VERİ BİLİMİ DERGİSİ
www.dergipark.gov.tr/veri

Forecasting Electric Vehicle Sales Using Optimized SARIMA Model: A Two-Year Predictive Analysis

Buse ÇETİN[1], Çağatay TAŞDEMİR[2,3,*]

[1] *Bursa Technical University, Department of Industrial Engineering (Grad. School), Bursa, Türkiye*
[2] *Bursa Technical University, Department of Industrial Engineering (Grad. School), Bursa, Türkiye*
[3] *Bursa Technical University, Department of Forest Industry Engineering, Bursa, Türkiye*

Abstract

The rapid expansion of the electric vehicle (EV) market underscores the need for accurate forecasting models to guide decision-making for manufacturers, policymakers, and stakeholders. This study leverages the Seasonal Autoregressive Integrated Moving Average (SARIMA) model to predict monthly EV sales for the next two years based on historical sales data from January 2021 to December 2023. The data is sourced from the U.S. Department of Energy's 'Monthly Sales of New Light-Duty EVs in the United States' report. The SARIMA model is optimized through a comprehensive grid search, resulting in an optimal configuration of (1, 0, 2) for the non-seasonal component and (1, 0, 1, 12) for the seasonal component. The methodology involves preprocessing the sales data to confirm stationarity using the Augmented Dickey-Fuller (ADF) test. A grid search identifies the optimal parameters, with model performance evaluated using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC). The chosen model exhibits an AIC of 739.51, BIC of 749.01, and HQIC of 742.82, indicating a good fit. The forecasting results reveal a consistent upward trend in EV sales over the next 24 months, with the model predicting sales to reach approximately 96.076 units by January 2024, peaking at 108,559 units in July 2024 and slightly tapering off to 100.676 units by December 2025. These projections underscore the increasing consumer adoption of electric vehicles and provide valuable insights for industry stakeholders. With its consistent upward trend, the predicted growth trajectory highlights the potential for continued market expansion, driven by advancements in EV technology, increasing environmental awareness, and supportive governmental policies. In conclusion, the SARIMA model provides a reliable forecast of EV sales, facilitating informed strategic planning and resource allocation for industry participants. This research contributes to the broader understanding of market dynamics in the rapidly evolving electric vehicle sector. It underscores the importance of robust predictive analytics in supporting sustainable growth, instilling a sense of optimism and hope for the industry's future.

Makale Bilgisi

Başvuru: 09/10/2024
Kabul: 01/12/2024

Keywords: *Electric Vehicles, Demand Forecasting, Time Series Analysis, SARIMA*

İletişim e-posta: 22435025014@ogrenci.btu.edu.tr and cagatay.tasdemir@btu.edu.tr

Optimize Edilmiş SARIMA Modeli Kullanarak Elektrikli Araç Satışlarının Tahmini: İki Yıllık Öngörü Analizi

Özet

Elektrikli araç (EV) pazarının hızla genişlemesi, üreticiler, politika yapıcılar ve paydaşlar için doğru tahmin modellerine olan ihtiyacı artırmaktadır. Bu çalışma, Ocak 2021'den Aralık 2023'e kadar olan tarihi satış verilerine dayanarak, önümüzdeki iki yıl için aylık EV satışlarını tahmin etmek amacıyla Mevsimsel Otoregresif Entegre Hareketli Ortalama (SARIMA) modelini kullanmaktadır. Veriler, ABD Enerji Bakanlığı'nın 'Amerika Birleşik Devletleri'nde Yeni Hafif Hizmet EV'lerinin Aylık Satışları' raporundan elde edilmiştir. SARIMA modeli, kapsamlı bir grid arama yoluyla optimize edilerek, mevsimsel olmayan bileşen için (1, 0, 2) ve mevsimsel bileşen için (1, 0, 1, 12) olarak belirlenen optimal konfigürasyona ulaşılmıştır. Yöntem, satış verilerinin Augmented Dickey-Fuller (ADF) testi kullanılarak durağanlığının doğrulanmasını içermektedir. Optimal parametrelerin belirlenmesi için grid arama kullanılmış ve modelin performansı Akaike Bilgi Kriteri (AIC), Bayesyen Bilgi Kriteri (BIC) ve Hannan-Quinn Bilgi Kriteri (HQIC) ile değerlendirilmiştir. Seçilen modelin AIC değeri 739.51, BIC değeri 749.01 ve HQIC değeri 742.82 olup, iyi bir uyum sağladığını göstermektedir. Tahmin sonuçları, önümüzdeki 24 ay boyunca EV satışlarında istikrarlı bir artış eğilimi ortaya koymakta; model, Ocak 2024'te yaklaşık 96.076 adet satış öngörmekte, Temmuz 2024'te 108.559 adede ulaşacağını ve Aralık 2025'te hafif bir düşüşle 100.676 adede gerileyeceğini tahmin etmektedir. Bu projeksiyonlar, elektrikli araçların artan tüketici benimsenmesini vurgulamakta ve sektör paydaşları için değerli içgörüler sunmaktadır. Tahmin edilen büyüme eğilimi, EV teknolojisindeki gelişmeler, artan çevresel farkındalık ve destekleyici hükümet politikaları ile yönlendirilen pazar genişlemesi potansiyelini gözler önüne sermektedir. Sonuç olarak, SARIMA modeli, EV satışlarının güvenilir bir şekilde tahmin edilmesini sağlamakta ve sektör katılımcıları için stratejik planlama ve kaynak tahsisine yönelik bilinçli kararlar alınmasına katkıda bulunmaktadır. Bu araştırma, hızla gelişen elektrikli araç sektöründe pazar dinamiklerinin daha geniş bir şekilde anlaşılmasına katkı sağlamakta ve sürdürülebilir büyümeyi desteklemede güçlü tahmin analitiklerinin önemini vurgulamaktadır. Bu da, sektörün geleceği için bir iyimserlik ve umut duygusu aşılamaktadır.

