Transactions on Economics, Business and Management Research
ISSN: 2960-1789, eISSN: 2960-2254 | Volume 10
MEEA 2024

Optimizing Retail Inventory Management Through Time Series Analysis

Ziyan Fu
School of Business University of Bristol, Bristol, BS81QU, United Kingdom
uh23242@bristol.ac.uk

Abstract. The purpose of this study is to investigate the application of time series analysis in the management of retail inventory, with a particular emphasis on the precision of forecasting as well as the applicability of ARIMA models and seasonal decomposition techniques. This study aimed to evaluate these models' effectiveness in improving the scientific and precise nature of inventory management. This was accomplished by using legitimate retail data. According to the findings, the ARIMA model is effective in managing data that is stable or does not exhibit seasonal patterns. On the other hand, the seasonal decomposition technique demonstrates strong forecasting accuracy in managing data that exhibits seasonal fluctuations. Not only does this research improve the theoretical implementation of time series analysis in the retail sector, but it also provides retailers with practical strategies for mitigating economic losses that are caused by inventory imbalance. When it comes to the management of retail inventory, the findings of the research are significant for improving both overall cost control and customer satisfaction.

Keywords: Time Series Analysis; ARIMA Model; Seasonal Decomposition; Forecasting Model Optimization.

1. Introduction

Efficient inventory management is crucial for businesses to remain competitive in the rapidly evolving retail market [1]. Having an excessive amount of inventory not only results in the unnecessary allocation of capital and higher expenses for storage, but it can also lead to missed sales opportunities and reduced customer satisfaction. Inventory management is a fundamental operational activity in the retail industry that has a direct impact on cost control and the quality of customer service [2]. Efficient inventory management not only minimizes unnecessary expenses related to storing goods and prevents excessive use of financial resources, but also guarantees that customer needs are met promptly, thereby improving customer satisfaction and market competitiveness [3]. Therefore, accurate forecasting of product demand is essential for optimizing inventory management, reducing costs, and enhancing customer satisfaction [4]. This study aims to investigate the use of time series analysis in retail inventory management to improve inventory levels by accurately predicting demand. This will result in better cost control and customer service levels. The retail industry has the potential to greatly improve inventory management strategies using data technology, particularly with the aid of big data and machine learning techniques. Utilizing historical sales data, and time-series analysis can assist companies in identifying cyclical and trending fluctuations in product demand, thereby enhancing the scientific and precise nature of inventory management.

While the significance of inventory management is widely recognized, accurately predicting product demand continues to be a difficult task. The presence of demand uncertainty, seasonal fluctuations, and rapidly changing market trends frequently pose challenges for traditional forecasting techniques [5]. Therefore, this can lead to either excessive or insufficient inventory levels, ultimately impacting operational effectiveness and customer satisfaction [6]. Nevertheless, although time series analysis techniques have undergone extensive research, their implementation in actual retail settings continues to encounter numerous obstacles, including the selection of appropriate models, fine-tuning of parameters, and adaptability to unforeseen market events [7]. This study aims to analyze the forecasting accuracy and operational feasibility of the autoregressive integrated moving average

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.

model (ARIMA) and seasonal decomposition. The analysis will be conducted using actual retail data. This study aims to evaluate the efficacy of time series analysis in practical inventory management through empirical research. Additionally, it seeks to offer actionable strategies to assist retailers in mitigating economic losses resulting from inventory imbalance [8]. Time series analysis is a statistical method employed to examine data points arranged in chronological sequence [9]. ARIMA and seasonal decompositions are commonly employed models for predicting economic and business trends over various time horizons. The ARIMA model, as suggested by Box and Jenkins, is appropriate for analyzing time series data that are stable or non-seasonal [10]. On the other hand, seasonal decompositions are better suited for handling data that exhibit notable seasonal fluctuations. Recent studies have shown that time series analysis is effective in predicting retail inventory demand. For instance, Taylor and Letham utilized Facebook's Prophet model in their study to predict retail sales data and demonstrated a significant level of precision [11]. Furthermore, studies have demonstrated that seasonal decomposition techniques are highly effective in handling demand forecasting during seasonal peaks, such as the Christmas and summer sales peaks.

