

# Supply Chain and Operations Decision Making

www.scodm.reapress.com

Sup. Chain Oper. Decis. Mak. Vol. 2, No. 4 (2025) 173–190.

Paper Type: Original Article

## Production Forecasting in the Automotive Industry: Performance Analysis of Time Series Models Using the Box-Jenkins Method

Seyed Kamal Chaharsoughi<sup>1\*</sup> , Niloofar Manzari Vahed<sup>1</sup>, Hassan Ashnavar<sup>2</sup>

<sup>1</sup> Department of Industrial Engineering, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran; skch@modares.ac.ir; Niloofar.Manzari@modares.ac.ir.

<sup>2</sup> Head of Production Control Department, SAIPA Automotive Group, Tehran, Iran; ashnavar@saipacorp.com.

### Citation:

Received: 01 Jun 2025  
Revised: 11 August 2025  
Accepted: 20 November 2025

Chaharsoughi, S. K., Manzari Vahed, N., & Ashnavar, H. (2025). Production forecasting in the automotive industry: Performance analysis of time series models using the Box-Jenkins method. *Supply chain and operations decision making*, 2(4), 173-190.

### Abstract

In the realm of industrial engineering and management, production forecasting is a pivotal topic that profoundly influences productivity and the optimization of production processes. By developing a precise prediction model, companies can significantly enhance production planning, control production, minimize stoppages, optimize inventories, and boost machinery productivity. This study delves into the production processes of the SAIPA Automotive Group and proposes an accurate prediction model. Using statistical time series forecast models, including the Box-Jenkins method and the Rolling forecast approach, the study reveals that these models, particularly the Auto Regressive Integrated Moving Average eXogenous (ARIMAX) model, excel at daily production prediction. Conversely, the Auto Regressive Integrated Moving Average (ARIMA) and Autoregressive (AR) models demonstrate superior efficiency in trend-based production prediction. Additionally, the Seasonal Auto Regressive Integrated Moving Average (SARIMA) and Seasonal Auto Regressive Integrated Moving Average eXogenous (SARIMAX) models performed worse. The application of time series tools and the rolling forecast approach has also led to a notable reduction in model errors.

**Keywords:** Production forecasting, Time series models, Box-Jenkins, Rolling forecasting.

## 1 | Introduction

In today's highly competitive global markets, timely and cost-effective production is increasingly important. It requires the collaborative efforts of many stakeholders to develop effective strategies for supply chain management, production planning and scheduling, and the control and allocation of equipment and labor [1].

 Corresponding Author: skch@modares.ac.ir

 <https://doi.org/10.48313/scodm.v2i4.42>

 Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

Production planning is a managerial process that involves designing and planning an organization's production activities to optimize the use of available resources and ensure the efficient production of high-quality products and services [2].

In industrial engineering and management, production forecasting is a critical topic in production planning, significantly impacting productivity and the optimization of production processes. By developing an accurate forecasting model, companies can enhance production planning and control, reduce production time, minimize stoppages, optimize inventories, and increase machinery productivity [3], [4].

All forecasting tasks must address issues such as what to forecast, the forecasting method, the level of aggregation, the data used, and the forecast accuracy. Forecasting can be at various levels of aggregation, ranging from fully aggregated to semi-aggregated or fully disaggregated forecasts for short- and long-term periods. The data used in forecasting can be based on historical production, sales, arrivals, or reservations. Additionally, forecasts should be adjusted for specific events, such as holidays and special occasions [5].

In the SAIPA Automotive Group, six companies, Saipa, Pars Khodro, Saipa Citroën, Bonro, Zamyad, and Saipa Diesel, engage in the daily production of various car models. As with any manufacturing unit, the SAIPA Automotive Group has established a Master Production Schedule (MPS) from the outset of operations, taking into account available capacity for machinery and equipment, labor hours, raw materials, market demand, and other factors. This MPS is then used to formulate monthly and weekly plans, with the weekly production plan detailing daily production targets.

However, in practice, production often deviates from the planned schedule, underscoring the necessity for accurate production forecasting to aid decision-makers and managers. This research aims to examine production in the SAIPA Automotive Group and to present an accurate production forecasting model. To this end, statistical time-series forecasting models, including the Box-Jenkins method and the Rolling forecasting approach, have been used to forecast production for the six car manufacturing companies within the SAIPA Automotive Group.

This paper is organized as follows: Section 2 reviews the related literature. Section 3 describes the research methodology, including the methods used for data collection and preprocessing, implementation of forecasting models, model evaluation, and the results obtained. Section 4 presents the discussion and conclusions of the research.

## 2 | Literature Review

Time series forecasting models use historical data to predict future values. These models include univariate linear or nonlinear time series models and multivariate linear and nonlinear time series models. *Table 1* briefly reviews the literature that has employed time-series forecasting models. Based on the studied articles, among the Box-Jenkins models, the Auto Regressive Integrated Moving Average (ARIMA) model has been most frequently used for forecasting, which is certainly related to the dataset's characteristics. Additionally, in recent years, hybrid models that combine several time series models have become more prevalent.

**Table 1. Examples of research on solving forecasting problems using time series models.**

	Reference	Description And Application	Solution Method	Data Characteristics
1	Singh and Mishra [6]	Retail price forecasting	Box–Jenkins	Univariate (with exogenous variable)
2	Ghomi and Forghani [7]	Forecasting the expected number of passengers using two techniques, Box-Jenkins and artificial neural networks. The superior performance of Box-Jenkins models is demonstrated.	Box–Jenkins and ANN	Univariate (with exogenous variable)
3	Ni et al. [8]	Forecasting metro passenger flow. A parametric and convex optimization approach, called Optimization and Prediction with Combined Loss Function (OPL), is proposed to combine SARIMA models jointly.	SARIMA + OPL	Univariate (with exogenous variable)
4	Milenkovic et al. [9]	Forecasting passenger flow	ARIMA and SARIMA	Univariate
5	Chen et al. [10]	Forecasting public transportation (bus) travel demand	ARMA and ARIMA	Univariate
6	Su and Ye [11]	Forecasting public transportation (bus) travel demand	ARMA and ARIMAX	Univariate (with exogenous variable)
7	Ghauri et al. [12]	Evaluating two econometric models for forecasting imports and exports for the Fiscal Year (FY) 2020	Box–Jenkins	Univariate (with exogenous variable)
8	Guleryuz [13]	Forecasting disease spread. Among box-Jenkins models, ARIMA, ETS, and RNN-LSTM are employed. ARIMA, with the lowest AIC values, is selected as the best model for the total number of cases, total case growth rate, number of new cases, total deaths, total death growth rate, and number of new deaths.	Box–Jenkins, ETS, and RNN-LSTM	Univariate (with exogenous variable)
9	Hadwan et al. [14]	Forecasting the expected number of cancer patients	ARIMA and BPNN	Multivariate (with exogenous variable)
10	Yasmin and Moniruzzaman [15]	Forecasting agricultural area, production, and yield in Bangladesh	Box–Jenkins	Univariate (with exogenous variable)

The time series models used in this research are briefly introduced below.

## 2.1 | Autoregressive

The Autoregressive (AR) model is a linear model that describes the relationship between a time series's previous values and its current value. In this model, the current value of the time series is modeled as a linear combination of its previous ( $p$ ) values and a random variable with mean zero (usually noise or error). The formula for the AR( $p$ ) model is as follows:

$$X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t + C \quad (1)$$

Here:  $X_t$  represents the current value of the time series.  $c$  is a constant.  $\varphi_i$  are the model parameters that indicate the dependency of previous values on the current value.  $\varepsilon_t$  is a random variable with a mean of zero, representing noise.

## 2.2 | Moving Average

The Moving Average (MA) is one of the primary models in time series analysis used for modeling and forecasting time series data. This model averages the values of sequential noise or errors in the time series. In an MA (q) model, the current value of the time series is modeled as a linear combination of the previous (q) values of noise or errors and a random variable with a mean of zero. The formula for the MA (q) model is as follows:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} = \mu + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (2)$$

Here:  $X_t$  represents the current value of the time series.  $\mu$  is the mean of the time series.  $\varepsilon_t$  is a random variable with a mean of zero, representing the noise or error of the model.  $\theta_i$  are parameters indicate the dependency of previous noise or error values on the current value.

## 2.3 | Auto Regressive Integrated Moving Average

The Auto Regressive Integrated Moving Average (ARIMA) model is a linear model composed of two main components: the AR and MA models. In this model, the relationship between the previous values of the time series and the MA of the sequential noise or errors is modeled. The ARIMA model is an extension of the ARMA model, incorporating three main elements: AR, MA, and I (integration). This model is used for modeling and forecasting non-stationary time series and provides more accurate predictions by fitting parameters to real data. The formula for this model is equivalent to the following expression:

$$X_t = \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (3)$$

## 2.4 | Seasonal Auto Regressive Integrated Moving Average

The Seasonal Auto Regressive Integrated Moving Average (SARIMA) model is an extension of the ARIMA model used for modeling time series with seasonal patterns. This model uses AR, MA, I (integration), and seasonal AR and MA components to model and forecast time series with seasonal patterns.

