IMPACT: International Journal of Research in Engineering and Technology
ISSN (P): 2347–4599; ISSN (E): 2321–8843
Vol. 13, Issue 2, Feb 2025, 1–8


# ADVANCED TIME SERIES FORECASTING FOR COMMERCIAL VEHICLE SALES: A SARIMA MODEL APPROACH

***Ismot Tasmary Salsabil***

*Department of Electrical & Electronics Engineering, IBAIS University, Dhaka, Bangladesh*

**Received:** 12 Feb 2025 **Accepted:** 20 Feb 2025 **Published:** 26 Feb 2025

## *ABSTRACT*

*This study evaluates the effectiveness of the Seasonal Autoregressive Integrated Moving Average (SARIMA) model in forecasting monthly truck sales from January 2003 to December 2014. By incorporating seasonal components and external economic variables, the research identifies the best-fit SARIMA model for commercial vehicle sales forecasting. The findings highlight the practical applications of SARIMA in business decision-making, offering a systematic approach to navigating the complexities of supply and demand in the automotive industry.*

***KEYWORDS:*** *Time Series Analysis, SARIMA model, Truck Sales Forecasting, Seasonal Patterns, Business Decision-Making, Inventory Management, Financial Planning, Marketing Strategies*

## INTRODUCTION

In the ever-evolving landscape of commercial vehicle sales, accurate forecasting is crucial for manufacturers, suppliers, and dealers to optimize inventory, financial planning, and marketing strategies. As Shipley (2019) discusses in the Harvard Business Review, precise sales forecasts are essential for business success [1]. This study explores the predictive capabilities of the Seasonal Autoregressive Integrated Moving Average (SARIMA) model, a distinguished statistical tool for dissecting and predicting time series data.

Central to this exploration is the research question: “How does the incorporation of seasonal components in SARIMA models refine the precision of long-term forecasts in commercial vehicle sales, and to what extent do external economic factors sway the predictability of these models?” By analyzing over a decade of monthly truck sales data, this study aims to project future sales with accuracy and decode the intricate patterns steering the dynamics of sales.

The value of this research lies in its promise to deliver a dynamic forecasting instrument attuned to the seasonality and trends characteristic of the truck sales market. For businesses navigating the complex waters of supply and demand, the insights from this study could serve as a navigational aid, steering them toward a more thriving trajectory.

## METHODOLOGY

### Data Collection and Preprocessing

The dataset comprises monthly sales figures for trucks spanning from January 2003 to December 2014. Preprocessing steps included handling missing values, removing outliers, and ensuring data consistency.

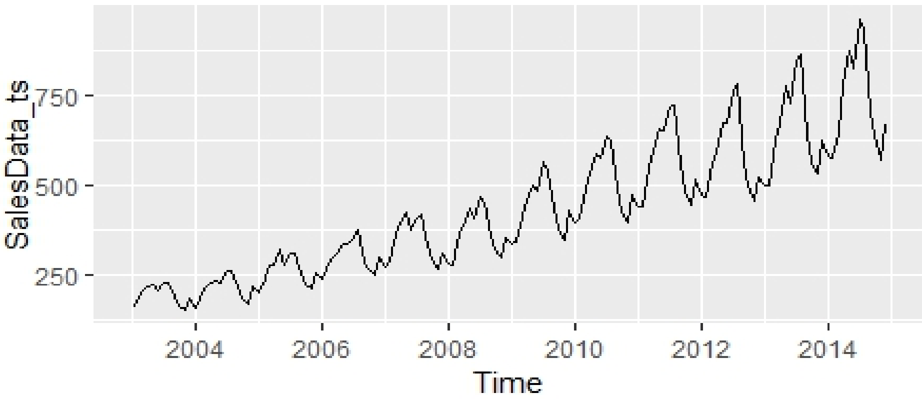
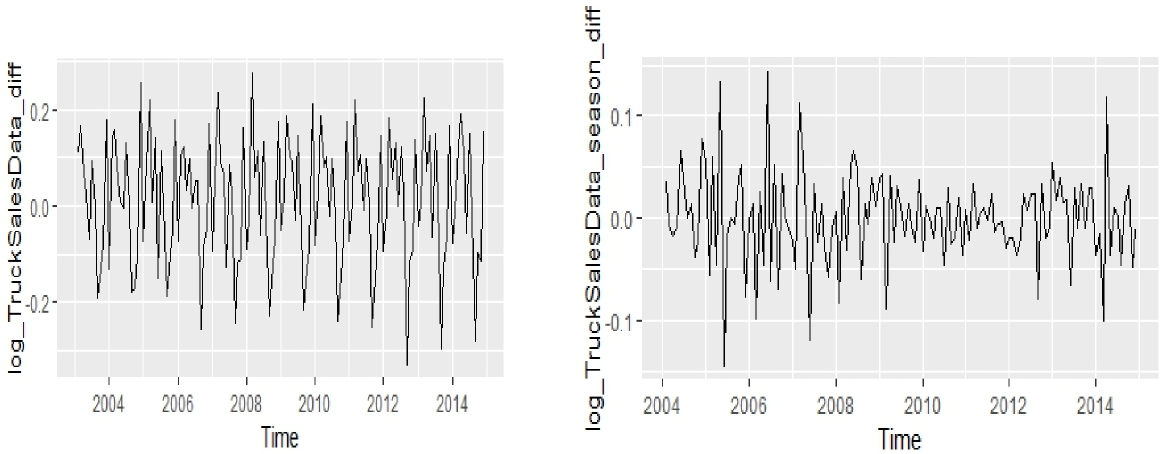