Anahtar Kelimeler: *Elektrikli Araçlar, Talep Tahmini, Zaman Serisi Analizi, SARIMA*

1 Introduction

Today's world faces significant environmental challenges, including climate change, increasing carbon emissions, and rapidly depleting natural resources. These global issues have driven governments and communities to seek sustainable and environmentally friendly solutions. The growth in global primary energy consumption in 2017 averaged 2.2%. This growth rate has averaged 1.7% per year over the past decade. Notably, vehicles hold a substantial share in energy consumption. As is well known, vehicles require a propulsion system to generate motion, typically achieved through two methods: the first and most important is the internal combustion engine, and the second is the electric motor [1]. This has demonstrated the need for a transformation in the automotive industry. The necessity for this transformation in the transportation sector has found a tangible solution with the rapid proliferation of electric vehicles (EVs). In particular, government incentive policies, innovative R&D investments by the automotive industry, and the increasing environmental awareness of consumers have enabled the electric vehicle market to grow with accelerating momentum worldwide. However, reliable future projections are essential to manage this growth sustainably and accurately determine long-term strategies. Specifically, forecasting the future trajectory of electric vehicle sales is critical for many stakeholders, ranging from manufacturers to policymakers.

The electric vehicle market is experiencing a steady expansion, driven by factors such as the rapid integration of new technologies, decreasing battery costs, and improvements in charging infrastructure. It is particularly noted that producing new-generation vehicle batteries has a more significant environmental impact than vehicles with internal combustion engines. Therefore, systems that extend the lifespan of batteries are crucial not only from an economic standpoint but also from an environmental perspective [2]. Accurately predicting the future trends of this growth is possible through short-term assessments based on current developments and a thorough analysis of historical data.

This study aims to forecast future sales volumes by analyzing the monthly electric vehicle sales data recorded in the United States between 2021 and 2023. Seasonal fluctuations and long-term trends observed in historical data are critical in understanding market dynamics and making strategic decisions. In the electric vehicle sector, the increase or decrease in consumer demand during specific periods necessitates accurate demand forecasting that accounts for seasonal effects. Effectively managing processes such as production planning, supply chain management, and resource allocation is directly related to the accuracy of these forecasts. Therefore, developing a forecasting model that considers overall trends and seasonal effects is vital in supporting the industry's future strategies.

This study utilizes the SARIMA (Seasonal Autoregressive Integrated Moving Average) model, widely applied in the analysis and forecasting of seasonal time series, to predict electric vehicle (EV) sales. Accurately capturing the seasonal cycles and long-term growth trends in EV sales, this model is a valuable tool for understanding market dynamics and generating reliable future projections. This paper provides a detailed overview of the dataset and the theoretical foundation of the SARIMA model, with analyses conducted using the Python programming language and Google Colab platform. The model setup, optimization processes, and resulting forecasts are presented, followed by a discussion of the study's findings, limitations, and recommendations for future research.

2 Literature Review

In the literature, studies that mainly employ the SARIMA method for forecasting have been reviewed and briefly discussed below.

Kim et al. (2019) developed a model to forecast short-term electricity load for institutional buildings by analyzing past electricity consumption data using ARIMA and Artificial Neural Networks (ANN). To evaluate forecast accuracy, they used Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics. Results indicated that the ANN model achieved lower error rates, with 2.5 kWh MAE and 3.8 kWh RMSE, compared to the ARIMA model. These findings suggest that ANN outperforms ARIMA for short-term electricity load forecasting in institutional settings [3].

In their research, C. Nwokike et al. (2020) compared the Seasonal Artificial Neural Network (SANN) and SARIMA models to forecast monthly rainfall in Umuahia. By analyzing historical rainfall data, they assessed the performance of each model using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics. The SARIMA model achieved an MAE of 12 mm and an RMSE of 18 mm, while the SANN model showed lower error values of 9 mm MAE and 14 mm RMSE. These results indicate that the SANN model was more effective in rainfall prediction, yielding lower errors than the SARIMA model [4].

Mugaloglu and Kilic (2021) utilized SARIMA-GARCH models to forecast unemployment rates in G7 countries. In their study, the seasonal unemployment data of each country were analyzed, and SARIMA and GARCH models were combined to generate short-term forecasts. The model's performance was assessed using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics. The results indicated that the SARIMA-GARCH models yielded an average MAE of 0.8 and an RMSE of 1.2 in unemployment rate forecasts. These findings demonstrate that the SARIMA-GARCH model effectively forecasts economic data characterized by seasonal fluctuations [5].

Falatouri et al. (2022) compared SARIMA and LSTM models for demand forecasting in the retail supply chain. Both models were tested on retail demand data, and their performance was evaluated using MAE and RMSE metrics. According to the results, the SARIMA model had an MAE of

125, while the LSTM model achieved a lower MAE of 95. Similarly, the LSTM model showed approximately 20% lower error rates in RMSE compared to the SARIMA model. These findings suggest that LSTM provides more accurate and effective results in demand forecasting [6].