This paper will examine the theoretical foundations of time series forecasting models and their application in retail case studies. By utilizing advanced statistical methods, the paper aims to optimize inventory management for retailers, offering a scientific and practical solution. This not only enhances the theoretical scope of time series analysis in various application domains, but also offers novel viewpoints and approaches for improving inventory optimization in the retail sector.

2. Methods

2.1. Data Source

The data originates from the sales records of Walmart. The data was converted to represent the average monthly sales prices in US dollars. It includes historical data from February 5, 2010, to November 1, 2012. The dataset consists of 6,436 distinct observations obtained from various stores' sales records.

2.2. Variable Selection

The sales data is categorized by various departments within a store to offer a deeper understanding of patterns within specific product categories. Shops can be classified based on their type, such as urban, suburban, or rural, and this classification can impact sales trends. Due to the influence of weather conditions on shopping patterns and sales, local temperatures exhibit irregular and unpredictable variations. As a result, prices experience frequent fluctuations and become challenging to forecast, as depicted in Figure 1.

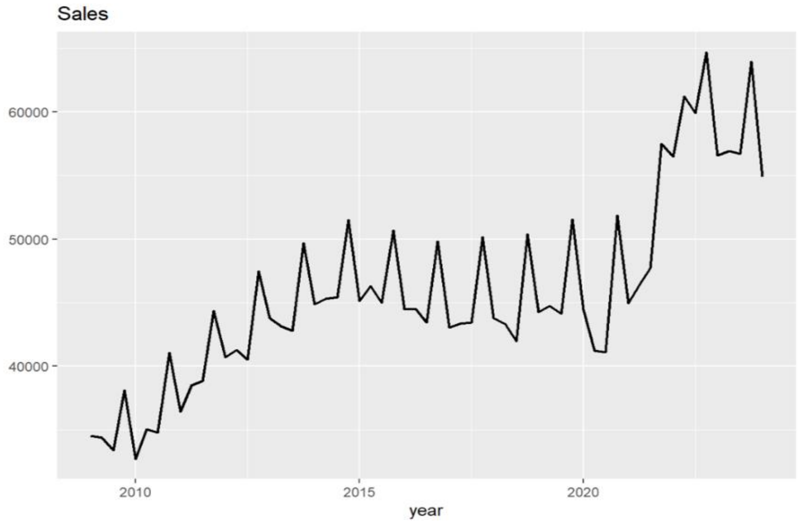

Figure 1. Walmart sales data

Figure 1 illustrates that Walmart's sales prices exhibited a cyclical increase until January 2010, followed by a cyclical decrease until after 2015. However, the sales figures remained consistently higher than those of 2010. Over the next five years, there is a consistent but cyclical pattern. Subsequently, there was a notable decline in 2020, which can be attributed to the economic vulnerability caused by the COVID-19 pandemic. Subsequently, there is a notable increase in the trend.

Based on the seasonal breakdown in Figure 2, sales are seasonal and on an upward trend.