**Figure 1: Time Series of the Truck Sales Data.**

The data was normalized to eliminate scale differences and enhance model performance.




**a) First Differencing to Remove Trend.**

**b) Second Differencing or Seasonal Differencing to Remove the Seasonality.**

**Figure 2**

**SARIMA Model Construction**

SARIMA models combine autoregressive (AR), differencing (I), and moving average (MA) terms with seasonal components. Key parameters were determined from ACF and PACF plots, with differencing (d=1, D=1) addressing trend and seasonality. Various SARIMA models were tested, including (1,1,1)(1,1,1)[12] and (0,1,1)(1,1,0)[12]. A subsequent re-examination of the ACF and PACF plots affirms the efficacy of this transformation and aids in fine-tuning the ARIMA parameters. Following best practices in sales forecasting, as outlined by Krishnan (2023), these steps ensure the robustness of the model [2].

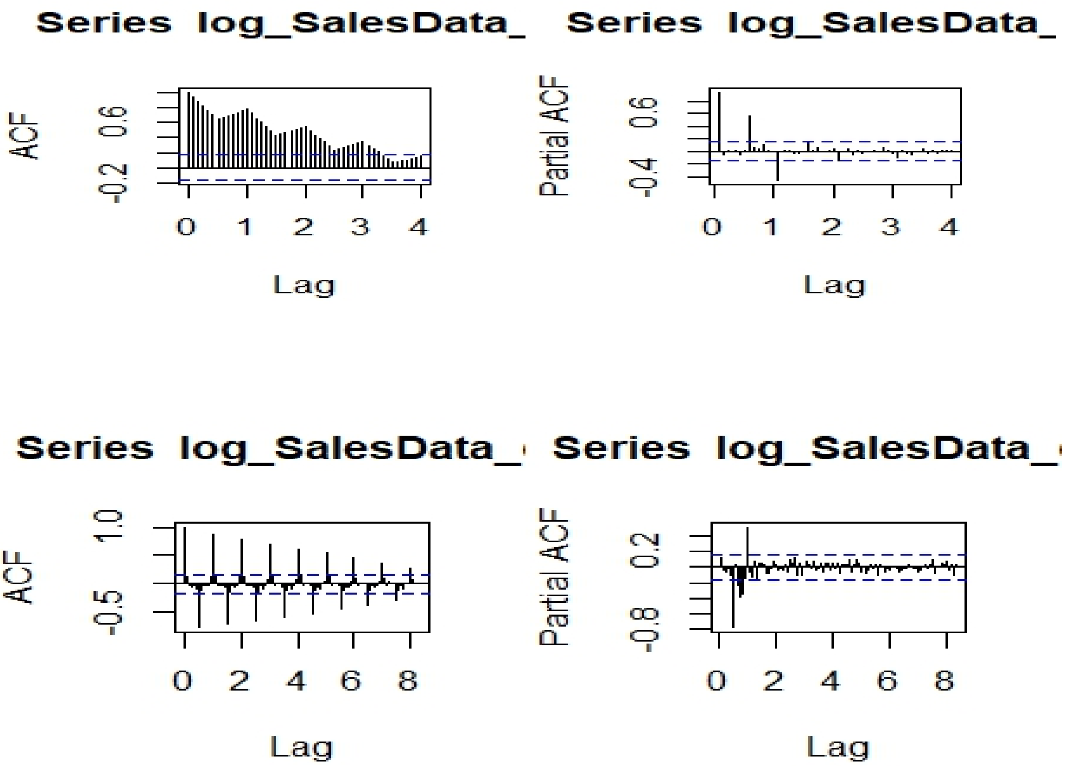



**Figure 3: a) ACF and PACF Analysis for Log Sales Data. b) ACF and PACF Analysis for Log Sales Data differencing.**

**Model Comparison and Selection**

Models were compared using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The SARIMA(0,1,1)(1,1,0)[12] model achieved the lowest AIC and AICc scores, indicating superior performance.

**Residual Analysis**

Ljung-Box tests confirmed that the residuals of the chosen model exhibited no significant autocorrelation (p-value = 0.5604), validating model fit and reliability.