Ciftci and Sir (2023) combined the SARIMA model with machine learning methods to predict the number of emergency room visits. The study used 396 days of "patient visit counts" data from a hospital in Ankara. The SARIMA and hybrid models were applied, and their results were compared. The analysis revealed that the SARIMA model provided higher accuracy. As a result, the proposed model was shown to predict emergency room visit volumes successfully and could be an effective tool in healthcare service planning [7].

In their 2023 study, Banda et al. used a hybrid transfer learning approach to forecast electric vehicle (EV) profiles and system voltage with limited charging data. They combined traditional models with transfer learning techniques to uncover data relationships, assessing accuracy through Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics. The hybrid model achieved 3.5V MAE and 5V RMSE, highlighting its effectiveness in accurately predicting EV profiles and system voltage, even with constrained data availability [8].

Kayran and U. Araz (2023) used the Box-Jenkins and Prophet methods to model passenger demand on the 127 Erdek-Bandırma line of the Balıkesir Metropolitan Municipality. The analysis compared both methods using MAPE, RMSE, and R^2 performance metrics. The SARIMA method showed smaller MAPE and RMSE values and a more considerable R^2 value than the Prophet method. These results indicate that the SARIMA method performed better on the dataset [9].

Agyemant et al. (2023) focused on the ability of SARIMA and Prophet models to process time series data with seasonal components. The study used monthly data from 2013 to 2018, with 2019 as the test period, and applied the Box-Jenkins method to develop various models. The findings revealed that the SARIMA $(0,1,1)x(1,0,0)_{12}$ model outperformed the Prophet model. Statistical tests, such as Ljung-Box and Box-Pierce, confirmed the adequacy of the SARIMA model, showing that it was independent and uncorrelated. Based on the high forecasting accuracy of the proposed SARIMA model, the study recommends its use in analyzing road traffic accidents in Ghana [10].

S. Saglam and Cavdur (2023) employed ARIMA and Artificial Neural Networks (ANN) models to predict the availability of parking spaces. By analyzing historical parking data, the study made short-term forecasts and compared the performance of ARIMA and ANN models. The results showed that the ARIMA model had an MAE of 10 and an RMSE of 15, while the ANN model achieved MAE and RMSE values of 7 and 10, respectively. These findings indicate that the ANN model outperformed the ARIMA model in terms of accuracy in parking space predictions [11].

Kumari and Muthulakshmi (2024) utilized the SARIMA model to forecast weather by analyzing historical data on temperature, humidity, and precipitation for short-term predictions. The model's accuracy was assessed using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The SARIMA model achieved an MAE of 1.5°C and an RMSE of 2.3°C, indicating its effectiveness in forecasting weather data with seasonal patterns [12].

As seen in Table 1, the reviewed studies show that models such as SARIMA, ARIMA, GARCH and LSTM are widely used in time series analysis and forecasting for seasonal and complex data. These studies show that SARIMA and hybrid models provide high accuracy for data influenced by seasonal effects, while deep learning methods such as LSTM and Artificial Neural Networks (ANN) perform better in short-term forecasts. Findings across various domains, from electric vehicles and weather forecasting to energy consumption and unemployment rates, underscore the critical role of data structure in model selection. Consequently, SARIMA and hybrid models effectively predict seasonal and complex time series data.

The objectives of this study were (1) to analyze monthly electric vehicle sales data recorded in the United States between 2021 and 2023 to forecast future market trends, (2) to use the SARIMA model to obtain reliable forecasts by considering seasonal fluctuations and long-term trends, (3) to identify anticipated increases and seasonal effects in electric vehicle sales, contributing to production planning, supply chain management and strategic decision-making processes, and (4) to evaluate the limitations of the model and provide recommendations for future research.

Table 1. Literature Research on Demand Forecasting

Reference No	Forecasting Method	Forecasting Horizon	Data Set Coverage Period	Forecasting Performance Assessment Criteria
[3]	ARIMA, GARCH, Holt-Winters, ANN	24 hours	2017	MAPE,RMSE
[4]	SARIMA, ANN	3 years	2007-2013	FE, MFE, MSE, RMSE
[5]	SARIMA, GARCH	2 years	1995-2019	MAPE, RMSE
[6]	SARIMA, LTSM	1 month	2017-2019	MAPE, RMSE
[7]	SARIMA, Holt-Winters, ANN	7-15-30 days	396 days	R2, MAPE
[8]	Hybrid Transfer Learning, Method	6200 hours	365 days	RMSE, MAPE, R2
[9]	SARIMA, PROPHET	5 months	2020-2022	MAPE, RMSE, R2
[10]	SARIMA, Facebook Prophet	1 year	2013-2018	MAE, MAPE, MSE, RMSE
[11]	ARIMA, SARIMA, ANN	1 day	2011-2013	-
[12]	SARIMA	1 year	2012-2022	-
[13]	SARIMA	30 days	275 days	-

3 Materials and Methods

The rapid growth of the electric vehicle (EV) market has made reliable future forecasts more critical than ever. In this study, as presented in Table 2, we utilized the 'Monthly Light-Duty Electric Vehicle Sales' report published by the U.S. Department of Energy’s Vehicle Technologies Office. This dataset spans the period from January 2021 to December 2023 and includes monthly sales figures for each year. This study used the Python programming language and the Google Colab platform. The list of Python libraries used for SARIMA analysis of EV sales forecasting are given below.

