9 (1): 109-117 (2025)

Journal of Aviation

https://dergipark.org.tr/en/pub/jav

e-ISSN 2587-1676

Forecasting Turkey's Air Cargo Tonnage: A Comparative Analysis of Statistical Techniques and Machine Learning Methods

Cüneyt Çatuk[1*]

[1*]Sirnak University, Department of Management and Organization, Silopi, Sirnak. (c.catuk@sirnak.edu.tr)

Article Info

Received: 11 November 2024
Revised: 15 December 2024
Accepted: 23 December 2024
Published Online: 24 February 2025

Keywords:
Air Logistic Prediction
SARIMA
Random Forest
ARIMAX
GBR
Support Vector Regression

Corresponding Author: *Cüneyt Çatuk*

RESEARCH ARTICLE

https://doi.org/10.30518/jav.1575091

Abstract

With the expanding global economy, the demand for air logistics continues to grow, further emphasizing its significance. However, this increased demand also presents a barrier to the growth of the air transportation sector, which is marked by a high degree of vulnerability. This study aims to forecast cargo volumes in the air logistics sector, which holds considerable growth potential. To achieve this, two statistical models (SARIMA and ARIMAX) and three machine learning methods (Gradient Boosting Regression Tree, Random Forest, and Support Vector Regression) were utilized in a comparative analysis, and forecasts for air cargo volumes were generated using the model with the best performance. The findings reveal that machine learning-based models outperform statistical models when applied to time series data. Specifically, the Random Forest model demonstrated superior performance in forecasting 1-10 month periods, while the Gradient Boosting Regressor (GBR) outperformed other models in 5-month periods. Additionally, the SARIMA model was found to be highly competitive for short-term forecasts. Based on these results, it was determined that the Random Forest model provides higher accuracy for 1-10 month periods, whereas the GBR model excels in 5-month periods. The results further indicate that dynamic modelling strategies achieved through machine learning methods yield more accurate predictions compared to statistical models.

1. Introduction

The positive developments that occurred in the industry with the industrial revolution led people to more consumption (Ekinler, 2022). The worldwide transportation sector is becoming significant due to the advent of mass manufacturing, rapid technical advancements, and heightened competitiveness (Papatya & Uygur, 2019). This situation contributed to the growth of the economy. The acceleration of economic growth and the increase in market demand driven by globalization have led to rapid expansion in the air logistics sector. The need for timely delivery of perishable goods, chemicals, and valuable items has significantly contributed to the swift development of air transportation (Nalçacıgil, 2023). This surge in market demand has resulted in a consistent rise in cargo volume over the years; however, the sector's vulnerability is perceived as a major barrier to its growth (Bakırcı, 2013). Therefore, accurately forecasting demand within the air logistics sector will support the ongoing development of the air cargo industry.

The obligation for air cargo companies to fulfill shipments within specified timeframes based on demand necessitates approaching air cargo volume forecasting as a regression problem within the framework of time series analysis. In this context, short-term cargo volume forecasts are conducted by accounting for temporal fluctuations in air logistics volume and the factors influencing these variations. Both statistical and artificial neural network methods can be employed in air cargo forecasting. Among statistical methods, and the Autoregressive Integrated Moving Average with Exogenous Input (Bierens, 1987) are prominent. Neural network-based methods include algorithms such as Gradient Boosting Regression Tree (Quinlan, 1986), Support Vector Regression (SVR) (Li-Xia et al., 2011), and Random Forest (Breiman, 2001).

One of the most widely used statistical methods in air logistics forecasting is the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The SARIMA model stands out for its ability to incorporate seasonal cycles in time series data, enabling future predictions based on historical data. Its high accuracy in datasets where seasonality and trends, as seen in air cargo data, are prominent is one of the main reasons for its preference. Additionally, the SARIMA model addresses seasonal components separately, providing higher accuracy for short- and medium-term forecasts. These features make SARIMA a widely utilized method for analyzing datasets with seasonality and trends, such as in air logistics.

The motivation of this study is to examine the advantages of statistical and machine learning models and analyze the complementary effects of each model. In this study, two statistical models SARIMA and ARIMAX and three machine learning methods Gradient Boosting Regression Tree (GBRT), Random Forest, and Support Vector Regression (SVR) were used for comparative analysis. SVR is a widely-used machine learning method that performs effectively on smaller datasets and in capturing complex relationships. In contrast to SVR, the GBRT method is more suitable for larger and more complex datasets, reducing prediction errors through sequentially constructed trees. Unlike GBRT and SVR, the Random Forest method operates by allowing each tree to work in parallel and independently. Each of these methods provides a robust foundation for modeling the dynamic and multidimensional nature of air cargo demand, which is shaped by economic, seasonal, and operational factors. The future air cargo volume forecasts were conducted using the model that yielded the most accurate results. To evaluate the predictive power of each model, comparisons were made across 1,5,10 monthly periods. Dynamic and static forecasting methods were employed as prediction strategies. In the second section of the study, a literature review of time series analysis is presented, while the third section provides theoretical foundations of the forecasting models. The fourth section explains the dataset and normality tests, and in the fifth section, the results obtained are discussed.