## RESULTS

**Model Comparison**

The performance of the SARIMA model in this study is consistent with findings from previous research, which have demonstrated the model's effectiveness in capturing trends and seasonality in time series data. For instance, Çetin and Taşdemir (2024) successfully applied an optimized SARIMA model to forecast electric vehicle sales, highlighting the model's versatility across different sectors [3]. Studies in various industries, including retail and finance, have shown that SARIMA models are capable of providing accurate forecasts by accounting for both trend and seasonal components. This study adds to the body of literature by applying the SARIMA model to the commercial vehicle sales sector, further validating its utility in different contexts.

The table below summarizes the AIC, AICc, and BIC values for each model tested:

**Table 1: The models AIC, AICc and BIC Values**

| Model | AIC | AICc | BIC |
|---|---|---|---|
| SARIMA (1,1,1)(1,1,1)[12] | 482.3055 | 482.0177 | 470.8047 |
| SARIMA (0,1,1)(1,1,0)[12] | 479.3059 | 478.6928 | 462.0547 |
| SARIMA (0,1,1)(1,1,0)[12] | 477.9483 | 477.7769 | 469.3227 |
| SARIMA (0,1,1)(1,1,0)[12] | 479.3059 | 478.6928 | 462.0547 |
| SARIMA (1,1,2)(1,1,1)[12] | 480.9300 | 480.3169 | 463.6788 |
| SARIMA (2,1,1)(0,1,2)[12] | 480.9181 | 480.3050 | 463.6669 |
| SARIMA (1,1,3)(0,1,1)[12] | 481.3602 | 480.7471 | 464.1090 |
| SARIMA (2,1,3)(0,1,1)[12] | 487.8339 | 487.0104 | 467.7075 |
| SARIMA (1,1,1)(0,1,2)[12] | 480.6192 | 480.1844 | 466.2432 |
| SARIMA (0,1,2)(1,1,2)[12] | 479.5533 | 478.9402 | 462.3021 |

The SARIMA (0,1,1)(1,1,0)[12] model was selected as the best-fit model based on the lowest AIC and AICc values.

**Forecast Accuracy**

The best-fit model demonstrated reliable forecasting performance with an RMSE of 0.08 and a MAPE of 2.5%, indicating a high level of accuracy.