- **pandas**: For data manipulation and handling time series data.
- **numpy**: For numerical operations and array handling.
- **matplotlib**: For visualizing the historical and forecasted sales data along with prediction intervals.
- **statsmodels**: Specifically for implementing the SARIMA model (statsmodels.tsa.statespace.sarimax.SARIMAX) and statistical tests like the Augmented Dickey-Fuller (ADF) test.
- **scikit-learn**: For calculating model evaluation metrics, such as Mean Squared Error (MSE).
- **warnings**: To suppress irrelevant warnings and improve readability of the output.

These libraries are commonly used for time series analysis and provide a comprehensive toolset for statistical modeling, data preprocessing, and visualization.

The data source is constantly updated by the US Department of Energy. This consistent EV sales reporting routine establishes a solid foundation for time series analysis, enabling robust and reliable future forecasts.

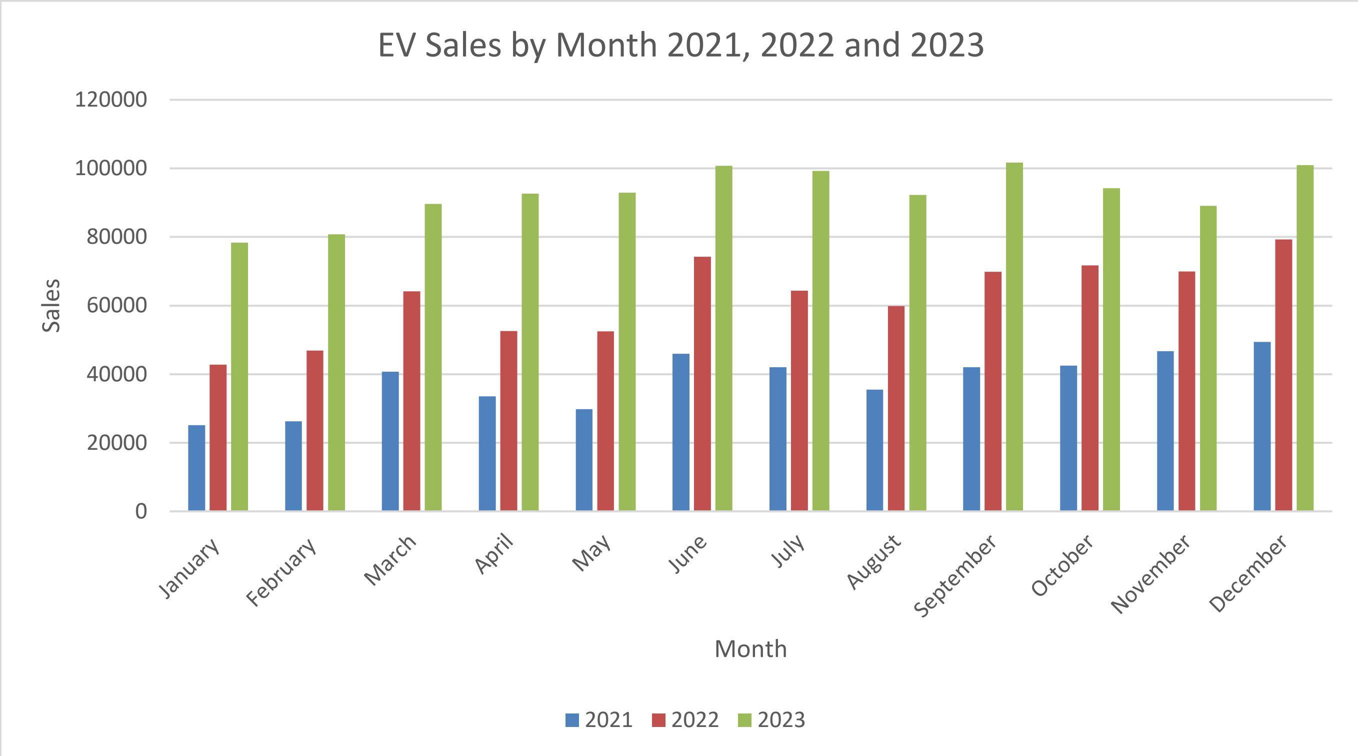

Figure 1. Graphical Representation of Monthly Light Commercial Electric Vehicle Sales from 2021 to 2023

Table 2. Monthly Light Commercial Electric Vehicle Sales from 2021 to 2023

Month	2021	2022	2023
January	25.103	42.780	78.382
February	26.215	46.859	80.759
March	40.755	64.160	89.648
April	33.547	52.537	92.631
May	29.796	52.502	92.897
June	45.913	74.262	100.745
July	42.013	64.310	99.259
August	35.499	59.836	92.277
September	42.020	69.811	101.719
October	42.485	71.739	94.192
November	46.687	69.924	89.082
December	49.441	79.262	100.982

To facilitate a more straightforward examination of the data, it has been presented in graphical form in Figure 1. This graph illustrates that electric vehicle sales are subject to seasonal fluctuations throughout the year. Specifically, sales significantly increase during the summer months (June, July, August) and decrease during the winter months (January, February). This suggests that consumer behavior varies with the seasons, with vehicle purchases being more popular in the summer. However, looking at the overall trend for each year, it is noteworthy that despite seasonal fluctuations, EV sales showed a significant increase in 2023.