2. Literature Review

A wide range of methods has been developed in the literature on time series analysis, yielding significant results. These models are categorized into two main groups: statistical forecasting models and machine learning models (Nacar and Erdebilli, 2021). The literature review presented below will cover studies conducted on both approaches.

The multiple linear regression model, one of the most commonly used statistical forecasting models, has long been established in the literatüre (Çubukcuoğlu, Ersöz et al. 2013, Kılıç 2013). This model examines the relationship between a dependent variable and multiple independent variables, enabling the analysis of associations between these variables (Yavuz, 2009). His model, with its functional structure, can be easily applied across various fields. In addition to the multiple linear regression model, another method commonly used in statistical forecasting is the ARIMA (AutoRegressive Integrated Moving Average) model, also known as the Box-Jenkins model. This model assumes a relationship between the predicted variable and past data values. To ensure accurate analyses, non-stationary time series are first transformed to achieve stationarity (Peter et al., 2012, p. 136). ARIMA has been widely utilized as a forecasting methodology and in time series analysis across multiple domains (Newbold, 1983). Tortum et al. (2014) attempted to forecast air transport demand in Turkey using ARIMA and Seasonal ARIMA (SARIMA) models, concluding that the SARIMA model could be effectively employed for air transport demand forecasting. (Önen, 2020) sought to forecast air cargo volume using data from 2020 to 2023 and found that the predicted values were within a 95% confidence interval, with both the Mean Absolute Percentage Error (MAPE) and Theil's inequality coefficient remaining within acceptable ranges. ARIMA and multiple linear regression models yield meaningful results only for linear relationships (Lee and Tong, 2011). The ARIMA model, using a single time series, does not effectively represent multivariate time series, necessitating the use of a multivariate model such as ARIMAX (Kongcharoen and Kruangpradit, 2013). Anggraeni, et al., (2017) compared the Vector Autoregressive (VAR) and ARIMAX models to forecast rice prices in Indonesia, finding that the ARIMAX model outperformed the VAR model by 15.27%, with a MAPE of 0.15%.

Recent advances in machine learning methods have significantly improved forecasting accuracy, especially for nonlinear datasets (Adetunji et al., 2022). These studies highlight that machine learning techniques often outperfom traditional statistical models in time series forecasting. Support Vector Regression (SVR), a popular machine learning method, performs effectively with smaller datasets and complex relationships. It is a nonlinear extension of the Generalized Portrait algorithm, originally developed in Russia in the 1960s (Vapnik, 1998; Vapnik and Lerner, 1963; Vapnik and Chervonenkis, 1964). The modern form of Support Vector Machines (SVMs) was largely developed by Vapnik and colleagues at AT&T Bell Laboratories (Boser, Guyon, and Vapnik, 1992). Huang et al. (2005) applied SVMs to forecast financial movements using NIKKEI 225 index data, finding that SVMs demonstrated superior performance in financial forecasting. Using SVR modeling, Yang et al. (2022) conducted an air freight forecasting study in which the model outperformed other methods, achieving a Mean Absolute Percentage Error (MAPE) of less than 2.5%, along with the lowest Mean Absolute Error and Root Mean Square Error.

In contrast to SVR, the Gradient Boosting Regression Tree (GBRT) method is suitable for larger and more complex datasets, reducing prediction errors through a series of sequentially built trees (Friedman, 2001, 2002). In a study on flight delay prediction, Manna et al. (2017) demonstrated that the Gradient Boosted Decision Tree model achieved the highest R-squared values, with 92.3185% accuracy for arrival delays and 94.8523% for departure delays. Furthermore, Persson et al. (2017) applied the GBRT model to forecast future electricity production from rooftop PV installations.

Unlike GBRT and SVR, Random Forest operates with each tree running in parallel and independently. Introduced by Breiman, Random Forest is a widely used machine learning technique (Breiman, Friedman, Olshen, and Stone, 1984). In a study on cargo weight predictions for flights, Pinheiro (2021) employed various machine learning models, concluding that Random Forest achieved the best performance, with a Root Mean Square Error of 33%. Using Random Forest on a U.S. airline's arrival data, Rahul et al. (2022) predicted delay durations with an accuracy rate of 86%. Additionally, Adetunji et al. (2022) used Random Forest to forecast housing prices in Boston, achieving an acceptable prediction accuracy with an error margin of ±5%.

3. Design and Methodology

3.1. Data Set

The dataset for this study covers the years 2012–2023. Data on Turkey's air cargo tonnage, which serves as the dependent variable, were obtained from the General Directorate of State Airports Authority (DHMI). Among the independent variables, exchange rates in USD were collected from the Electronic Data Distribution System (EVDS) of the Central Bank of the Republic of Turkey. Crude oil purchase prices in USD were sourced from the U.S. Energy Information Administration, while GDP data were obtained from the Turkish Statistical Institute (TUIK). Since the predictive model is also implemented on a quarterly basis, a dataset comprising more than 30 time series observations is deemed sufficient (Gujarati, 2014).