## 2.5 | Auto Regressive Integrated Moving Average eXogenous

The Auto Regressive Integrated Moving Average eXogenous (ARIMAX) model is an extension of the ARIMA model that incorporates external or exogenous variables for time series modeling. This model allows for the modeling of the effect of external variables on the time series.

## 2.6 | Seasonal Auto Regressive Integrated Moving Average eXogenous

The Seasonal Auto Regressive Integrated Moving Average eXogenous (SARIMAX) model is an extension of the SARIMA model that incorporates exogenous variables to model seasonal patterns in time series. This model uses AR, MA, and I components, seasonal AR and MA components, and external variables to model time series with seasonal patterns and external-variable influences.

## 2.7 | Model Performance Evaluation Criteria

To evaluate model performance and select the best model, evaluation criteria from the time series literature, including Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Log-Likelihood (LLF), have been used. AIC and BIC are criteria for evaluating the quality of a statistical model. These criteria compare the model's simplicity with the accuracy of its predictions. Lower AIC and BIC values indicate a better, simpler model. LLF is the logarithm of the likelihood function, which indicates how well the model fits the data. Higher LLF values indicate a better fit of the model to the data. In addition to the AIC, BIC, and LLF performance evaluation criteria used in statistical models, this study uses Mean Absolute Error (MAE) which is the average absolute error between predicted and actual values; Mean Squared Error (MSE) which is the average squared error between predicted and actual values; Root Mean Squared Error (RMSE)

which is the square root of the average squared error between predicted and actual values; and R-squared which shows the ratio of the MSE to the mean squared actual values, to evaluate the prediction models.

### 3 | Research Methodology

The Box-Jenkins method is a well-known approach to time series analysis and forecasting, used to model and predict time series data. In this study, the Box-Jenkins method is employed to implement time series models. This method first fits an ARIMA model to the time series under investigation and then provides better predictions by estimating the model parameters from the actual data. The research method is shown in Fig. 1. The main steps of the Box-Jenkins method are as follows:

#### Data understanding

In this step, the time series under investigation is analyzed to identify various structures such as stationarity, trend changes, seasonal patterns, and random values. In this study, after data cleaning, additive models and exponential smoothing are used to plot decomposed time series charts to identify trends and seasonal patterns. Additionally, to examine the stationarity of the time series and determine the appropriate time lag, the Dickey-Fuller statistical test, autocorrelation, and correlation coefficient plots are used.

#### Model implementation and parameter estimation

After understanding the dataset, ARIMA model parameters  $p$ ,  $d$ , and  $q$  are calculated using methods such as autocorrelation functions, statistical tests, or the auto-ARIMA model. In this study, the initial parameters of the time series models are determined by implementing and running the auto-ARIMA model. Auto-ARIMA is a model that uses grey search and local optimization methods to find the best combination of hyperparameters for univariate time series models. This method searches for the best time series model that fits the data within the specified range of hyperparameter values, and reports the best model. Given the use of local optimization and top-down search concepts, the auto-ARIMA model may also identify simpler models with fewer coefficients than the reported model for the time series data. Therefore, the upper bound of the ARIMA model coefficients is determined using this model, and then models with smaller coefficients are implemented and executed.

#### Model validation and evaluation

Time series forecasting models are created using the estimated parameters. In this study, all time series models, including AR, MA, ARMA, ARIMA, SARIMA, ARIMAX, and SARIMAX, are implemented and evaluated over the hyperparameter range. To validate and evaluate the model, residual correlation tests and normality checks of the error distribution are performed using correlation and Q-Q plots of the model errors.

#### Selecting the best model

At this stage, after fine-tuning the hyperparameters, the best models are identified based on evaluation criteria, including AIC, BIC, LLF, MSE, RMSE, and MAE.

#### Forecasting and conclusion

Using the final time-series model, future time-series predictions are provided, and the model's validity is assessed. In this study, to examine the best model's results and improve its accuracy, a rolling (moving) forecast approach is used.

### 3.1 | Data Understanding

The data used in this study includes the daily production statistics of companies in the SAIPA automotive group, encompassing daily production characteristics, annual production plans, and weekly production plans for Saipa, Pars Khodro, Saipa Citroën, Bonro, Zamyad, and Saipa Diesel from the beginning of 2023 to the end of December 2024. Fig. 2 shows the production trend over time for the companies in the SAIPA automotive group.

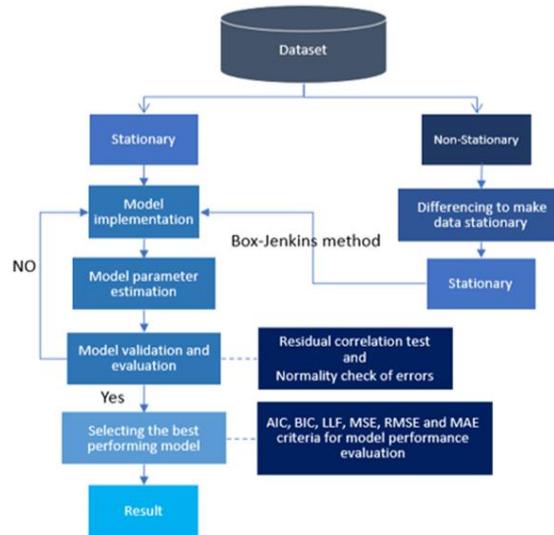


Fig. 1. Research method structure.

By examining the covariance and correlation values shown in Fig. 3, contrary to expectations, there is no strong positive correlation between the production plan and the production of SAIPA group companies, especially at Zamyad and Saipa Diesel. It indicates that these companies have fallen short of their production targets. Additionally, negative correlations between the production or production plans of one company and another suggest a high likelihood of unfairness in the supply chain in some cases.

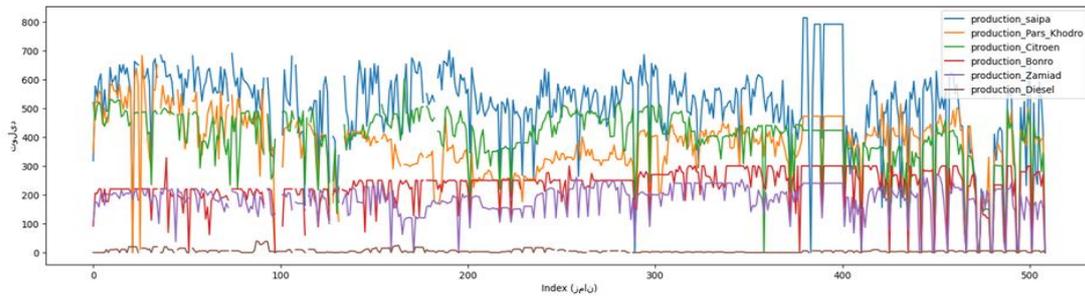


Fig. 2. Production trend over time for SAIPA automotive group companies.

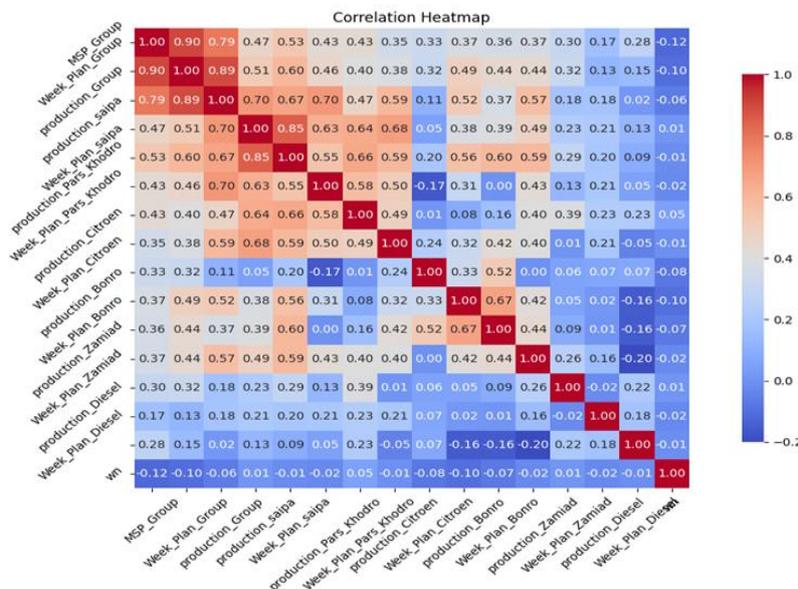


Fig. 3. Covariance values between production plan and production for SAIPA group companies.

### 3.1.1 | Data cleaning

Data cleaning is the process of identifying and handling errors, duplicates, and irrelevant data from a raw dataset to prepare it for analysis. Generally, data cleaning involves three steps: 1) removing duplicate data, 2) identifying and handling missing data, and 3) identifying and handling outliers. The operations performed to clean the dataset used in this study are briefly explained below.

- I. Identifying duplicate data: the dataset contains no duplicates.
- II. Identifying and handling missing data: the missing data identified in the dataset pertains to holidays during the specified time period when the factories were not in production. Given the time series nature of the dataset, two scenarios are considered: 1) Missing values are replaced with zero due to factory closures. This scenario preserves the information and knowledge within the data by recording the actual event in the dataset. Still, it may complicate data patterns, potentially affecting the accuracy of predictive models, and 2) Missing values are replaced with the production value of the nearest date after the holiday, based on the assumption that no production occurs on holidays. This argument reduces data complexity and may increase the accuracy of predictive models.