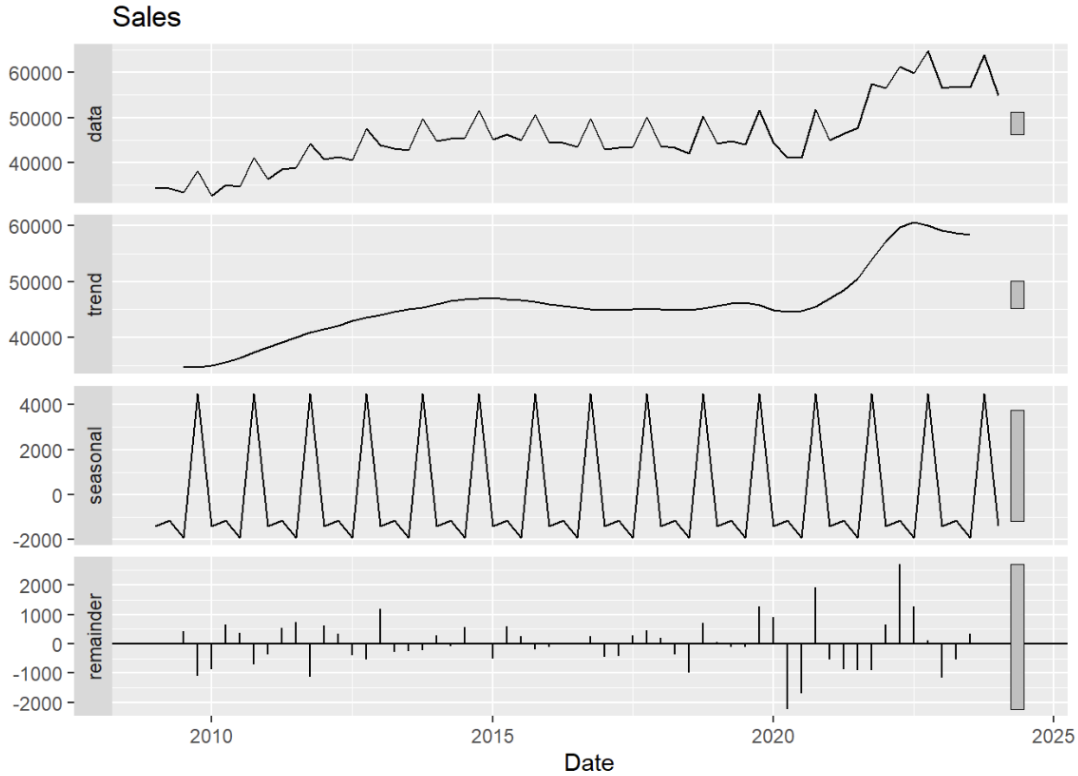

Figure 2. Seasonal Breakdown

2.3. Model Selection

This paper employs the Autoregressive Integrated Moving Average (ARIMA) model to analyze the time series data of retail sales in Wal-Mart stores. ARIMA is a comprehensive and straightforward model used for analyzing time series data [12]. It incorporates three main components: Autoregressive (AR), Integrated (I), and Moving Average (MA). The AR component is utilized to handle the autoregressive segment of the time series, which considers the influence of past period observations on the current observations [13]. The I component is utilized to transform a non-stationary time series into a stationary state by eliminating the trend and seasonal components through differential processing. On the other hand, the MA component is employed to address the moving average component of the time series, considering the impact of previous forecasting errors on the current values.

3. Results and Discussion

3.1. Data Processing

The data used for ARIMA needs to be ensured to be stationary, therefore, the ACF plot (Figure 3) and PACF plot (Figure. 4) meet the requirements based on the results of the first-order differencing. p-value is 0.3. therefore, the resultant data obtained from the first-order differencing can be stationary.