This study follows a multi-stage analytical process to forecast future sales in the U.S. electric vehicle market. In the initial stage, monthly electric vehicle sales data from January 2021 to December 2023 were collected from U.S. Department of Energy reports. The collected data were prepared for time series analysis, and necessary stationarity checks were conducted for the SARIMA model that accounts for seasonal effects. The Augmented Dickey-Fuller (ADF) test was applied to determine whether the data was stationary, and differencing was performed as needed. The ADF test is a standard linear unit root test proposed by Dickey and Fuller. The Augmented Dickey-Fuller (ADF) test is a statistical method used to determine if a time series is stationary or possesses a unit root, which would indicate non-stationarity. This test helps in assessing whether the data maintains consistent statistical properties over time or whether its behavior changes, impacting forecasting models that assume stationarity. Stationarity means the time series has a constant mean, variance, and autocorrelation over time. At this stage, first-order differencing was applied to the time series. Differencing is a method to make the data stationary by removing trends present in the time series [18]. First-order differencing involves calculating the differences between consecutive data points. First order differencing transformation $\nabla y_t = y_t - y_{t-1}$ [13]. Subsequently, the optimal parameters for the seasonal and non-seasonal components of the SARIMA model were optimized using a grid search method. The model's performance was evaluated using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC). These criteria assess both the model's fit and complexity. The parameter combinations with the lowest AIC, BIC, and HQIC values are generally considered the most suitable. These criteria were used to evaluate model performance and determine the best-fitting parameters. The best-performing model was selected in the final stage, and monthly sales forecasts for 2024-2025 were generated. This comprehensive analytical process provides objective and systematic forecasts of future electric vehicle sales trends.

3.1 ARIMA Model

The ARIMA model, also known as the Box-Jenkins method, is a statistical technique used for forecasting. The Box-Jenkins approach predicts the future values of univariate time series data. It provides a systematic method for creating forecasting models from discrete and stationary time series data collected at equal time intervals [7]. ARIMA models are widely used models that explain the variability in the series with past period effects and randomness [14].

Figure 2. AR-I-MA Model

Applying the ARIMA model begins with collecting the dataset and performing exploratory analysis. Next, the stationarity of the time series is checked, and differencing is applied if necessary. The model's parameters p, d, and q are determined using ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots to build the appropriate model. The model's performance is then evaluated using AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). Finally, forecasts are made, and these forecasts are tested for accuracy and reported.

3.2 SARIMA Model

The SARIMA model, which extends the ARIMA model to incorporate seasonal effects, is generally expressed as SARIMA (p, d, q) (P, D, Q). Compared to the ARIMA model, SARIMA includes four additional parameters to account for the seasonal component of the time series [20]. These parameters are **p, d**, and **q**: These represent the autoregressive order, differencing order, and moving average order of the trend component, respectively. **P, D, Q**: These are the orders of the seasonal autoregressive, seasonal differencing, and seasonal moving average components [13] **m**: This parameter indicates the number of time steps in a single seasonal period. The SARIMA model is expressed as given in Equation 1 below:

$$\emptyset_p(B^m)\emptyset(B)\nabla^D\nabla^d x_t = O_Q(B^m)\theta(B)w_t \tag{1}$$

x_t is the non-stationary time series,

w_t is the Gauss white noise process,

$\emptyset_p(B^m)$ and $O_Q(B^m)$ are autoregressive seasonal component and its moving average.

∇^D and ∇^d show the differing components of the seasonal and trend components.

B is the back shift operator.

$\emptyset(B)$ and $\theta(B)$ are polynomials representing the trend component's autoregressive and moving average processes.

Time series subject to SARIMA model structure Box-Jenkins method is utilised for estimation. This method has three stages: identification, estimation and testing, and application. In the identification stage, in order to make the variance of the series constant, the logarithm of the relevant series is taken and the transformation process is applied. Then, with the help of various tests, it is examined whether the series are stationary or not. Differencing is applied to non-stationary series until they become stationary [15]. Thus, the popularity of this model seems to be due to its ability to adapt very well to different time series models.

4 Results and Discussion

Initially, the study examined whether the monthly light commercial electric vehicle sales data for 2021-2023, as presented in Table 1, are stationary. This step is crucial for ensuring the accuracy and validity of the SARIMA model, as the SARIMA model is designed to work with stationary data. Stationarity refers to the property of a time series where the mean, variance, and autocorrelation structure remain constant over time. If the data is not stationary, trends and seasonal effects can complicate the modeling process and negatively impact the model's forecasting accuracy.

Initially, the ADF test was applied to the raw data, and the results indicated that the dataset was non-stationary. This finding suggests a trend or seasonality in the data, which must be addressed before direct modeling. Table 3 presents the results of the ADF test for both the raw and the differenced data.

The ADF test results indicated that the raw data was non-stationary; thus, the data needed to be differenced before proceeding with the modeling process. For the raw data, the ADF test statistic was -1.560 with a p-value of 0.497, suggesting that the data was non-stationary. Compared to the ADF test's critical values, the test statistic was above the critical values, so the null hypothesis of non-stationarity could not be rejected.