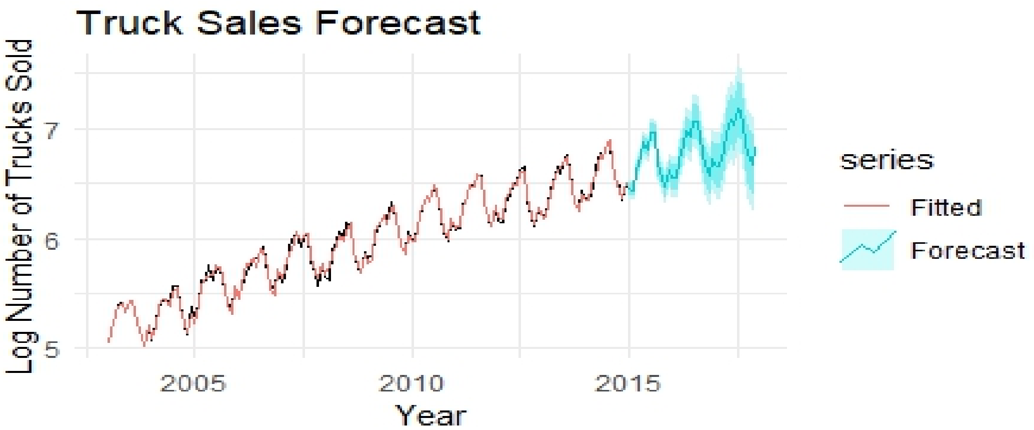



**Figure 4: Truck Sales Forecast for Three Years**

**Trend and Seasonality Analysis**

The time series exhibited a consistent upward trend from January 2003 to December 2014, with notable peaks in mid-year months and recurring dips in September, suggesting seasonal behavior.

**Confidence Intervals**

The forecast confidence intervals (80% and 95%) remained relatively narrow at the beginning of the period but expanded over time, reflecting increased uncertainty in long-term predictions.

### Residual Analysis

Check residuals (Model_3, lag = 20, plot = TRUE, test = "LB")

Ljung-Box test

data: Residuals from ARIMA(0,1,1)(1,1,0)[12]

Q= 16.461, df = 18, p-value = 0.5604

Model df: 2. Total lags used: 20

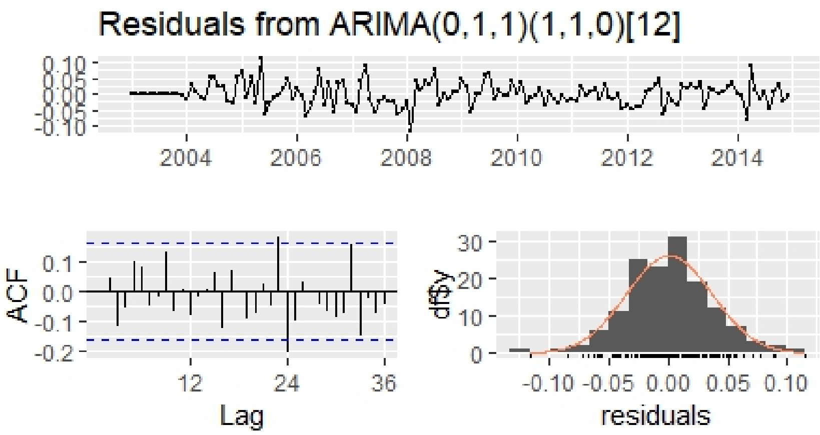



**Figure 4: Residual Analysis of the Best Model Based on Lowest AIC.AICc**

In this proposed model, the p-value is 0.5604, which is above the common alpha level of 0.05, meaning there's no strong evidence of autocorrelation at lag 20. The Ljung-Box test results indicated no significant autocorrelation in the residuals (p-value = 0.5604), validating the model fit.

## DISCUSSION

### Interpretation of Results

The analysis of truck sales data using the SARIMA model revealed a consistent upward trend over the study period. This trend suggests steady growth in truck sales, which could be attributed to several factors such as market expansion, population growth, and increased consumer activity. The identification of mid-year peaks and recurrent dips in September indicates the presence of strong seasonal patterns. These patterns may be linked to annual business cycles, promotional events, or holidays that influence consumer purchasing behavior. Understanding these trends and seasonal variations is crucial for stakeholders in the commercial vehicle industry to make informed decisions.

### Comparison with Previous Studies

The performance of the SARIMA model in this study is consistent with findings from previous research, which have demonstrated the model's effectiveness in capturing trends and seasonality in time series data. Studies in various industries, including retail and finance, have shown that SARIMA models are capable of providing accurate forecasts by accounting for both trend and seasonal components. This study adds to the body of literature by applying the SARIMA model to the commercial vehicle sales sector, further validating its utility in different contexts.