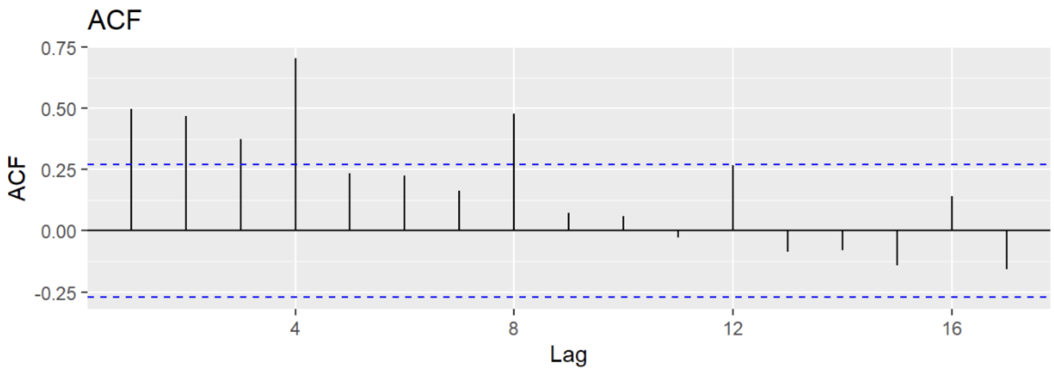

Figure 3. ACF diagram

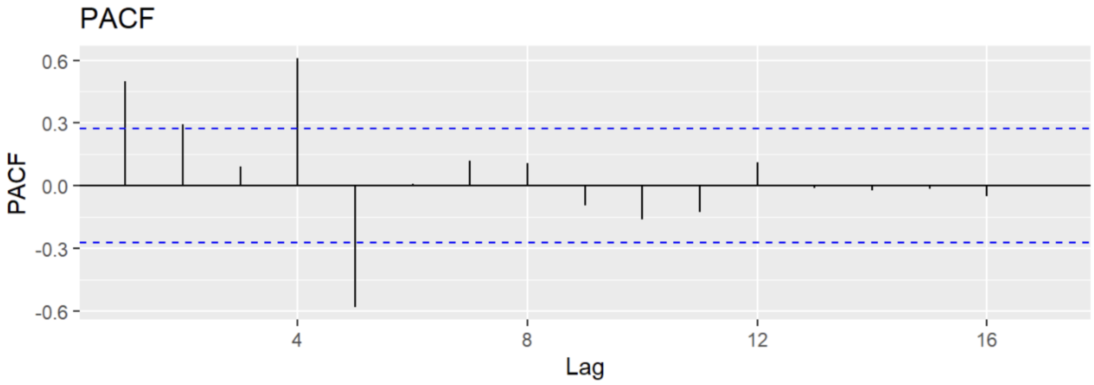

Figure 4. PACF diagram

3.2. Model Evaluation

The model's accuracy is reflected in the deviation of the relationship between the fitted and actual values. The smaller the deviation, the better the fit and the more accurate the model [12]. According to the previous analysis, the following part is a refinement of the above part to analyze the whole model more comprehensively.

Table 1. Model evaluation

ARIMA Model	AIC
(2, 0, 2) (1, 1, 1)	834.5917
(0, 0, 0) (0, 1, 0)	888.4188
(1, 0, 0) (1, 1, 0)	826.6812
(0, 0, 1) (0, 1, 1)	852.622
(1, 0, 0) (0, 1, 0)	833.7081
(1, 0, 0) (2, 1, 0)	829.0314
(1, 0, 0) (1, 1, 1)	828.9623
(1, 0, 0) (0, 1, 1)	827.0891
(1, 0, 0) (2, 1, 1)	831.2837
(0, 0, 0) (1, 1, 0)	888.4772
(2, 0, 0) (1, 1, 0)	827.8877
(1, 0, 1) (1, 1, 0)	827.6916
(0, 0, 1) (1, 1, 0)	852.4305
(2, 0, 1) (1, 1, 0)	829.2253

Root Mean Square Error (RMSE) is a common measure of the difference between measurements and is used to measure the deviation between the fitted value and the actual value. Also, Akaike Information Criterion (AIC) can be used to select the model with the least amount of own parameter shucking in the data, to avoid overfitting conditions.

Combining these two points, the smaller the size of the model, the better the fit to the data. From the above Table 1, by using the ARIMA model for forecasting, the optimal model is when the AIC is minimum, the result shows that the model is optimal when the parameter is ARIMA (1, 0, 0) (1, 1, 0), currently, the AIC is 826.6812. Secondly, the value of the Root Mean Square Error (RMSE) for calculating the optimal model is 1162.152. Therefore, this paper concludes that the value of the Final model from the sales volume data, the Final model is chosen for forecasting and analysis.