Table 3. ADF Test Result

	ADF Test Statistic	p-value	Critical Value (1%)	Critical Value (5%)	Critical Value (10%)
Raw Data	-1.560	0.497	-3.449	-2.870	-2.571
1. Differencing Data	-5.233	0.000	-3.449	-2.870	-2.571

After applying differencing, the dataset was re-evaluated with the Augmented Dickey-Fuller (ADF) test, and the results indicated that the data had become stationary. The ADF test statistic for the differenced data was -5.233, with a p-value of 0.000 below the 0.05 threshold. This confirms that the differencing process made the data suitable for SARIMA modeling. In contrast, the ADF test statistic for the raw data was -1.560, which did not fall below the critical values at the 1%, 5%, or 10% levels, indicating that the series was non-stationary. However, after differencing, the test statistic value of -5.233 was below all critical values, confirming that the series is stationary.

Optimizing the model parameters became crucial for ensuring the model's best performance in the next step. Parameter optimization involves trying various combinations to select parameter values that minimize errors and provide the best model fit. This process enhances the model's predictive accuracy and identifies parameters that best fit the data structure. Selecting the correct parameters improves the model's reliability and performance [17].

The SARIMA model, which incorporates seasonal and non-seasonal components, requires careful parameter optimization to deliver optimal results. This study used a grid search method to determine the parameters for the SARIMA model. Grid search systematically explores combinations of parameters (p, d, q) and (P, D, Q, m) to identify the optimal set. During the grid search, the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan- Quinn Information Criterion (HQIC) were calculated for each parameter combination.

Table 4. SARIMA Model Parameters

Parameters	AIC	BIC	HQIC
(1, 0, 1)(1, 0, 1, 12)	742.30	752.68	746.35
(1, 0, 2)(1, 0, 1, 12)	739.51	749.01	742.82
(2, 0, 2)(1, 0, 1, 12)	741.02	751.52	745.33

Different SARIMA model parameters were tested, and the performance of each model was evaluated based on the AIC, BIC, and HQIC criteria. The model with the lowest AIC, BIC, and HQIC values was considered the most suitable. The results presented in Table 4 indicate that the parameters (1, 0, 2) and (1, 0, 1, 12) provided the best performance.

Based on the optimized SARIMA model parameters, monthly electric vehicle (EV) sales forecasts between January 2024 and December 2025 revealed a specific growth trend and seasonal fluctuations. Table 5 provides the forecasted monthly EV sales and a summary of these predictions.

The forecasts indicate that from January 2024 to December 2025, electric vehicle (EV) sales are expected to increase overall, reflecting a growth trend in the market. Notably, sales are projected to peak at 108.559 units in July 2024, after which a slight decline towards the end of the year is anticipated. This decline suggests seasonal fluctuations throughout the year, though overall high sales levels are expected to persist. The confidence intervals of the forecasts reflect the uncertainties of the model and provide valuable insights into the market dynamics. The results are visually presented in Figure 3 for a more precise interpretation.