In this study, a combination of these two scenarios is used to handle missing values. During official national holidays, missing values are set to 0. However, on days when one or more factories were closed for any reason, these values are replaced with the next period's value in the time series.

- III. Identifying and handling outliers: given that data collection was conducted with great care and reviewed multiple times, no erroneous records are present in this dataset, and all records accurately reflect real events. Therefore, outliers identified by statistical tests contain useful information about production trends in the companies, and no action has been taken regarding these data in this study.

### 3.1.2 | Identifying seasonal patterns in time series

In the literature, the behavioral patterns or change models of a time series are divided into four components: trend, cycle, season, and irregular variations or residuals. By plotting the time series against time, these components can be distinguished, leading to a better understanding of the time series data. Identifying seasonal patterns in time series data can be done using several different methods. In this study, additive models, which consider the time series as the sum of three components: trend, seasonal patterns, and random values, and the Local Regression (LOESS) method, which decomposes the time series into seasonal and trend components using local smoothing, are used to discover and identify seasonal patterns. *Fig. 4* shows the trends and patterns discovered in the daily production dataset of the Saipa company using additive models and exponential smoothing.

According to the charts plotted by these models, the dataset does not exhibit a clear upward or downward trend and shows a nearly sinusoidal behavior over time. Therefore, regarding seasonality, as indicated, the data has been identified to have seasonal patterns. This pattern is particularly associated with additive charts for weekend holidays. Irregular variations are changes caused by random and unpredictable factors. In research, these values are typically removed from the time series and treated as outliers. Still, in this study's dataset, for the reasons previously mentioned, no action has been taken regarding these values.

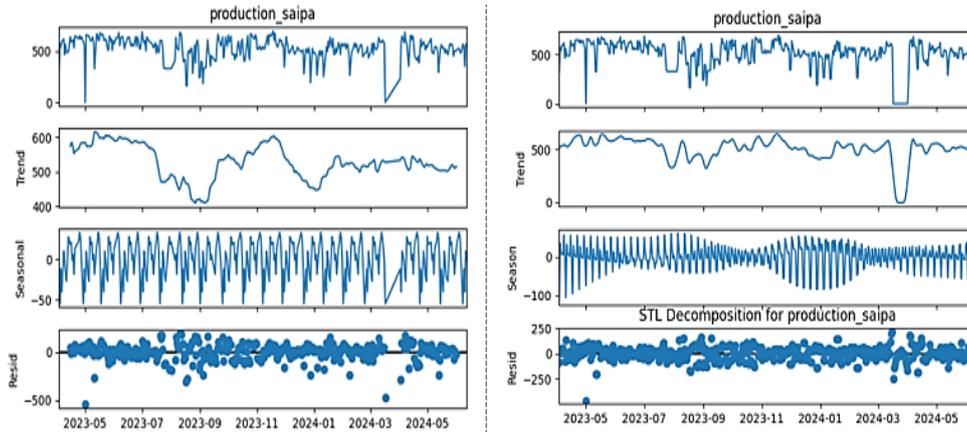


Fig. 4. Display of production trend, seasonal patterns, and residuals using the exponential smoothing model (right) and additive model (left) for the Saipa company.

### 3.1.3 | Stationarity

Stationarity of a time series is a necessary condition for achieving reliable forecasts. There are several statistical tests for examining the stationarity of time series data; the most popular is the Dickey-Fuller test. The Dickey-Fuller test formula generally examines the presence of a unit root in a time series. This test can be performed in various forms, but one of the most common forms, the basic Dickey-Fuller test model, is as follows:

$$Y_t = \alpha + \beta t + \rho Y_{t-1} + \epsilon_t \tag{4}$$

$Y_t$  is the value of the time series at time  $t$ ,  $\alpha$  is the intercept,  $\beta$  is the trend coefficient,  $\rho$  is the unit root parameter,  $Y_{t-1}$  is the previous value of the time series, and  $\epsilon_t$  is the random error.

Test hypotheses: null hypothesis ( $\rho = 1$ )  $H_0$  the time series has a unit root and is non-stationary. Alternative hypothesis: ( $\rho < 1$ )  $H_1$  the time series is stationary and does not have a unit root. The Dickey-Fuller test statistic is calculated as follows:

$$FDA = \frac{1-\hat{\rho}}{(\hat{\rho})ES} \tag{5}$$

where:  $\hat{\rho}$  is the estimated unit root parameter (using regression), and  $(\hat{\rho})ES$  is the standard error of the estimate  $\hat{\rho}$ . In the augmented version, to control for autocorrelation, several lags are added to the model:

$$Y_t = \alpha + \sum_{i=1}^p \phi_i Y_{t-i} + \rho Y_{t-1} + \epsilon_t \tag{6}$$

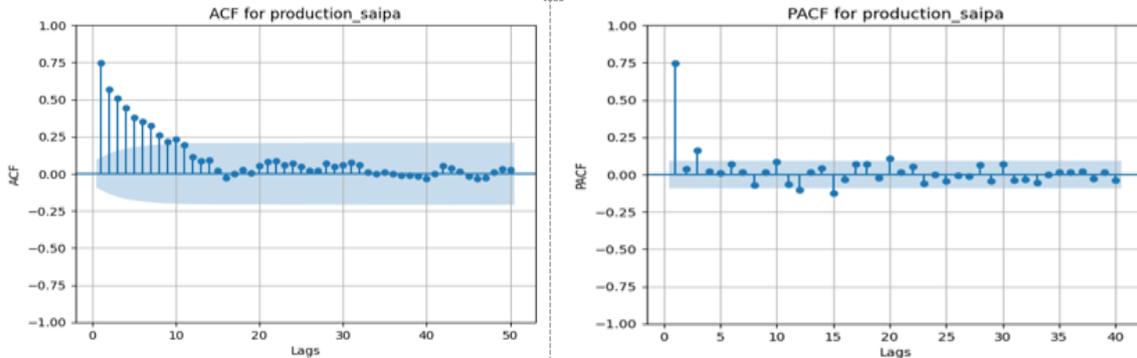
where  $p$  is the number of lags and  $\phi_i$  are the coefficients corresponding to the lags?

By calculating the test statistic and comparing it with the critical values, the null hypothesis can be rejected or accepted. If the test statistic is less than the critical value (more negative), the null hypothesis is rejected, and we conclude that the time series is stationary. *Table 2* shows the results of the Dickey-Fuller test. The results of the Dickey-Fuller test indicate that all production data columns (except Saipa Diesel, which is marginally significant at the 1% level) are stationary. It means that time-series models like Box-Jenkins can be used for forecasting.

After examining the stationarity of the time series, the lag can be determined. The lag in a time series is the time interval between observations over which meaningful relationships for predicting future values exist. There are various methods for identifying lag in a time series. In this study, the Dickey-Fuller test and plotting autocorrelation and partial autocorrelation charts have been used. *Fig. 5* shows the autocorrelation and partial autocorrelation values for the Saipa company.

**Table 2. Dickey-Fuller test results for examining dataset stationarity.**

Company	Dickey-Fuller Test Statistic	P-Value	Critical Region	Time Lag
Saipa	-5.5858	1.36E-06	{'1%': -3.4413, '5%': -2.86637, '10%': -2.56934}	3
Pars Khodro	-4.3346	0.000386	{'1%': -3.4412, '5%': -2.86636, '10%': -2.569337}	3
Saipa Citroën	-5.6880	8.19E-07	{'1%': -3.44133, '5%': -2.86638, '10%': -2.56935}	8
Bonro	-4.3631	0.000344	{'1%': -3.44127, '5%': -2.86636, '10%': -2.569337}	13
Zamyad	-4.8148	5.09E-05	{'1%': -3.44127, '5%': -2.86636, '10%': -2.56933}	5
Saipa Diesel	-3.9388	0.001766	{'1%': -3.4413, '5%': -2.86637, '10%': -2.56934}	10



**Fig. 5. Autocorrelation charts (with 10 significant time lags) and partial autocorrelation charts (with 15 significant time lags) for the Saipa Automotive Company.**

### 3.2 | Model Implementation and Parameter Estimation

For the implementation, training, parameter estimation, and performance evaluation of the predictive models, the data is split into two sets: training (75%) and test (25%). Then, by implementing and running the auto-ARIMA model, the initial parameter values for the time series models are determined. Using this model, the upper bound of the ARIMA model coefficients is determined, so that only the performance of models with smaller coefficients is examined.

The approach to model validation and evaluation, along with a detailed explanation of the process for selecting the best predictive model, is provided. *Tables 3-5* sequentially present the training and implementation results of the model, using the Box-Jenkins method, for predicting the daily production of six automotive group companies: single-target models without exogenous variables and single-target models with exogenous variables (weekly plan).