3.3. Forecasting Results

From the above analysis ARIMA (1, 0, 0) (1, 1, 0) is selected for forecasting the sales volume data and the results are shown in Figure 5 and Table 2:

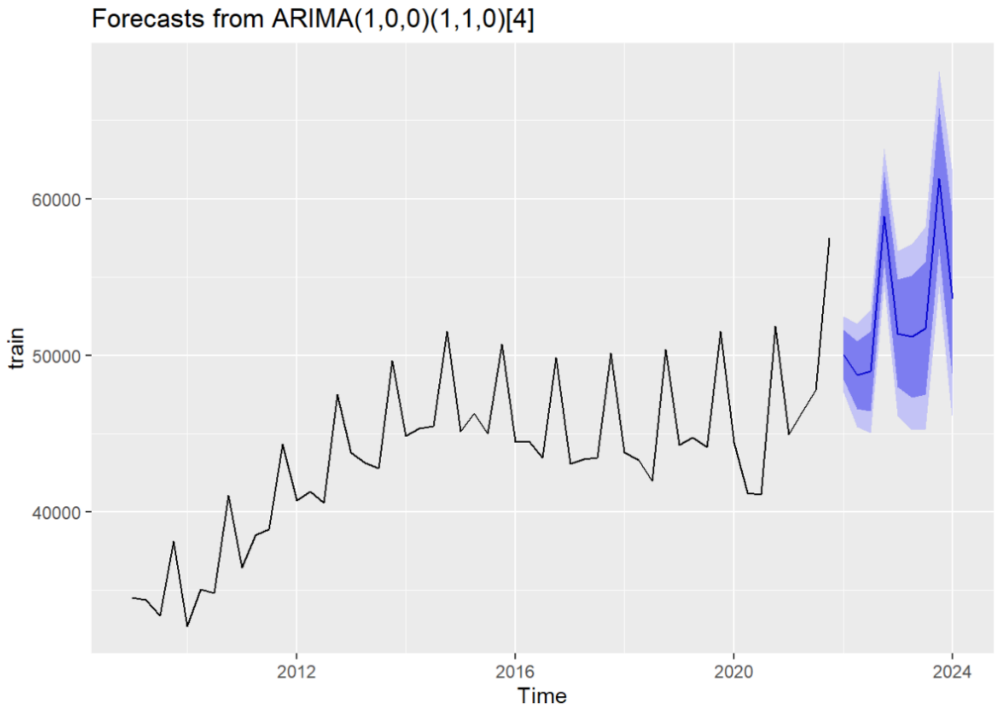

Figure 5. ARIMA (1, 0, 0) (1, 1, 0) prediction results

Table 2. ARIMA (1, 0, 0) (1, 1, 0) prediction results

Time	Point Forecast	L 80	H 80	L 95	H 95
2022 Q1	50050.00	48466.49	51633.52	47628.23	52471.78
2022 Q2	48719.09	46561.19	50876.98	45418.87	52019.30
2022 Q3	48956.31	46407.15	51505.47	45057.71	52854.92
2022 Q4	58868.10	56026.16	61710.03	54521.73	63214.46
2023 Q1	51393.45	47951.96	54834.94	46130.15	56656.75
2023 Q2	51176.78	47294.46	55059.10	45239.28	57114.28
2023 Q3	51720.11	47496.44	55943.78	45260.56	58179.65
2023 Q4	61288.78	56793.15	65784.41	54413.31	68164.26
2024 Q1	53602.80	48402.78	58802.82	45650.06	61555.55

The blue line indicates the prediction results of the model, and the black line indicates the actual data. From the above results, the predicted sales show an upward trend in the later stages and remain cyclical.