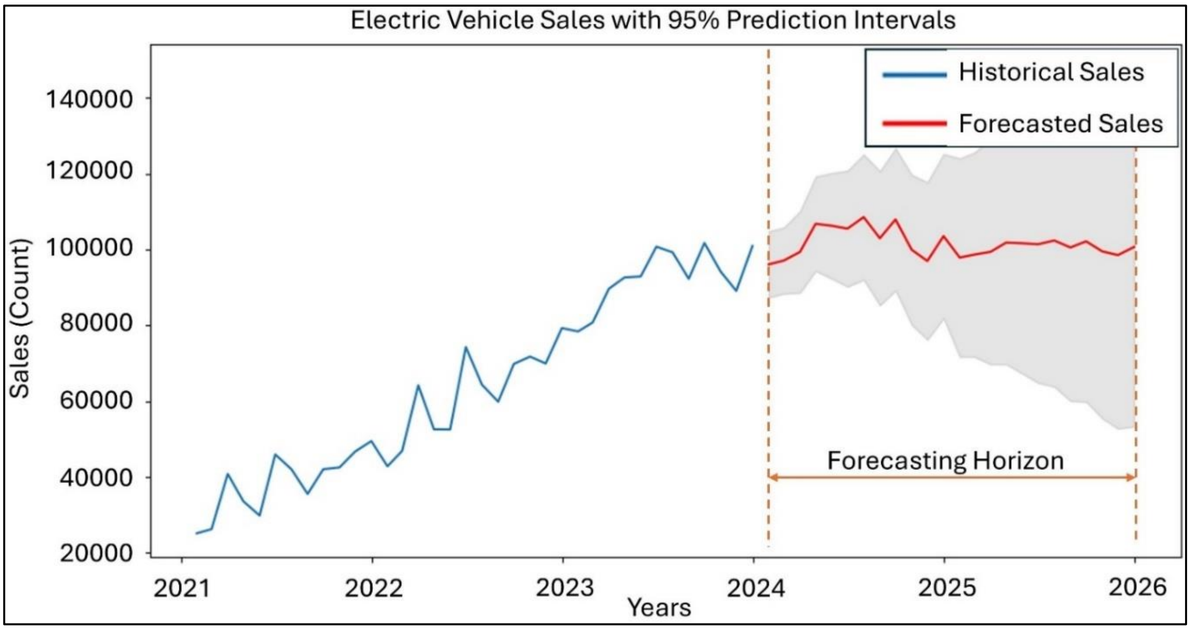

Figure 1. EV Sales Forecasts

Table 5. EV Sales Forecasts for 2024-2025

	Predicted_Sales	Lower CI	Upper CI
31.01.2024	96.076	87.314	104.839
29.02.2024	97.092	88.342	105.842
31.03.2024	99.355	88.580	110.130
30.04.2024	106.792	94.316	119.267
31.05.2024	106.260	92.290	120.230
30.06.2024	105.513	90.194	120.831
31.07.2024	108.559	92.001	125.117
31.08.2024	102.976	85.266	120.685
30.09.2024	107.945	89.154	126.736
31.10.2024	99.932	80.119	119.745
30.11.2024	96.970	76.185	117.754
31.12.2024	103.538	81.823	125.254
31.01.2025	97.875	71.663	124.087
28.02.2025	98.655	71.676	125.635
31.03.2025	99.391	69.722	129.060
30.04.2025	101.833	69.699	133.967
31.05.2025	101.647	67.224	136.070
30.06.2025	101.390	64.823	137.957
31.07.2025	102.384	63.792	140.977
31.08.2025	100.533	60.017	141.048
30.09.2025	102.161	59.810	144.512
31.10.2025	99.508	55.398	143.618
30.11.2025	98.521	52.720	144.321
31.12.2025	100.676	53.246	148.107

The forecast results align with the seasonal patterns observed in the historical data set. Sales tend to be higher in the spring and summer, indicating increased consumer demand. A detailed analysis of seasonal fluctuations is crucial for understanding market dynamics.

The graph compares historical and forecasted sales, showing that seasonal fluctuations will continue, with distinct periods of increase and decrease in sales throughout the year. Notably, the higher sales in the summer months indicate that consumer demand is more robust during this period.

The projected sales figures provide critical information for strategic planning in the electric vehicle (EV) market. The anticipated growth trend could significantly guide manufacturers in optimizing their capacity planning, supply chain management, and marketing strategies.

The forecasted results suggest that the growth in the EV market is sustainable and is supported by supportive government policies, technological innovations, and increasing environmental awareness. Industry stakeholders should be aware of seasonal fluctuations and prepare for high-demand periods, aligning their strategic planning accordingly. These forecasts offer valuable insights into future growth opportunities and potential challenges.

This study serves as a critical guide for stakeholders seeking to understand the future development of the electric vehicle market and to develop appropriate strategies. In this context, the forecasts provided by the SARIMA model serve as a reference point for academic research and support practical applications in the industry. Given the dynamic nature of the electric vehicle market, the potential of such a model to offer advanced analysis and reduce uncertainties in the sector is notably high.

5 Conclusions

The electric vehicle (EV) market has significantly transformed in recent years due to technological advancements,increased environmental wareness, and government incentives. This study aimed to forecast future trends in the market by analyzing monthly EV sales data from the US for the years 2021-2023. The focus of the study was the SARIMA (Seasonal Autoregressive Integrated Moving Average) model, which successfully captured both the seasonal cycles and long-term growth trends in the EV market, providing meaningful insights for the sector.

The research findings indicate a stable growth trend in EV sales, which will continue over the next two years. Specifically, a sales forecast of approximately 96.076 units is projected for January 2024, with an anticipated peak in July 2024, reaching 108.559 units due to increased demand during the summer months. These forecasts reflect the seasonal cycles and consumer behavior in the sector while highlighting the overall market trends. For instance, sales are expected to decrease towards the end of the year, with a forecasted drop to 100.676 units by December 2025 due to seasonal effects.

The results indicate that the electric vehicle (EV) market growth will continue, highlighting how consumer demand responds to seasonal changes. The increase in sales during the summer months suggests that consumers are more inclined to upgrade or purchase new vehicles during this period. Furthermore, the EV market is being more widely adopted by consumers, driven by technological advancements, battery cost decreases, and charging infrastructure improvements.

One of the strengths of this study is the successful application of the SARIMA model in capturing seasonal and trend components from historical data, enabling accurate forecasts of future trends. The optimization of the model using criteria such as AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), and HQIC (Hannan-Quinn Information Criterion) has improved the accuracy of the forecasts and ensured the selection of the most suitable parameters considering the complexity of the data set. As a result, the forecasts can guide decision-makers in various aspects, including production planning, supply chain management, sales strategies, and resource allocation.

However, the study has some limitations. The SARIMA model does not account for potential structural changes or sudden market shocks in the future. For example, unexpected economic crises, changes in government policies, or significant technological breakthroughs could negatively impact the validity of the model's predictions. Additionally, the model is based solely on US data, needing a broader perspective on global market dynamics. Future research could address these limitations using more diverse data sets encompassing different geographies.

In conclusion, this study provides significant insights into understanding the growth potential of the electric vehicle market and forecasting future trends. The findings offer valuable insights for strategic decision-makers, supply chain managers, and investors in the EV sector, contributing to the development of sustainable growth strategies.

However, to keep forecasts current and accurate in a rapidly changing industry, models must be continuously updated and various external factors considered. Future studies could enhance the validity of predictions by incorporating broader data sets and alternative forecasting methods, offering a more comprehensive view of market dynamics.

Reference

[1] H. Ozbay, C. Kozkurt, A. Dalcali, and M.Tektas, “Geleceğin ulaşım tercihi: Elektrikli araçlar.”

[2] H. Kaymaz and Y. Hançar, “Elektrikli Araç Batarya Yönetim Sistemleri için Hücre Eşitleme Yöntemleri,” Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi, vol. 4, no. 1, pp. 59–73, Apr. 2021, doi: 10.51513/jitsa.897126.

