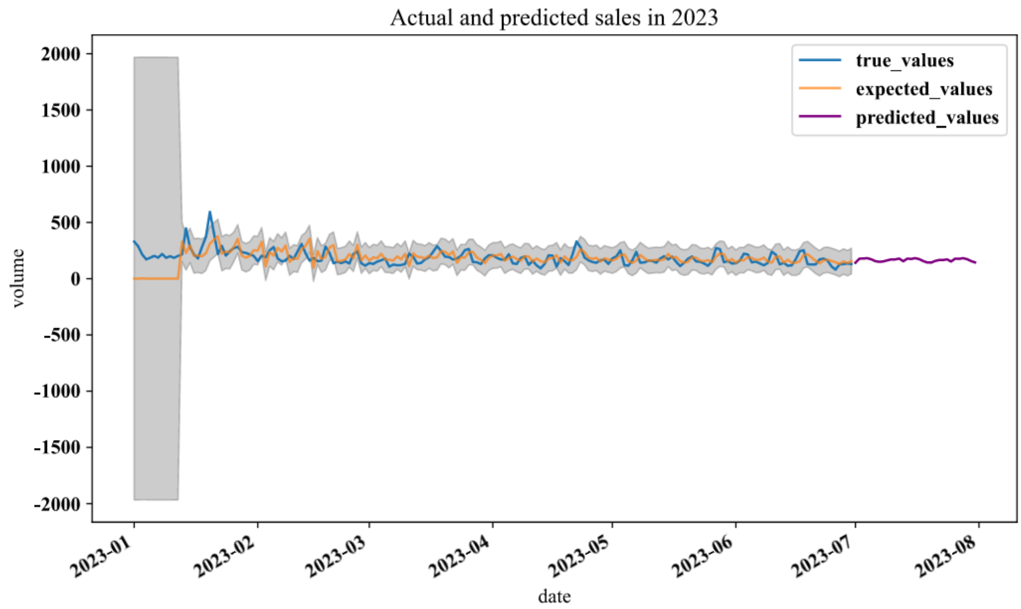

Figure 3: Real and forecasted sales volume curves for July 2023

Upon reviewing Figure 3, an overall improvement in the fit is evident. In January, due to inherent systematic errors within the model, there is a considerable error interval resulting in poor fitting. However, progressing to February, the fitting between the expected and true values notably improves, accompanied by a gradual reduction in the error interval. Nevertheless, the periodicity of the fit remains inadequate during this period, leading to inconsistency in the alignment of peaks and troughs. Contrastingly, from March to June, the expected value demonstrates superior alignment with the true value, exhibiting a minimal error interval. Notably, the periodicity performs exceptionally well during this phase, displaying a high degree of consistency in the peaks and troughs across the curve. Consequently, the accuracy of predicted data for July showcases a significant improvement.

3.2 Analysis of experimental results

Based on the known projected gross sales and sales volume, the ultimate projected sales volume and pricing are depicted in Figures 4 and 5:

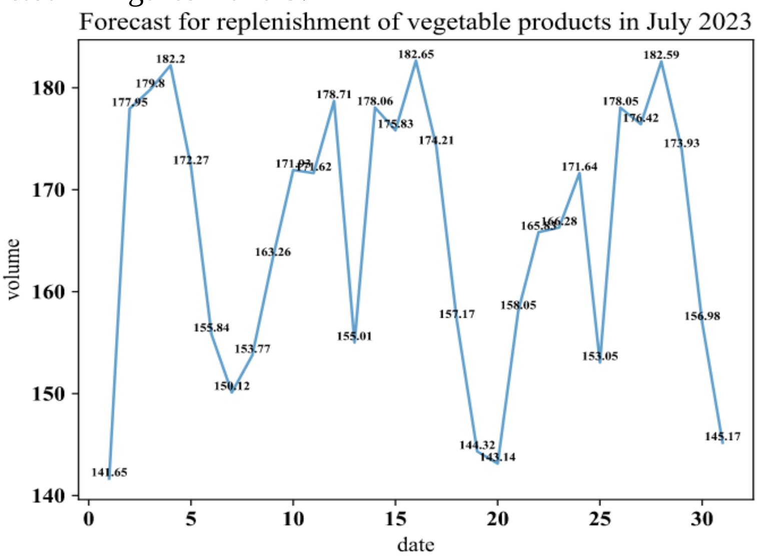

Figure 4: July 2023 replenishment forecast curve

The results indicate a noticeable fluctuation in daily replenishment during the month of July, primarily influenced by cyclicality. Considering practical circumstances, increased replenishment over a certain period results in surplus fresh products, leading to stagnant inventory and reduced replenishment in subsequent days. Subsequently, due to the perishable nature of the products, these items either sell or expire, surpassing their shelf life within a few additional days. Consequently, a renewed surge in replenishment initiates the cycle once again. This cyclic pattern contributes to an erratic up-and-down cycle in monthly sales volume.