**Table 3. Selected single-target time series models without exogenous variables for predicting the SAIPA automotive group production.**

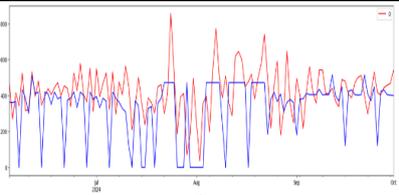
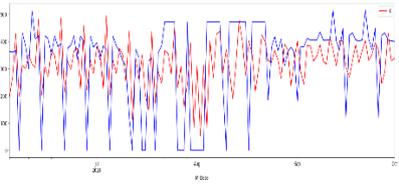
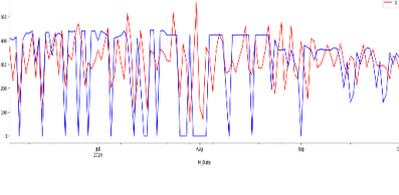
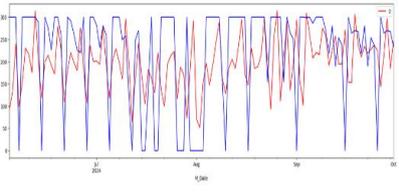
Company	Coefficient Table	Forecast Chart for Test Data	Evaluation Criteria Values
Saipa	<pre> SARIMAX Results ----- Dep. Variable:      y      No. Observations:  428 Model:              SARIMAX(4, 1, 5)  Log Likelihood: -2930.634 Date:               Mon, 06 Jan 2025    AIC:          5843.328 Time:               17:49:02           BIC:          5881.796 Sample:             04-03-2023        HQIC:         5857.251 Sample:             - 06-03-2024 Covariance Type:   opg ----- coef      std err      z      P&gt; z       [0.025      0.975] ----- ar.L1     -1.2758     0.104    -12.279   0.000    -1.479    -1.072 ar.L2     -1.1646     0.105    -11.081   0.000    -1.408    -0.921 ar.L3     -0.8660     0.105    -8.236    0.000    -1.100    -0.632 ar.L4     -0.6182     0.100    -6.195    0.000    -0.814    -0.423 ma.L1      0.3599     0.076     4.754    0.000     0.212     0.508 ma.L2      0.1352     0.080     1.698    0.090    -0.021     0.291 ma.L3     -0.2870     0.072    -3.986    0.000    -0.428    -0.146 ma.L4     -0.2084     0.076    -2.747    0.007    -0.357    -0.061 ma.L5     -0.8317     0.070   -11.921   0.000    -0.971    -0.697 sigma2     5.530e+04  4893.037  11.318   0.000    4.58e+04  6.5e+04 ----- Ljung-Box (L1) (Q):      3.40  Jarque-Bera (JB):      40.87 Prob(Q):                  0.07  Prob(JB):              0.00 Heteroskedasticity (H):  1.16  Skew:                  -0.71 Prob(H) (two-sided):     0.18  Kurtosis:              2.54                     </pre>		<p>MSE=84228.015</p> <p>RMSE=290.220</p> <p>MAE=257.60</p>
Pars khodro	<pre> SARIMAX Results ----- Dep. Variable:      y      No. Observations:  428 Model:              SARIMAX(4, 1, 3)  Log Likelihood: -2755.300 Date:               Mon, 06 Jan 2025    AIC:          5528.599 Time:               16:03:19           BIC:          5565.110 Sample:             04-03-2023        HQIC:         5543.020 Sample:             - 06-03-2024 Covariance Type:   opg ----- coef      std err      z      P&gt; z       [0.025      0.975] ----- Intercept  -0.1923     2.867    -0.067   0.947    -5.812     5.420 ar.L1     -1.4245     0.064   -22.088   0.000    -1.551    -1.298 ar.L2     -0.4634     0.119    -3.881   0.000    -0.697    -0.229 ar.L3      0.1220     0.117     1.045   0.296    -0.107     0.351 ar.L4     -0.1385     0.061    -2.253   0.024    -0.259    -0.018 ma.L1      0.8187     0.036    22.688   0.000     0.748     0.890 ma.L2     -0.6931     0.049   -14.066   0.000    -0.790    -0.597 ma.L3     -0.8774     0.035   -25.370   0.000    -0.945    -0.810 sigma2     2.971e+04  2969.736  10.004   0.000    2.30e+04  3.55e+04 ----- Ljung-Box (L1) (Q):      0.00  Jarque-Bera (JB):      17.46 Prob(Q):                  0.96  Prob(JB):              0.00 Heteroskedasticity (H):  0.60  Skew:                  -0.49 Prob(H) (two-sided):     0.00  Kurtosis:              2.86                     </pre>		<p>MSE=45189.1578</p> <p>RMSE=672.231</p> <p>MAE=449.285</p>
Saipa citroën	<pre> SARIMAX Results ----- Dep. Variable:      y      No. Observations:  428 Model:              SARIMAX(4, 1, 4)  Log Likelihood: -2784.264 Date:               Mon, 06 Jan 2025    AIC:          5588.527 Time:               18:34:18           BIC:          5639.095 Sample:             04-03-2023        HQIC:         5604.551 Sample:             - 06-03-2024 Covariance Type:   opg ----- coef      std err      z      P&gt; z       [0.025      0.975] ----- Intercept  -0.5232     4.936    -0.106   0.916   -10.196     9.153 ar.L1     -0.9648     0.199    -4.847   0.000    -1.355    -0.575 ar.L2      0.1322     0.329     0.402   0.688    -0.513     0.777 ar.L3      0.1634     0.184     0.890   0.374    -0.197     0.524 ar.L4     -0.3728     0.079    -4.710   0.000    -0.528    -0.218 ma.L1      0.3162     0.193     1.635   0.102    -0.063     0.695 ma.L2     -1.0605     0.178    -5.965   0.000    -1.415    -0.718 ma.L3     -0.4167     0.125    -3.333   0.001    -0.662    -0.172 ma.L4      0.5540     0.172     3.230   0.001     0.218     0.890 sigma2     3.451e+04  3518.430  9.807    0.000    2.70e+04  4.14e+04 ----- Ljung-Box (L1) (Q):      0.03  Jarque-Bera (JB):      27.87 Prob(Q):                  0.87  Prob(JB):              0.00 Heteroskedasticity (H):  1.07  Skew:                  -0.62 Prob(H) (two-sided):     0.70  Kurtosis:              3.08                     </pre>		<p>MSE= 24852.285</p> <p>RMSE=157.649</p> <p>MAE=116.2545</p>
Bonro	<pre> SARIMAX Results ----- Dep. Variable:      y      No. Observations:  428 Model:              SARIMAX(1, 0, 0)  Log Likelihood: -2572.114 Date:               Mon, 06 Jan 2025    AIC:          5150.227 Time:               20:51:43           BIC:          5162.405 Sample:             04-03-2023        HQIC:         5155.037 Sample:             - 06-03-2024 Covariance Type:   opg ----- coef      std err      z      P&gt; z       [0.025      0.975] ----- Intercept  141.5831     8.677    16.317   0.000    124.576    158.590 ar.L1      0.1964     0.850     0.230   0.819    -0.517     0.909 sigma2     9720.0660  1441.326  6.744    0.000  6895.118  1.25e+04 ----- Ljung-Box (L1) (Q):      0.04  Jarque-Bera (JB):      65.2 Prob(Q):                  0.83  Prob(JB):              0.00 Heteroskedasticity (H):  1.87  Skew:                  -0.9 Prob(H) (two-sided):     0.00  Kurtosis:              2.4                     </pre>		<p>MSE=15018.9059</p> <p>RMSE=122.5576</p> <p>MAE=115.9825</p>

Table 3. Continued.

Company	Coefficient Table	Forecast Chart for Test Data	Evaluation Criteria Values
Zamyad	<pre> SARIMAX Results ===== Dep. Variable:      y      No. Observations:  428 Model:             SARIMAX(1, 0, 0)  Log Likelihood: -2493.213 Date:             Mon, 06 Jan 2025  AIC: 4992.425 Time:             21:19:18  BIC: 5004.602 Sample:           04-03-2023  HQIC: 4997.234 Covariance Type:  opg ===== coef  std err  z      P&gt; z     [0.025   0.975] ----- Intercept  116.0183   6.765   17.148   0.000   102.751   129.270 ar.l1      0.2869    0.050    4.103   0.000    0.188    0.386 sigma2     6717.8233  836.711   8.029   0.000  5077.899  8357.747 ===== Ljung-Box (L1) (Q):      0.81  Jarque-Bera (JB):      48.4 Prob(Q):                 0.91  Prob(JB):              0.00 Heteroskedasticity (H):  1.70  Skew:                 -0.71 Prob(H) (two-sided):     0.00  Kurtosis:              2.31                     </pre>		MSE=7969.475 RMSE=89.2719 MAE=79.6305
Saipa diesel	<pre> SARIMAX Results ===== Dep. Variable:      y      No. Observations:  428 Model:             SARIMAX(0, 1, 0)  Log Likelihood: -1331.341 Date:             Mon, 06 Jan 2025  AIC: 2604.682 Time:             21:22:01  BIC: 2608.739 Sample:           04-03-2023  HQIC: 2666.285 Covariance Type:  opg ===== coef  std err  z      P&gt; z     [0.025   0.975] ----- sigma2     29.9816   0.855   34.966   0.000   28.226   31.578 ===== Ljung-Box (L1) (Q):      0.03  Jarque-Bera (JB):      1589.95 Prob(Q):                 0.86  Prob(JB):              0.00 Heteroskedasticity (H):  0.15  Skew:                 0.81 Prob(H) (two-sided):     0.00  Kurtosis:              12.45                     </pre>		MSE=13.9650 RMSE=3.7369 MAE=2.4685

Table 4. Selected single-target time series models with exogenous variables (ARIMAX) for predicting the SAIPA automotive group production.