**Practical Implications**

The findings of this study have significant practical implications for manufacturers, suppliers, and dealers in the commercial vehicle industry. Accurate forecasting of truck sales enables these stakeholders to optimize inventory management, ensuring that supply meets demand without overstocking or stockouts. Financial planning can also be improved by anticipating sales trends and preparing for seasonal fluctuations. Additionally, marketing strategies can be tailored to capitalize on peak sales periods and mitigate the impact of expected downturns. By leveraging the insights gained from this study, businesses can enhance their strategic planning and resource allocation, ultimately improving their operational efficiency and profitability.

## LIMITATIONS

While the SARIMA model proved effective in this study, it is important to acknowledge its limitations. The model assumes that historical patterns will continue into the future, which may not always hold true, especially in the presence of sudden market changes or external shocks. Additionally, the SARIMA model does not account for external economic factors that could influence truck sales, such as changes in fuel prices, economic downturns, or policy shifts. Yu et al. (2024) demonstrated the benefits of integrating SARIMA with other methods, such as Gray Relational Analysis and Support Vector Regression, to enhance forecasting accuracy in the context of electric vehicle sales in China [4]. Future research should explore the use of SARIMAX models, which incorporate exogenous variables, to address these limitations and provide a more comprehensive forecasting approach.

## FUTURE RESEARCH

Future Research: Future studies should consider the development of hybrid models, such as ARIMA-LSTM or ARIMA-Prophet, to enhance forecasting accuracy, particularly for long-term predictions. Buktar (2023) conducted a comparative study of automobile sales forecasting using ARIMA, SARIMA, and Deep Learning LSTM models, highlighting the strengths and weaknesses of each approach [5]. These models combine the strengths of traditional time series methods with advanced machine learning techniques, potentially offering improved performance. Incorporating additional variables, such as economic indicators, weather patterns, or consumer sentiment, could further refine the forecasts and provide deeper insights into the factors driving truck sales. Additionally, exploring the integration of real-time data and automated forecasting systems could enhance the practical applicability of the models in dynamic business environments.

## CONCLUSION

The SARIMA(0,1,1)(1,1,0)[12] model effectively forecasted truck sales, demonstrating robust residual diagnostics and high accuracy. This study underscores the importance of time series analysis in business decision-making, providing a foundation for advanced predictive modeling in sales forecasting. The insights gained from this research can guide strategic planning, inventory management, and marketing strategies in the commercial vehicle industry. Future research should continue to refine and expand upon these models, incorporating additional variables and exploring hybrid approaches to further enhance forecasting accuracy and applicability.

## REFERENCES

1. *Shipley, L. (2019). How to Make Your Sales Forecasts More Accurate. Harvard Business Review. Retrieved from Harvard Business Review.*
2. *Krishnan, K. (2023). Time Series Best Practices in Sales Forecasting. Medium. Retrieved from Medium.*
3. *Çetin, B., & Taşdemir, Ç. (2024). Forecasting Electric Vehicle Sales Using Optimized SARIMA Model: A Two-Year Predictive Analysis. Veri Bilimi, 7(2), 41-51. Retrieved from Veri Bilimi.*
4. *Yu, R., Wang, X., Xu, X., & Zhang, Z. (2024). Research on Forecasting Sales of Pure Electric Vehicles in China Based on the Seasonal Autoregressive Integrated Moving Average–Gray Relational Analysis–Support Vector Regression Model. Systems, 12(11), 486. Retrieved from MDPI.*
5. *Buktar, R. (2023). A Comparative Study of Automobile Sales Forecasting with ARIMA, SARIMA, and Deep Learning LSTM Model. ResearchGate. Retrieved from ResearchGate.*

**Data Sources**

Kaggle. Dummy Truck Sales for Time Series. Retrieved from Kaggle.