Furthermore, each low point consistently hovers around 150 units, while each high point remains around 180 units. This similarity in purchase volume at each cycle's peak and the uniformity in cycle duration, approximately 10 days, substantiates the highly cyclical nature of vegetable product sales depicted in this graph.

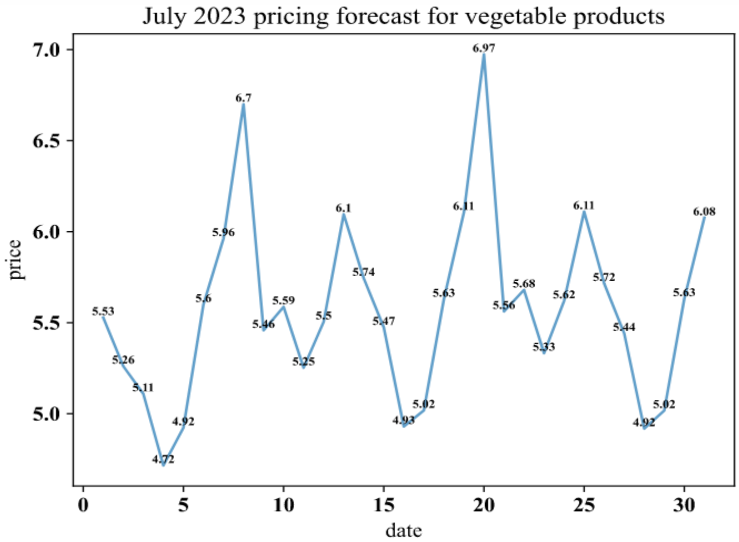

Figure 5: July 2023 pricing forecast curve

Figure 5 demonstrates a parallel cycle in the prediction of pricing and replenishment volume, albeit exhibiting contrasting trends in their peaks and troughs. Notably, as the predicted replenishment volume increases, the predicted pricing value decreases correspondingly. Conversely, a decrease in the predicted replenishment volume results in an increase in the predicted pricing value. This inverse relationship primarily stems from the relatively consistent total daily supermarket sales within a month, despite the constant flux in sales volume due to the aforementioned objective factors. Hence, merchants often adjust prices accordingly to facilitate sales: reducing prices during periods of significant replenishment to encourage sales, and conversely, increasing prices during smaller replenishment periods to ensure sales. As a result, the prediction results mirror the pattern depicted in the figure.

Following the forecast results for July, verifying the accuracy of the obtained data involves comparing the data for July 2023 with data from previous years. Similar trends in the comparison results substantiate the accuracy of our findings. Hence, data from July 2020, July 2021, July 2022, and July 2023 are concurrently plotted to illustrate the comparison, as depicted in Figure 6:

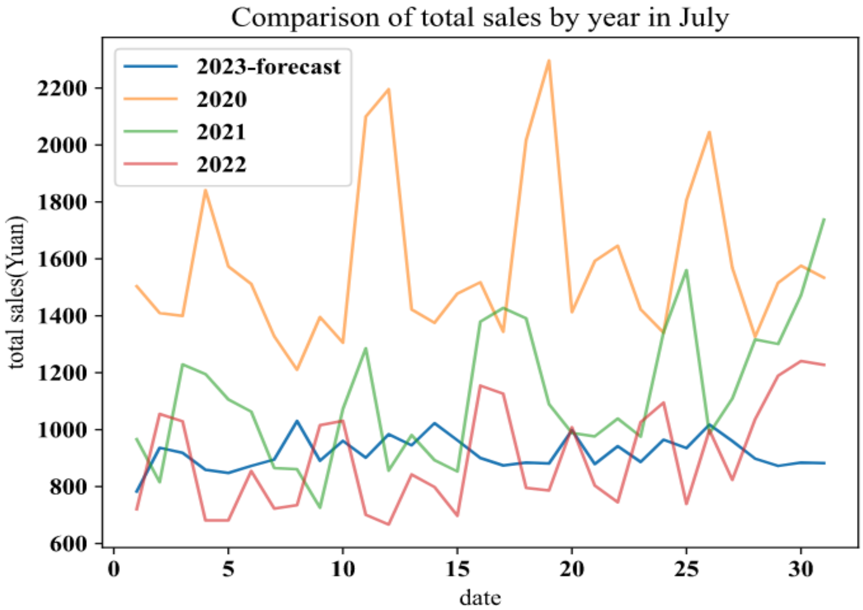

Figure 6: Comparison of total sales in July of different years

Upon reviewing the results, it's evident that the total sales' value range varies across different years within the same period. However, despite these differences, there's a notable similarity in the overall trend. The total sales exhibit strong consistency in the peaks and troughs, depicting a consistent cyclic pattern of rise and fall with a high degree of consistency. This consistency enhances the fitting effect, resulting in better alignment with the observed data.