3.4. Residuals Checking

To ensure that the model is effective in interpreting the data information, the next step is to check whether its residuals are white noise. The detection results are shown in Figure 6 and Figure 7:

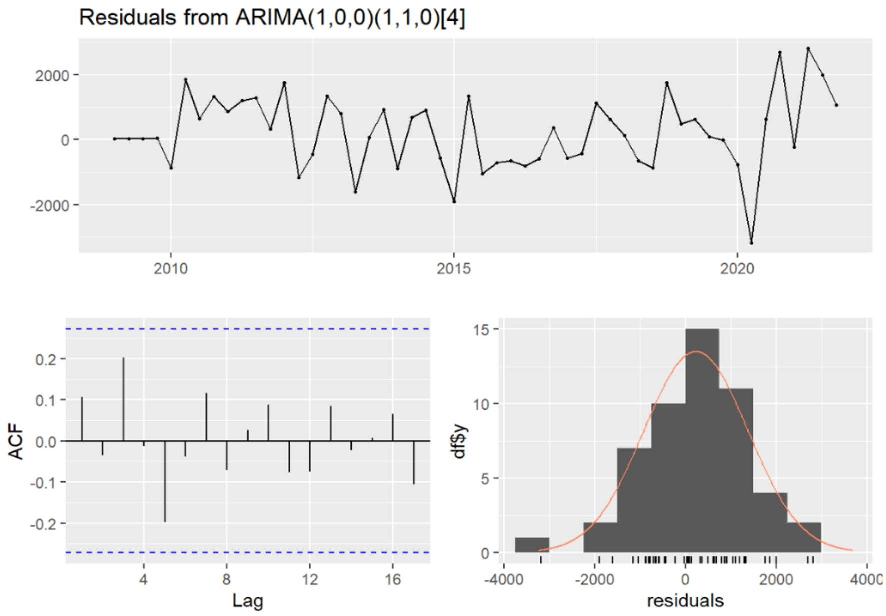

Figure 6. Residual test

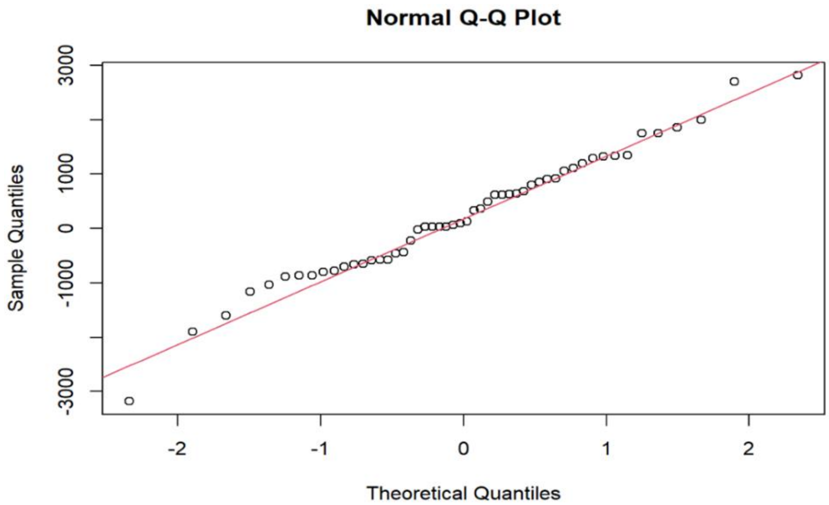

Figure 7. Displays the ARIMA model

The coordinates (1, 0, 0) represent a point in a three-dimensional space. The coordinates (1, 1, 0) represent a point in a three-dimensional space. Figure 6 illustrates that the magnitude of the residuals' changes after 2015 appears to level off, which aligns with the distribution pattern of white noise. All the values in the ACF plot are below the critical value. The data conforms to a normal distribution. Furthermore, in Figure 7, the QQ plot exhibits a linear pattern, indicating that the residuals can be considered as white noise. Furthermore, the error analysis demonstrates that the ARIMA error exhibits characteristics of white noise and adheres to the assumption of normality, indicating that the model performs effectively.