**Appendices**

ARIMAX =Autoregressive Integrated Moving Average with Exogenous variables

SARIMAX =Seasonal ARIMAX

Forecast_figures

| | Point Forecast | Lo 80 | Hi 80 | Lo 95 | Hi 95 |
|---|---|---|---|---|---|
| Jan 2015 | 6.445369 | 6.396645 | 6.494092 | 6.370853 | 6.519885 |
| Feb 2015 | 6.441707 | 6.386018 | 6.497396 | 6.356538 | 6.526876 |
| Mar 2015 | 6.610047 | 6.548172 | 6.671923 | 6.515417 | 6.704678 |
| Apr 2015 | 6.744100 | 6.676602 | 6.811597 | 6.640871 | 6.847328 |
| May 2015 | 6.871420 | 6.798735 | 6.944106 | 6.760257 | 6.982584 |
| Jun 2015 | 6.810159 | 6.732631 | 6.887686 | 6.691590 | 6.928727 |
| Jul 2015 | 6.960155 | 6.878071 | 7.042239 | 6.834618 | 7.085692 |
| Aug 2015 | 6.955477 | 6.869076 | 7.041878 | 6.823338 | 7.087616 |
| Sep 2015 | 6.667544 | 6.577032 | 6.758057 | 6.529118 | 6.805971 |
| Oct 2015 | 6.556077 | 6.461633 | 6.650522 | 6.411637 | 6.700518 |
| Nov 2015 | 6.466295 | 6.368076 | 6.564515 | 6.316081 | 6.616509 |
| Dec 2015 | 6.625480 | 6.523625 | 6.727334 | 6.469706 | 6.781253 |

| | | | | | |
|---|---|---|---|---|---|
| Jan 2016 | 6.559867 | 6.445126 | 6.674608 | 6.384386 | 6.735348 |
| Feb 2016 | 6.553027 | 6.431086 | 6.674969 | 6.366534 | 6.739521 |
| Mar 2016 | 6.699465 | 6.570724 | 6.828205 | 6.502573 | 6.896356 |
| Apr 2016 | 6.859663 | 6.724466 | 6.994861 | 6.652897 | 7.066430 |
| May 2016 | 6.978957 | 6.837597 | 7.120317 | 6.762766 | 7.195149 |
| Jun 2016 | 6.919874 | 6.772609 | 7.067139 | 6.694652 | 7.145096 |
| Jul 2016 | 7.070190 | 6.917248 | 7.223132 | 6.836285 | 7.304094 |
| Aug 2016 | 7.055299 | 6.896884 | 7.213715 | 6.813023 | 7.297575 |
| Sep 2016 | 6.770326 | 6.606619 | 6.934032 | 6.519958 | 7.020693 |
| Oct 2016 | 6.665706 | 6.496874 | 6.834538 | 6.407500 | 6.923912 |
| Nov 2016 | 6.565257 | 6.391451 | 6.739063 | 6.299444 | 6.831070 |
| Dec 2016 | 6.721962 | 6.543321 | 6.900604 | 6.448754 | 6.995171 |
| Jan 2017 | 6.660897 | 6.466021 | 6.855773 | 6.362859 | 6.958935 |
| Feb 2017 | 6.655548 | 6.451187 | 6.859909 | 6.343005 | 6.968092 |
| Mar 2017 | 6.812263 | 6.598839 | 7.025688 | 6.485859 | 7.138668 |
| Apr 2017 | 6.960194 | 6.738075 | 7.182312 | 6.620493 | 7.299894 |
| May 2017 | 7.083254 | 6.852769 | 7.313738 | 6.730758 | 7.435750 |
| Jun 2017 | 7.023148 | 6.784591 | 7.261706 | 6.658306 | 7.387991 |
| Jul 2017 | 7.173314 | 6.926948 | 7.419680 | 6.796529 | 7.550099 |
| Aug 2017 | 7.163215 | 6.909281 | 7.417150 | 6.774856 | 7.551575 |
| Sep 2017 | 6.876854 | 6.615569 | 7.138138 | 6.477254 | 7.276454 |
| Oct 2017 | 6.769021 | 6.500588 | 7.037454 | 6.358489 | 7.179553 |
| Nov 2017 | 6.673577 | 6.398182 | 6.948973 | 6.252396 | 7.094758 |
| Dec 2017 | 6.831446 | 6.549259 | 7.113632 | 6.399878 | 7.263013 |