Company	Coefficient Table	Forecast Chart for Test Data	Evaluation Criteria Values
Saipa	<pre> SARIMAX Results ===== Dep. Variable:      production_saipa  No. Observations:  428 Model:             ARIMA(4, 1, 5)  Log Likelihood: -2343.769 Date:             Mon, 06 Jan 2025  AIC: 4780.538 Time:             18:17:23  BIC: 4754.162 Sample:           04-03-2023  HQIC: 4727.164 Covariance Type:  opg ===== coef  std err  z      P&gt; z     [0.025   0.975] ----- Week_Plan_saipa  0.7848   0.012   67.725   0.000   0.761   0.807 ar.l1            -0.1824   0.728  -0.144   0.888  -1.530   1.225 ar.l2            -0.4321   0.308  -1.336   0.181  -1.016   0.192 ar.l3            -0.2946   0.538  -0.547   0.584  -1.350   0.761 ar.l4            0.2661   0.212   1.256   0.209  -0.149   0.681 ma.l1            -0.5203   0.727  -0.716   0.474  -1.945   0.904 ma.l2            0.1810   0.458   0.395   0.693  -0.716   1.078 ma.l3            -0.0543   0.433  -0.125   0.900  -0.903   0.704 ma.l4            -0.6801   0.366  -1.655   0.098  -1.324   0.112 ma.l5            0.1714   0.264   0.648   0.517  -0.347   0.689 sigma2          3471.3854  196.531  17.663   0.000  3086.192  3856.578 ===== Ljung-Box (L1) (Q):      0.00  Jarque-Bera (JB):      66.18 Prob(Q):                 0.96  Prob(JB):              0.00 Heteroskedasticity (H):  1.15  Skew:                 -0.37 Prob(H) (two-sided):     0.41  Kurtosis:              4.78                     </pre>		MSE=50939.139 RMSE=225.667 MAE=147.098
Pars khodro	<pre> SARIMAX Results ===== Dep. Variable:      production_Pars_Khodro  No. Observations:  428 Model:             ARIMA(3, 1, 3)  Log Likelihood: -2396.064 Date:             Mon, 06 Jan 2025  AIC: 4880.129 Time:             17:15:19  BIC: 4880.583 Sample:           04-03-2023  HQIC: 4820.948 Covariance Type:  opg ===== coef  std err  z      P&gt; z     [0.025   0.975] ----- Week_Plan_Pars_Khodro  0.7473   0.014   52.789   0.000   0.720   0.775 ar.l1            0.8091   0.235   3.444   0.000   0.344   1.274 ar.l2            0.8787   0.127   6.830   0.000   0.621   1.121 ar.l3            -0.1882   0.114  -1.645   0.100  -0.412   0.036 ma.l1            -0.0544   0.221  -0.247   0.805  -0.500   0.391 ma.l2            -0.9599   0.032  -29.675   0.000  -1.023  -0.897 ma.l3            0.6330   0.214   2.960   0.003   0.214   1.052 sigma2          4357.9904  184.893  23.570   0.000  3995.607  4720.374 ===== Ljung-Box (L1) (Q):      0.87  Jarque-Bera (JB):      638.33 Prob(Q):                 0.79  Prob(JB):              0.00 Heteroskedasticity (H):  0.24  Skew:                 -1.05 Prob(H) (two-sided):     0.00  Kurtosis:              8.61                     </pre>		MSE=31094.542 RMSE=176.336 MAE=150.572

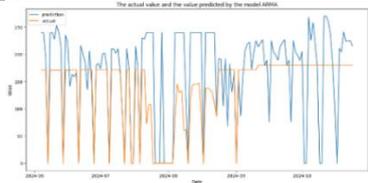
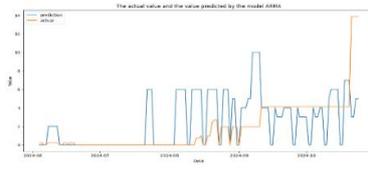
Table 4. Continued.

Company	Coefficient Table	Forecast Chart for Test Data	Evaluation Criteria Values
Saipa citroën	<pre> SARIMAX Results ===== Dep. Variable:   production_Citroen  No. Observations:   428 Model:          ARIMA(2, 1, 2)      Log Likelihood      -2301.074 Date:           Mon, 06 Jan 2025    AIC                 4614.148 Time:           20:39:24            BIC                 4638.488 Sample:         04-03-2023         HQIC                4623.762 Covariance Type: opg ===== coef    std err      z    P&gt; z     [0.025    0.975] ----- Week_Plan_Citroen ar.L1    0.6489    0.407    1.596    0.111    -0.148    1.446 ar.L2   -0.0928    0.165   -0.817    0.987    -0.327    0.321 ma.L1   -1.2051    0.409   -2.947    0.003   -2.006   -0.484 ma.L2    0.2535    0.370    0.686    0.493   -0.471    0.978 sigma2   2090.6763  112.059   24.868    0.000   2579.869  3021.484 ===== Ljung-Box (L1) (Q):    0.80  Jarque-Bera (JB):    358.92 Prob(Q):               1.00  Prob(JB):            0.00 Heteroskedasticity (H): 0.74  Skew:                -0.70 Prob(H) (two-sided):  0.07  Kurtosis:            7.22                     </pre>		<p>MSE= 547.079</p> <p>RMSE=97.709</p> <p>MAE=61.217</p>
Bonro	<pre> SARIMAX Results ===== Dep. Variable:   production_Bonro  No. Observations:   428 Model:          ARIMA(1, 0, 0)     Log Likelihood      -2186.323 Date:           Mon, 06 Jan 2025    AIC                 4228.646 Time:           21:09:49            BIC                 4236.882 Sample:         04-03-2023         HQIC                4227.058 Covariance Type: opg ===== coef    std err      z    P&gt; z     [0.025    0.975] ----- const    5.0132    6.791    0.741    0.459   -8.276    18.343 Week_Plan_Bonro ar.L1    0.8207    0.023   35.868    0.000    0.775    0.867 sigma2  1101.5779  42.108   26.160    0.000  1019.047  1184.109 ===== Ljung-Box (L1) (Q):    0.20  Jarque-Bera (JB):    2368.51 Prob(Q):               0.65  Prob(JB):            0.00 Heteroskedasticity (H): 0.29  Skew:                -2.32 Prob(H) (two-sided):  0.00  Kurtosis:            13.53                     </pre>		<p>MSE=5848.7371</p> <p>RMSE=76.4770</p> <p>MAE=49.706</p>
Zamyad	<pre> SARIMAX Results ===== Dep. Variable:   production_Zamyad  No. Observations:   428 Model:          ARIMA(1, 0, 0)     Log Likelihood      -2218.419 Date:           Mon, 06 Jan 2025    AIC                 4444.839 Time:           21:51:55            BIC                 4461.075 Sample:         04-03-2023         HQIC                4451.251 Covariance Type: opg ===== coef    std err      z    P&gt; z     [0.025    0.975] ----- const   21.8228    6.765    3.255    0.001    8.763   35.283 Week_Plan_Zamyad ar.L1    0.6826    0.017   39.552    0.000    0.649    0.716 sigma2  1858.4529  81.867   22.701    0.000  1697.996  2018.909 ===== Ljung-Box (L1) (Q):    6.41  Jarque-Bera (JB):    173.53 Prob(Q):               0.01  Prob(JB):            0.00 Heteroskedasticity (H): 3.55  Skew:                -0.12 Prob(H) (two-sided):  0.00  Kurtosis:            6.11                     </pre>		<p>MSE= 4749.805</p> <p>RMSE=68.9188</p> <p>MAE=48.5793</p>
Saipa diesel	<pre> SARIMAX Results ===== Dep. Variable:   production_Diesel  No. Observations:   428 Model:          ARIMA(0, 1, 0)     Log Likelihood      -1306.642 Date:           Mon, 06 Jan 2025    AIC                 2617.285 Time:           21:55:28            BIC                 2625.398 Sample:         04-03-2023         HQIC                2620.489 Covariance Type: opg ===== coef    std err      z    P&gt; z     [0.025    0.975] ----- Week_Plan_Diesel sigma2   26.6358    0.773   34.458    0.000   25.120   28.158 ===== Ljung-Box (L1) (Q):    0.25  Jarque-Bera (JB):    1832.68 Prob(Q):               0.62  Prob(JB):            0.00 Heteroskedasticity (H): 0.12  Skew:                -0.44 Prob(H) (two-sided):  0.00  Kurtosis:            13.11                     </pre>		<p>MSE=10.6592</p> <p>RMSE=3.2648</p> <p>MAE=1.9755</p>