3.5. Analytical Reasoning

The ARIMA model is renowned for its efficient handling of univariate time series data. It can handle a diverse set of inventory data, particularly those that exhibit notable patterns and recurring features, thereby serving as a solid foundation for predicting demand. Furthermore, the ARIMA model's simplicity facilitates its practical operation and interpretation, which is especially crucial in business settings that demand swift implementation and comprehension [13]. Nevertheless, this paper has certain constraints due to the impact of the COVID-19 pandemic on the year 2020, which resulted in subpar performance in adapting to abrupt market fluctuations or exceptional cases, as indicated by the images. This occurs because the model relies on past data to make predictions, which hinders its ability to promptly adapt and modify its forecasts in response to abrupt changes. Furthermore, ARIMA models necessitate a substantial amount of past data to guarantee the dependability of predictions. If there is a lack of sufficient or incomplete data collection, the model's effectiveness will be significantly diminished, particularly when dealing with new products or incomplete data records.

4. Conclusion

To summarise, while the market is influenced by global economic fluctuations, the sales volume follows a cyclical pattern. Nevertheless, it is challenging to handle unforeseen events like the COVID-19 pandemic, where the worldwide economy is undergoing a process of recuperation following the outbreak. This situation hampers the ability to precisely forecast and analyze forthcoming data.

Furthermore, the forecasts and evaluations presented in this paper can offer valuable recommendations for enhancement. To address the limitations of the ARIMA model, one can consider employing machine learning techniques like Random Forest or Support Vector Machine to improve the prediction's flexibility and accuracy. By incorporating the ARIMA model, this approach enhances the capacity to effectively manage nonlinearities and abrupt fluctuations without sacrificing the benefits associated with it. In addition, implementing advanced techniques such as detailed outlier detection and seasonal adjustment during data preprocessing can enhance the smoothness of the data and consequently improve the forecasting performance of the ARIMA model. Therefore, it is imperative to predict consumer behavior and anticipate fluctuations in global market risk.

References

[1] E. Bayraktar, M. C. Jothishankar, E. Tatoglu, and T. Wu, Evolution of operations management: past, present and future. Management Research News, 30 (2007) 843–871.

[2] A. M. Mkonu and J.O. Gichana, Relationship between inventory management policies and supply chain performance of retail supermarkets in Nairobi, Kenya. International Journal of Supply Chain and Logistics, 3 (2019) 63.

[3] M. Bonney and M. Y. Jaber, Environmentally responsible inventory models: Non-classical models for a non-classical era. International Journal of Production Economics, 133 (2011) 43–53.

[4] C. Deng and Y. Liu, A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. Wireless Communications and Mobile Computing, (2021) 1–14.

[5] S. Ren, H.L. Chan, and T. Siqin, Demand forecasting in retail operations for fashionable products: methods, practices, and real case study. Annals of Operations Research, 291 (2019).

[6] A. S. Otto, D. M. Szymanski, and R. Varadarajan, Customer satisfaction and firm performance: insights from over a quarter century of empirical research. Journal of the Academy of Marketing Science, 48 (2019) 543–564.

[7] R. Aldrighetti, D. Battini, D. Ivanov, and I. Zennaro, Costs of resilience and disruptions in supply chain network design models: A review and future research directions. International Journal of Production Economics, 235 (2021).

[8] S. Gupta and D. Ramachandran, Emerging Market Retail: Transitioning from a Product-Centric to a Customer-Centric Approach. Journal of Retailing, 97 (2021).

[9] A. L. Schaffer, T. A. Dobbins, and S.A. Pearson, Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. BMC Medical Research Methodology, 21 (2021).

[10] T. Dimri, S. Ahmad, and M. Sharif, Time series analysis of climate variables using seasonal ARIMA approach. Journal of Earth System Science, 129 (2020).

[11] L. Liço, I. Enesi, and H. Jaiswal, Predicting Customer Behavior Using Prophet Algorithm In A Real Time Series Dataset. European Scientific Journal ESJ, 17 (2021).

[12] E. Beard, et al. Understanding and using time series analyses in addiction research. Addiction, 114 (2019) 1866–1884.

[13] I. O. Alade, M. A. Abd Rahman and T. A. Saleh, Modeling and prediction of the specific heat capacity of Al2 O3/water nanofluids using hybrid genetic algorithm/support vector regression model. Nano-Structures & Nano-Objects, 17 (2019) 103–111.