Comparatively, the two standard root-mean-square errors (RMSE) calculated based on curve errors are contrasted with the normalized values of the mean absolute error (MAE). The obtained error table is presented in Table 2:

Table 2: Comparison of errors between different year curves and the 2023 curve

Year	RMSE	MAE
2020	0.722	0.663
2021	0.335	0.264
2022	0.193	0.167

The table primarily calculates two types of errors for the curves of 2020, 2021, and 2022 in comparison to the curve of 2023. Upon analysis, it becomes apparent that from 2020 to 2022, the error in the curve progressively diminishes compared to 2023. Specifically, 2022 exhibits the smallest error when compared to 2023, signifying that closer proximity between the years correlates with higher accuracy in the derived data. This observation supports the conclusion that the predicted values for July 2023 are accurate.

4. Conclusions

This study primarily focuses on utilizing the SARIMA model to devise replenishment and pricing strategies for supermarket products. Beginning with the cost-plus pricing method as the foundational pricing determinant, the study employs a SARIMA model to forecast the total sales volume for the upcoming month, subsequently deriving predictions for restocking volume and pricing. The comparative analysis of historical data from corresponding periods concludes that the supermarket replenishment and pricing strategy formulated based on the SARIMA model proves to be accurate.

The findings reveal fluctuating trends in the predicted results, showcasing inverse peaks and valleys between replenishment volume and pricing. This observed pattern aligns with anticipated norms, displaying a consistent trend akin to the same historical period, exhibiting minimal curve errors. Consequently, the SARIMA model demonstrates its capability to accurately formulate effective replenishment and pricing strategies for supermarkets.

References

[1] Lin Feng, Wan-Chih Wang, Jinn-Tsair Teng, Leopoldo Eduardo Cárdenas-Barrón. Pricing and lot-sizing decision for fresh goods when demand depends on unit price, displaying stocks and product age under generalized payments. European Journal of Operational Research [J].2022, 296(3):940-952.

[2] Cenying Yang, Yihao Feng, Andrew Whinston. Dynamic Pricing and Information Disclosure for Fresh Produce: An Artificial Intelligence Approach. Production and operations management [J].2022, 31(1):155-171.

[3] Rayner Alfred, Christylyn Leikson, B. Boniface, G. Tanakinjal, A. Kamu, Mori Kogid, S. Sondoh, Nolila Mohd Nawi, N. Arumugam, R. Andrias. Modelling and Forecasting Fresh Agro-Food Commodity Consumption Per Capita in Malaysia Using Machine Learning. Hindawi [J].2022, 2022(1):17.

[4] Lobna Nassar, Ifeanyi Emmanuel Okwuchi, Muhammad Saad, Fakhri Karray, Kumaraswamy Ponnambalam. Deep Learning Based Approach for Fresh Produce Market Price Prediction. International Joint Conference on Neural Networks (IJCNN) [J].2020:1-7.

[5] Islam Nasr, L. Nassar, F. Karray. Transfer Learning Framework for Forecasting Fresh Produce Yield and Price. International Joint Conference on Neural Networks (IJCNN) [J].2022, (3):1-8.

[6] Manish Shukla, Sanjay Jharkharia. Applicability of ARIMA Models in Wholesale Vegetable Market: An Investigation. International Journal of Information Systems and Supply Chain Management [J].2013, 6(3):105-119.

[7] Xuping Wang, Dongping Lin, Wenping Fan, Tianteng Wang. Research on Sales Forecast of Fresh Produce Considering Weather Factors. International Conference on Electronic Business (ICEB) [J].2018:541-548.

[8] Siti Nurhaliza, Wan Junita Raflah. The Application of a Cost-Plus Pricing Method in Determining the Selling Price of Dried Lomek Products (Case Study at Bumdes Kuala Alam). Inovbiz Jurnal Inovasi Bisnis Seri Manajemen Investasi dan Kewirausahaan (INOVBIZ) [J].2022:154-161.

[9] E. Kartika, M. R. Bakhtiar. Penentuan Cost of Goods Sold dan Penerapan Cost Plus Pricing Method dalam Menentukan Harga Jual: Study Penggilingan Padi UD Budi Luhur. Media Akuntansi Universitas Muhammadiyah Semarang [J]. 2020, 10(2):67-73.

[10] Nari Sivanandam Arunraj, Diane Ahrens, Michael Fernandes. Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry. International Journal of Operations Research and Information Systems (IJORIS)[J]. 2016, 7(2):21.