**Table 5. Selected single-target time series models with exogenous variables (SARIMAX) for predicting the SAIPA automotive group production.**

Company	Coefficient Table	Forecast Chart for Test Data	Evaluation Criteria Values
Saipa	<p>SARIMAX Results</p> <pre> Dep. Variable: production_saipa No. Observations: 428 Model: SARIMAX(4, 1, 5)(4, 1, 5, 15) Log Likelihood: -2333.977 Date: Mon, 06 Jan 2025 AIC: 4667.954 Time: 18:26:32 BIC: 4748.375 Sample: 04-03-2023 HQIC: 4699.764 - 06-03-2024  Covariance Type: opg  coef std err z P&gt; z  [0.025 0.975] ----- Week_Plan_saipa 0.7862 0.0819 48.786 0.000 0.7480 0.8244 ar.L1 -1.1352 0.678 -1.674 0.094 -2.465 0.194 ar.L2 -0.8035 0.520 -0.807 0.995 -1.822 1.015 ar.L3 1.1592 0.617 1.896 0.071 -0.098 2.799 ar.L4 0.9373 0.419 2.239 0.025 0.117 1.758 ma.L1 0.2874 5.934 0.048 0.961 -11.343 11.918 ma.L2 -1.0227 7.359 -0.138 0.991 -15.436 13.411 ma.L3 -1.1727 2.345 -0.598 0.617 -5.768 3.423 ma.L4 0.8479 5.168 0.089 0.993 -10.866 10.162 ma.L5 4.7777 4.777 1.000 0.318 -0.512 10.212 ar.S.L15 -1.6977 1.021 -1.663 0.096 -3.698 0.303 ar.S.L30 -1.6727 1.793 -0.933 0.351 -5.187 1.842 ar.S.L45 -1.1769 1.153 -1.488 0.139 -3.067 0.553 ar.S.L60 -0.8115 0.587 -1.681 0.109 -1.805 0.182 ma.S.L15 0.7666 2.688 0.284 0.769 -4.344 5.877 ma.S.L30 0.8428 3.322 0.253 0.990 -4.468 0.552 ma.S.L45 0.1628 3.241 0.050 0.968 -6.189 6.515 ma.S.L60 -0.8436 3.011 -0.280 0.779 -6.746 5.059 ma.S.L75 0.7368 3.648 0.202 0.889 -6.684 5.222 sigma2 5094.2497 2.8e+04 0.182 0.856 -4.98e+04 6e+04                 </pre> <p>Ljung-Box (L1) (Q): 21.55 Jarque-Bera (JB): 64.55  Prob(Q): 0.00 Prob(JB): 0.00  Heteroskedasticity (H): 1.06 Skew: -0.44  Prob(H) (two-sided): 0.73 Kurtosis: 4.73</p>		<p>MSE=53531.336</p> <p>RMSE=231.368</p> <p>MAE=152.981</p>
Pars khodro	<p>SARIMAX Results</p> <pre> Dep. Variable: production_Pars_Khodro No. Observations: 428 Model: SARIMAX(3, 1, 3)(3, 1, 3, 15) Log Likelihood: -2341.176 Date: Mon, 06 Jan 2025 AIC: 4710.351 Time: 17:48:33 BIC: 4766.646 Sample: 04-03-2023 HQIC: 4732.619 - 06-03-2024  Covariance Type: opg  coef std err z P&gt; z  [0.025 0.975] ----- Week_Plan_Pars_Khodro 0.7466 0.0815 51.482 0.000 0.718 0.775 ar.L1 -1.5586 0.853 -29.137 0.000 -1.663 -1.454 ar.L2 -0.3898 0.895 -3.261 0.001 -0.496 -0.124 ar.L3 0.3572 0.849 7.291 0.000 0.261 0.453 ma.L1 1.8019 0.858 21.950 0.000 0.984 1.108 ma.L2 -0.7475 0.847 -15.566 0.000 -0.819 -0.056 ma.L3 -0.9849 0.847 -19.351 0.000 -0.997 -0.813 ar.S.L15 -0.2248 4.784 -0.847 0.963 -9.681 9.153 ar.S.L30 0.8387 3.536 0.235 0.814 -6.899 7.760 ar.S.L45 0.8638 0.287 0.220 0.826 -0.499 0.625 ma.S.L15 -0.8724 4.635 -0.188 0.851 -9.956 8.211 ma.S.L30 -0.9869 0.116 -0.122 0.903 -16.894 14.920 ma.S.L45 0.8691 3.826 0.227 0.820 -6.630 8.369 sigma2 4463.1933 1920.130 2.324 0.020 699.888 8226.579                 </pre> <p>Ljung-Box (L1) (Q): 0.22 Jarque-Bera (JB): 293.68  Prob(Q): 0.64 Prob(JB): 0.00  Heteroskedasticity (H): 0.26 Skew: -0.91  Prob(H) (two-sided): 0.00 Kurtosis: 6.71</p>		<p>MSE=11833.48</p> <p>RMSE=108.782</p> <p>MAE=69.5627</p>
Saipa Citroën	<p>SARIMAX Results</p> <pre> Dep. Variable: production_Citroen No. Observations: 428 Model: SARIMAX(2, 1, 2)(2, 1, 2, 15) Log Likelihood: -2248.084 Date: Mon, 06 Jan 2025 AIC: 4516.089 Time: 20:46:28 BIC: 4556.219 Sample: 04-03-2023 HQIC: 4531.914 - 06-03-2024  Covariance Type: opg  coef std err z P&gt; z  [0.025 0.975] ----- Week_Plan_Citroen 0.7955 0.0811 70.589 0.000 0.773 0.818 ar.L1 0.9188 0.221 4.129 0.000 0.479 1.343 ar.L2 -0.8968 0.123 -0.787 0.431 -0.338 0.144 ma.L1 -1.4808 0.215 -6.893 0.000 -1.901 -1.059 ma.L2 0.4025 0.286 2.393 0.017 0.889 0.896 ar.S.L15 0.5838 3.198 0.158 0.875 -5.764 6.772 ar.S.L30 -0.8181 0.198 -0.851 0.959 -0.398 0.377 ma.S.L15 -1.3987 3.230 -0.433 0.665 -7.730 4.933 ma.S.L30 0.4241 3.845 0.139 0.889 -5.544 6.392 sigma2 2931.8657 224.755 13.045 0.000 2491.355 3372.377                 </pre> <p>Ljung-Box (L1) (Q): 0.08 Jarque-Bera (JB): 185.41  Prob(Q): 0.78 Prob(JB): 0.00  Heteroskedasticity (H): 0.75 Skew: -0.56  Prob(H) (two-sided): 0.09 Kurtosis: 6.09</p>		<p>MSE=5351.3364</p> <p>RMSE=231.368</p> <p>MAE=152.9816</p>
Bonro	<p>SARIMAX Results</p> <pre> Dep. Variable: production_Bonro No. Observations: 428 Model: SARIMAX(1, 0, 0)(1, 0, 0, 15) Log Likelihood: -2187.197 Date: Mon, 06 Jan 2025 AIC: 4222.394 Time: 21:15:44 BIC: 4238.631 Sample: 04-03-2023 HQIC: 4228.887 - 06-03-2024  Covariance Type: opg  coef std err z P&gt; z  [0.025 0.975] ----- Week_Plan_Bonro 0.8352 0.0811 75.812 0.000 0.814 0.857 ar.L1 0.3273 0.022 14.641 0.000 0.284 0.371 ar.S.L15 0.8076 0.050 0.154 0.878 -0.890 0.105 sigma2 1186.8652 42.378 26.188 0.000 1023.086 1189.125                 </pre> <p>Ljung-Box (L1) (Q): 0.29 Jarque-Bera (JB): 2482.38  Prob(Q): 0.59 Prob(JB): 0.00  Heteroskedasticity (H): 0.28 Skew: -2.38  Prob(H) (two-sided): 0.00 Kurtosis: 13.79</p>		<p>MSE=5921.850</p> <p>RMSE=76.953</p> <p>MAE=49.4894</p>

Table 5. Continued.