[3] Y. Kim, H. Son, and S. Kim, “Short Term Electricity Load Forecasting for Institutional Buildings,” Energy Reports, vol. 5, pp. 1270–1280, Nov. 2019, doi: 10.1016/j.egyr.2019.08.086.

[4] C. C. Nwokike, B. C. Offorha, M. Obubu, C. B. Ugoala, and H. I. Ukomah, “Comparing SANN and SARIMA

for Forecasting Frequency of Monthly Rainfall in Umuahia," Sci Afr, vol. 10, p. e00621, Nov. 2020, doi: 10.1016/j.sciaf.2020.e00621.

[5] E. Mugaoglu and E. Kilic, "G7 Ülkeleri İşsizlik Oranı Tahminleri:SARIMA-GARCH Model Karşılaştırması," Journal of Yaşar University, vol. 16, no. 61, pp. 228–247, Jan. 2021, doi:10.19168/jyasar.803807.

[6] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, "Predictive Analytics for Demand Forecasting – A Comparison of SARIMA and LSTM in Retail SCM," Procedia Comput Sci, vol. 200, pp. 993–1003, 2022, doi: 10.1016/j.procs.2022.01.298.

[7] S. Ciftci and G. D. Batur Sir, "An Application for Forecasting The Number of Applications to the Emergency Department with Time Series Analysis and Machine Learning Methods," Pamukkale University Journal of Engineering Sciences, vol. 29, no. 7, pp. 667–679, 2023, doi: 10.5505/pajes.2022.18488.

[8] P. Banda, M. A. Bhuiyan, K. N. Hasan, and K. Zhang, "Assessment of hybrid transfer learning method for forecasting EV profile and system voltage using limited EV charging data," Sustainable Energy, Grids and Networks, vol. 36, p. 101191, Dec. 2023, doi: 10.1016/j.segan.2023.101191.

[9] G. Kayran, O. Uzun Araz, M. Celal Bayar Üniversitesi, M. Fakültesi, and E. Mühendisliği Bölümü, "Toplu Taşıma Yolcu Talebinin SARIMA ve PROPHET Yöntemleri İle Tahmini," Endüstri Mühendisliği / Journal of Industrial Engineering, vol. 34, no. 3, pp. 334–356, 2023, doi: 10.46465/endustrimuhendisligi.1281456.

[10] E. F. Agyemang, J. A. Mensah, E. Ocran, E. Opoku, and E. N. N. Nortey, "Time series based road traffic accidents forecasting via SARIMA and Facebook Prophet model with potential changepoints," Heliyon, vol.9, no.12, p. e22544,Dec.2023,doi:10.1016/j.heliyon.2023.e22544.

[11] A. Sebatlı Sağlam and F. Çavdur, "Prediction Of Parking Space Availability Using ARIMA And Neural Networks," Endüstri Mühendisliği, vol. 34,no.1,pp.86–108,Apr.2023,doi:10.46465 /endustrimuhendisligi.1241453.10.46465

[12] S. Kumari and P. Muthulakshmi, "SARIMA Model: An Efficient Machine Learning Technique for Weather Forecasting," Procedia Comput Sci, vol. 235, pp. 656–670, 2024, doi: 10.1016/j.procs.2024.04.064.

[13] J. Liu, Z. Zhao, Y. Zhong, C. Zhao, and G. Zhang, "Prediction of the dissolved gas concentration in power transformer oil based on SARIMA model," Energy Reports, vol. 8, pp. 1360–1367, Aug. 2022, doi: 10.1016/j.egyr.2022.03.020.

[14] Akay, E. Ç., Topal, K. H., Kizilarslan, S., Bulbul, "Türkiye Konut Fiyat Endeksi Öngörüsü: ARIMA, Rassal Orman ve ARIMA-RASSAL ORMAN." PressAcademia Procedia, 10(1), 7-11, 2019, doi:10.17261/Pressacademia. 2019.1134.

[15] Tayyar, A. E., "Türkiye'de Politik Parasal Dalgalanmalar: 1986-2016 Dönemi İçin SARIMA Analizi", Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi, 9(5), 1381-1401, 2020, doi: 10.18506/ anemon.761615.

[16] Akgül, I., "Zaman Serisi Analizi ve Öngörü Modelleri", Öneri Dergisi, 1(1), 52-69, 1994, doi:10.14783/maruoneri.698511.

[17] Yüceer, M., Atasoy, I., Berber, R., "A software for parameter estimation in dynamic models", Brazilian J. of Chem. Engineering, Cilt. 25(4), 813 – 821, 2008.

[18] Huang, W., Li, Y., Zhao, Y., Zheng, L., "Time Series Analysis and Prediction on Bitcoin", BCP Business & Management, 34, 1223-1234, 2022, doi: 10.54691/bcpbm.v34i.3163.

[19] İnce, M., Taşdemir, Ç., "Forecasting Retail Sales for Furniture and Furnishing Items through the Employment of Multiple Linear Regression and Holt–Winters Models", Systems, 12(6), 1-29, 2024, doi: 10.3390/systems12060219.

[20] Yuan, Z., Jia, G., "Systematic investigation of keywords selection and processing strategy on search engine forecasting: a case of tourist volume in Beijing", Information Technology & Tourism, 24(4), 547-580, 2022, doi: 10.1007/s40558-022-00238-5.