Company	Coefficient Table	Forecast Chart for Test Data	Evaluation Criteria Values
Zamyad	<pre> SARIMAX Results ===== Dep. Variable:      production_Zamiad  No. Observations:      428 Model:              SARIMAX(1, 0, 0)(1, 0, 0, 15)  Log Likelihood          -2223.834 Date:               Mon, 06 Jan 2025  AIC          4455.668 Time:               21:58:46  BIC          4471.905 Sample:             04-03-2023  HQIC         4462.081                    - 06-03-2024 Covariance Type:    opg =====                    coef  std err  z      P&gt; z     [0.025  0.975] ----- Week_Plan_Zamiad   0.7442   0.014   58.372   0.000   0.686   0.742 ar.1               0.6545   0.033   19.827   0.000   0.590   0.719 ar.5.115          -0.0023   0.054   -0.238   0.820   -0.118   0.094 sigma2            1914.4419  85.515  22.387   0.000  1746.835  2082.049 ===== Ljung-Box (L1) (Q): 12.31  Jarque-Bera (JB): 198.66 Prob(Q):           0.00  Prob(JB):      0.00 Heteroskedasticity (H): 3.63  Skew:        -0.11 Prob(H) (two-sided): 0.00  Kurtosis:     6.33                     </pre>		MSE= 425.470 RMSE= 3.6577 MAE=51.297
Saipa Diesel	<pre> SARIMAX Results ===== Dep. Variable:      production_Diesel  No. Observations:      428 Model:              SARIMAX(0, 1, 0)(0, 1, 0, 15)  Log Likelihood          -1382.166 Date:               Mon, 06 Jan 2025  AIC          2768.332 Time:               22:02:11  BIC          2776.374 Sample:             04-03-2023  HQIC         2771.513                    - 06-03-2024 Covariance Type:    opg =====                    coef  std err  z      P&gt; z     [0.025  0.975] ----- Week_Plan_Diesel   0.2814   0.024   8.370   0.000   0.154   0.249 sigma2             48.0229  1.831   26.233   0.000  44.435  51.611 ===== Ljung-Box (L1) (Q): 0.79  Jarque-Bera (JB): 380.12 Prob(Q):           0.37  Prob(JB):      0.00 Heteroskedasticity (H): 0.14  Skew:         0.15 Prob(H) (two-sided): 0.00  Kurtosis:     7.70                     </pre>		MSE=150.50 RMSE=12.2678 MAE=10.6670

### 3.3 | Model Validation and Evaluation

To validate the models and examine the normality of the residual distribution, a Q-Q plot (Fig. 6) has been used. A Q-Q plot shows the quantiles of a distribution. To check the normality of the data using a Q-Q plot, if the points lie approximately on a straight line, it indicates that the data are normally distributed. Deviation from the straight line may indicate that the data distribution is not normal. By examining the plots, the error distributions in all production data columns (except Saipa Diesel, which is weaker at the 1% level) appear normal. Regarding Saipa Diesel's production, due to frequent holidays and almost no production variance, the predictive model is equal to the company's constant average production, which has also affected the error distribution. According to industry experts, this model has been accepted for predicting Saipa Diesel.

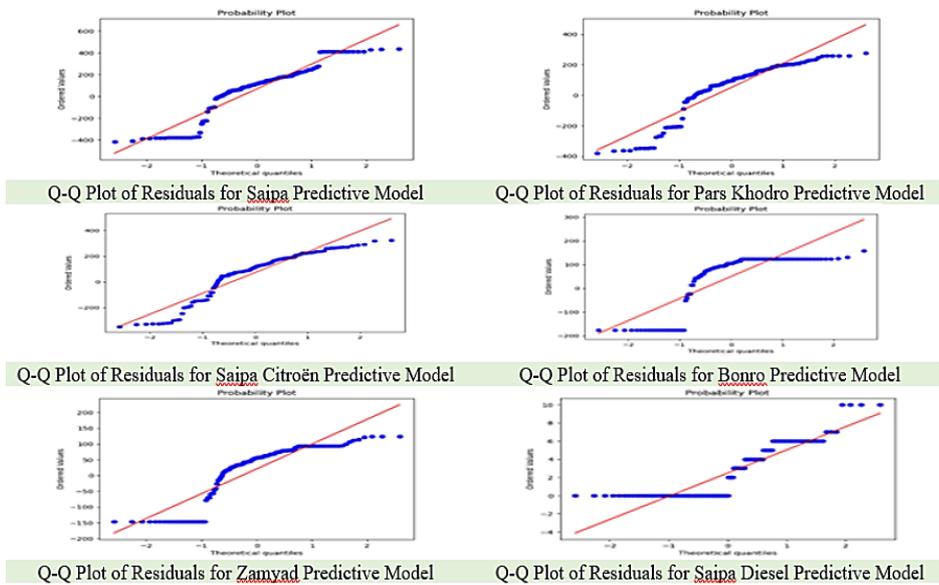


Fig. 6. Q-Q Plots of residuals for predictive models in the SAIPA automotive group companies.

### 3.4 | Selecting the Best Performing Model

To fully explain the process of selecting the best predictive model, the process for selecting the best-performing model for the Saipa company is provided. Using the Box-Jenkins method and the auto-ARIMA model, time series models have been trained on the dataset. The auto-ARIMA model reports the best performing model based on BIC and AIC values. The best-performing model for predicting SAIPA's daily production, based on the auto-ARIMA results, is the ARIMA (4,1,5) model, with evaluation criteria values of AIC=5841.227, BIC=5881.795, MSE=84228.0155, RMSE=290.220, and MAE=257.60.

Then, the proposed model and models with smaller coefficients are implemented on the dataset, and the best performing model is selected based on AIC, BIC, MAE, MSE, and RMSE criteria, as well as its simplicity. As shown in *Fig. 7* and *Fig. 8*, models with smaller coefficients have been executed on the dataset, and the evaluation criteria values for these models have been obtained. Compared to the proposed auto-ARIMA model, the ARIMA (4,1,5) model still performs best at predicting the data. Additionally, the predictive model with the exogenous variable (weekly plan) for each company was executed. The best-performing model for predicting Saipa's daily production is the ARIMAX (4,1,5) model, with evaluation criteria values of AIC=4709.518, BIC=4754.164, MSE=50939.139, RMSE=225.667, and MAE=147.098. Following this, the SARIMAX (4,1,5) model, with evaluation criteria values of AIC=4667.954, BIC=4748.375, MSE=53531.336, RMSE=231.368, and MAE=152.981, performs worse. This process has also been used to select the best-performing model for other companies.

```
Auto_ARIMA (4,1,5) Model Information: AIC=5841.22766423708, BIC=5881.795504369366, LLF=-2910.61383211854
AR (1,0,0) Model Information: AIC=5987.714305927808, BIC=5995.827873954265, LLF=-2991.857152963904
AR (2,0,0) Model Information: AIC=5961.032564027664, BIC=5973.20291606735, LLF=-2977.516282013832
AR (3,0,0) Model Information: AIC=5947.820274608606, BIC=5964.047410661521, LLF=-2969.910137304303
MA (0,0,1) Model Information: AIC=5870.586387574196, BIC=5878.699955600653, LLF=-2933.293193787098
MA (0,0,2) Model Information: AIC=5870.10545643893, BIC=5882.2758084786155, LLF=-2932.052728219465
MA (0,0,3) Model Information: AIC=5868.648654231038, BIC=5884.8757902839525, LLF=-2930.324327115519
ARIMA (1,1,1) Model Information: AIC=5869.610984906441, BIC=5881.781336946127, LLF=-2931.8054924532207
ARIMA (2,1,2) Model Information: AIC=5869.610984906441, BIC=5881.781336946127, LLF=-2931.8054924532207
ARIMA (3,1,3) Model Information: AIC=5869.610984906441, BIC=5881.781336946127, LLF=-2931.8054924532207
```

Fig. 7. Evaluation criteria values for AIC, BIC, and LLF.

```
Auto_ARIMA (4,1,5) validation_metrics:
r2= -0.49279136078162167 mse= 84228.01551677268 rmse= 290.2206324794512 mae= 257.59889041424776
None
AR (1,0,0) validation_metrics:
r2= -1.3550599844134354 mse= 132879.94164585962 rmse= 364.5270108590852 mae= 331.9571083988821
None
AR (2,0,0) validation_metrics:
r2= -0.7628467483858967 mse= 99465.39561898005 rmse= 315.38134951036665 mae= 288.1231392875356
None
AR (3,0,0) validation_metrics:
r2= -0.7647001680484549 mse= 99569.97142521683 rmse= 315.5470985846912 mae= 288.1903746548931
None
MA (0,0,1) validation_metrics:
r2= -0.07304471590679507 mse= 60544.58068022681 rmse= 246.05808395626187 mae= 207.76272901143716
None
MA (0,0,2) validation_metrics:
r2= -0.07144903080376088 mse= 60454.54707395785 rmse= 245.87506395313423 mae= 207.37736224519284
None
MA (0,0,3) validation_metrics:
r2= -0.08506443583652423 mse= 61222.771339247236 rmse= 247.43235709835372 mae= 210.20843964241755
None
ARIMA (1,1,1) validation_metrics:
r2= -0.07371964280099053 mse= 60582.6621927613 rmse= 246.13545496892823 mae= 207.85412400005396
None
ARIMA (2,1,2) validation_metrics:
r2= -0.07237401324123427 mse= 60506.73751204797 rmse= 245.98117308454314 mae= 207.32911891502846
None
ARIMA (3,1,3) validation_metrics:
r2= -0.07478657685084289 mse= 60642.86199031325 rmse= 246.25771458030152 mae= 207.91718355002098
.....
```

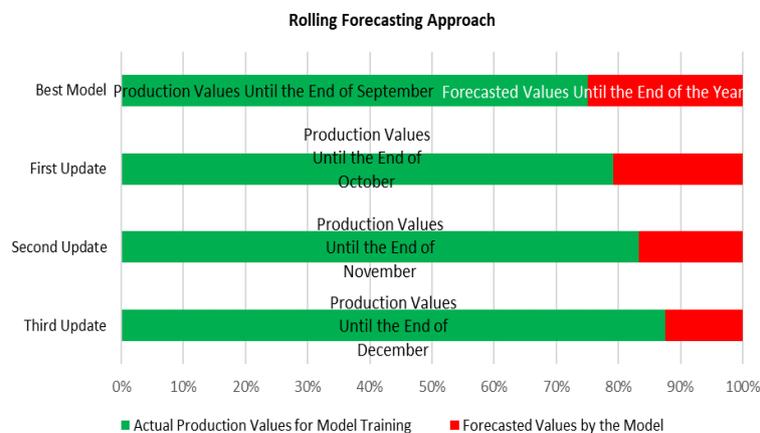
Fig. 8. Evaluation criteria values for MAE, MSE, RMSE, and R-squared.

### 3.5 | Result and Rolling Forecast

After implementation, validation, evaluation, and comparison of the results, the best time series forecasting models were selected as shown in Fig. 6. In this section, using a rolling or moving window approach, the daily production values for the SAIPA automotive group were forecast. To implement the rolling forecasting approach, the model parameters were updated every 30 days over the last three months of the dataset. Fig. 9 shows the rolling forecasting method used in the research.

**Table 6. Evaluation criteria values of the best forecasting model for the SAIPA automotive group companies.**

	Saipa		Pars Khodro		Saipa Citroën	
	ARIMA	ARIMAX	ARIMA	SARIMAX	ARIMA	ARIMAX
MSE	84228.015	50939.139	45189.1578	11833.48	24852.285	5547.079
RMSE	290.220	225.667	672.231	108.782	157.649	97.709
MAE	257.60	147.098	449.285	69.5627	116.2545	61.217
	Bonro		Zamyad		Saipa Diesel	
	AR	ARIMAX	AR	ARIMAX	---	ARIMAX
MSE	15018.9059	5848.7371	7969.475	4749.805	13.9650	10.6592
RMSE	122.5576	76.4770	89.2719	68.9188	3.7369	3.2648
MAE	115.9825	49.706	79.6305	48.5793	2.4685	1.9755

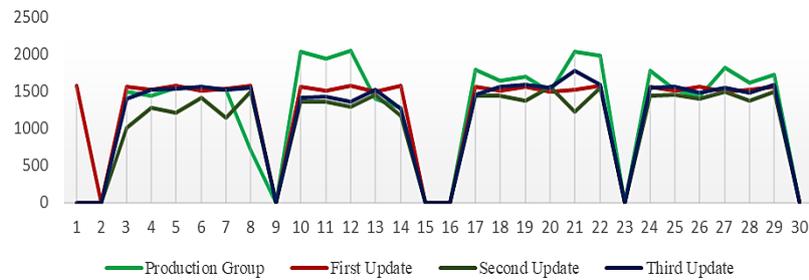


**Fig. 9. Rolling forecasting approach and updating parameters of the best model in October, November, and December.**

In Table 7, the mean absolute errors for the parameter update stages of the selected forecasting models are presented. It is expected that by updating the parameters based on new data, the forecasting models' error will decrease. However, this did not happen in the second update. This argument stems from an unexpected production decrease in October, highlighting the importance of patterns and data analysis for accurate forecasting. In Fig. 10, the forecasted production values for December for the update stages, the actual production statistics in this period, and the better performance of the third update are shown.

**Table 7. MAE values based on parameter update stages of selected forecasting models for SAIPA.**

Performance Evaluation Metric	First Update	Second Update	Third Update
MAE for October	253.1745	-----	-----
MAE for November	361.9097	366.9636	-----
MAE for December	227.9702	255.3307	209.3926



**Fig. 10. Comparison of forecasted production values for update stages and actual production statistics in December.**

## 4 | Conclusion

This study examined and forecasted production for the SAIPA automotive group using statistical time-series models. The results indicate that the ARIMAX model is the best for forecasting daily production. This model has provided more accurate forecasts by incorporating an exogenous variable (weekly production schedule). However, to achieve accurate forecasting models, examining and controlling production based solely on production trends, the ARIMA and AR models have performed better. The SARIMA and SARIMAX models, which are typically used for seasonal data, performed poorly in this study. This result may be due to the lack of strong seasonal patterns in daily production data. However, this result may change when examining cumulative forecasts of weekly or monthly values.

Using the rolling forecasting approach only in the first and second stages has reduced errors and improved forecast accuracy. It indicates that not only is the selection of an appropriate model and parameter optimization crucial for achieving more accurate forecasts, but also the examination and analysis of data patterns and unexpected trends are essential for achieving an optimal forecasting model.

Suggestions for future research include using more advanced models, such as artificial intelligence and machine learning, and employing more complex models, such as neural networks and deep learning, for production forecasting. Other practical applications of this research include developing software tools to assist managers in decision-making based on production forecasts. These tools can help improve production planning and inventory management. Given the results and suggestions, this research is expected to improve production processes in the automotive industry and other similar industries and to pave the way for further research in this field.

## Acknowledgments

The authors would like to express their sincere appreciation to all individuals whose valuable comments and support contributed to the successful completion of this study.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

## References

- [1] Geurtsen, M., Didden, J. B. H. C., Adan, J., Atan, Z., & Adan, I. (2023). Production, maintenance and resource scheduling: A review. *European journal of operational research*, 305(2), 501–529. <https://doi.org/10.1016/j.ejor.2022.03.045>
- [2] Lohmer, J., & Lasch, R. (2020). Blockchain in operations management and manufacturing: Potential and barriers. *Computers & industrial engineering*, 149, 106789. <https://doi.org/10.1016/j.cie.2020.106789>
- [3] Wang, G. (2021). Integrated supply chain scheduling of procurement, production, and distribution under spillover effects. *Computers & operations research*, 126, 105105. <https://doi.org/10.1016/j.cor.2020.105105>
- [4] Estes, A., Peidro, D., Mula, J., & Diaz-Madroño, M. (2023). Reinforcement learning applied to production planning and control. *International journal of production research*, 61(16), 5772–5789. <https://doi.org/10.1080/00207543.2022.2104180>
- [5] Chiang, W. C., Chen, J. C. H., & Xu, X. (2007). An overview of research on revenue management: Current issues and future research. *International journal of revenue management*, 1(1), 97–128. <https://doi.org/10.1504/IJRM.2007.011196>
- [6] Singh, A., & Mishra, G. C. (2015). Application of box-jenkins method and artificial neural network procedure for time series forecasting of prices. *Statistics in transition new series*, 16(1), 83–96. <https://doi.org/10.21307/stattrans-2015-005>
- [7] Ghomi, S. M. T. F., & Forghani, K. (2016). Airline passenger forecasting using neural networks and box-jenkins. *2016 12th international conference on industrial engineering (ICIE)* (pp. 10–13). IEEE. <https://doi.org/10.1109/INDUSENG.2016.7519342>
- [8] Ni, M., He, Q., & Gao, J. (2016). Forecasting the subway passenger flow under event occurrences with social media. *IEEE transactions on intelligent transportation systems*, 18(6), 1623–1632. <https://doi.org/10.1109/TITS.2016.2611644>
- [9] Milenković, M., Švadlenka, L., Melichar, V., Bojović, N., & Avramović, Z. (2018). SARIMA modelling approach for railway passenger flow forecasting. *Transport*, 33(5), 1113–1120. <https://doi.org/10.3846/16484142.2016.1139623>
- [10] Chen, E., Ye, Z., Wang, C., & Xu, M. (2019). Subway passenger flow prediction for special events using smart card data. *IEEE transactions on intelligent transportation systems*, 21(3), 1109–1120. <https://doi.org/10.1109/TITS.2019.2902405>
- [11] Su, Y., & Ye, Y. (2020). Daily passenger volume prediction in the bus transportation system using arimax model with big data. *2020 international conference on cyber-enabled distributed computing and knowledge discovery (CYBERC)* (pp. 291–300). IEEE. <https://doi.org/10.1109/CyberC49757.2020.00055>
- [12] Ghauri, S. P., Ahmed, R. R., Streimikiene, D., & Streimikis, J. (2020). Forecasting exports and imports by using autoregressive (AR) with seasonal dummies and Box-Jenkins approaches: A case of Pakistan. *Engineering economics*, 31(3), 291–301. [10.5755/j01.ee.31.3.25323](https://doi.org/10.5755/j01.ee.31.3.25323)
- [13] Guleryuz, D. (2021). Forecasting outbreak of COVID-19 in Turkey; comparison of box-jenkins, brown's exponential smoothing and long short-term memory models. *Process safety and environmental protection*, 149, 927–935. <https://doi.org/10.1016/j.psep.2021.03.032>
- [14] Hadwan, M., Al-Maqaleh, B. M., Al-Badani, F. N., Khan, R. U., & Al-Hagery, M. A. (2022). A hybrid neural network and box-jenkins models for time series forecasting. *Computers, materials & continua*, 70(3), 10.32604/cmc.2022.017824
- [15] Yasmin, S., & Moniruzzaman, M. (2024). Forecasting of area, production, and yield of jute in Bangladesh using box-jenkins ARIMA model. *Journal of agriculture and food research*, 16, 101203. <https://doi.org/10.1016/j.jafr.2024.101203>