3.2. Research Methodology

To forecast monthly air cargo tonnage, the Box-Jenkins methodology was employed using SARIMA and ARIMAX models. The best-fit parameters for these models were selected based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The final SARIMA model was determined as SARIMA (1,1,1) (0,1,1) [12] based on the lowest AIC score. While machine learning methods such as Random Forest, Gradient Boosting Regression Tree (GBRT), and Support Vector Regression (SVR) were also utilized. Eviews 12 software was used to determine and forecast SARIMA and ARIMAX models, while Python was employed for RF, GBRT, and SVR predictions.

3.2.1. SARIMA

SARIMA is formed by adding seasonal terms in the ARIMA models:

SARIMA (p, d, q) (P, D, Q) [S],

In the SARIMA model, p represents the non-seasonal autoregressive (AR) order, P denotes the seasonal autoregressive order, q is the non-seasonal moving average (MA) order, and Q indicates the seasonal moving average order. Meanwhile, d and D represent the overall differencing and seasonal differencing orders, respectively (Pepple and Harrison, 2017).

SARIMA(p,d,q)(P,D,Q)[S] models are written as (Pankratz, 1983);

$$\varphi_p(B)\phi_P(B^S)\nabla^d\nabla_s^D y_t = \theta_q(B)\Theta_Q(B^S)\mathrm{T}_t \quad (1)$$

∅ is the non-seasonal parameter of autoregression and θ is the non-seasonal parameter of moving average, φ is the seasonal parameter of autoregression and Θ is the seasonal parameter of moving average, ω is frequency and B is the diferential variable (Pepple and Harrison 2017).

3.2.2. ARIMAX

The ARIMAX model is an extension of the ARIMA model. This model incorporates additional independent variables, represented by the X at the end, which stands for "exogenous variables." This involves adding a separate exogenous variable to the model to aid in measuring the endogenous variable (Adu, Appiahene et al., 2023).

The ARIMAX (p,d,q) model consists of four main components (Almaleck, Massucco et al., 2024):

- An autoregressive component of order p,
- An order of differencing d,
- A moving average component q,
- A dataset comprising exogenous inputs.

The ARIMAX model equation is expressed as follows:

$$\Delta P_t = c + \beta X + \emptyset \Delta P_{t-1} + \theta_{1\,\epsilon_{t-1}} + \epsilon_t \quad (2)$$

Here, P_t and P_{t-1} represent the values in the current and previous periods, respectively. Similarly, ϵ_t and $\epsilon_{t-1}$are the error terms for these two periods. c denotes a constant term. $\emptyset_1$ and θ_1 indicate the influence of the previous period's value P_{t-1} and error ϵ_{t-1} in predicting the current value. $\beta$is a coefficient to be estimated based on model selection and data, and X is the exogenous variable of interest

The ARIMAX model is valuable as it integrates time series and regression components, allowing for a more comprehensive forecasting approach (Moslemi et al., 2024).

3.2.3. Random Forest

The Random Forest model is a widely used machine learning algorithm that reaches a single outcome by aggregating the outputs of multiple decision trees. Decision trees start with the most fundamental question and follow a series of questions, which form the decision nodes of the tree. Each question contributes to determining the final answer. In this structure, observations that meet certain criteria follow the "yes" branch, while those that do not meet these criteria follow an alternative branch. Decision trees use these questions to find the optimal method for training subsets and achieving the best results (Melzer, 2023).

The Random Forest equation is expressed as follows (Xing and Zhang, 2024):

$$F = S(T_1(d_1), T_2(d_2), \ldots, T_n(d_n)) \quad (3)$$

Here, F represents the final class, S denotes the selection function, T_n is the decision tree processing function, d_n represents the input data for each decision tree, and n is the number of decision trees. Based on these functions, the corresponding Random Forest prediction model can be constructed.

3.2.4. Support vector regression (SVR)

Support Vector Regression (SVR) is one of the most important branches of Support Vector Machines (SVMs). The classical regression model constructs the loss function by calculating the difference between the actual value and the predicted value. For continuous-valued functions, the mathematical representation can be simplified by incorporating the x value into the w vector and adding b for multidimensional data, as shown in Equation 1. This results in a multivariate regression, illustrated in Equation 1.2 (Awad et al., 2015).

Equation 1:

$$y = f(x) = < w, x > +b = \sum_{j=!}^{M} w_j\, x_j + b, y, b \in R, x, w \in R^m \quad (4)$$

Equation 1.2 (Multivariate Regression):

$$f(x)[^{w^T}_{b}[^{x}_{i} = w^{Tx} + b\ x, w \in R^{M+1}] \quad (5)$$

3.2.5. Gradient boosting regression tree

A decision tree is a predictive model proposed by Quinlan (1986). As illustrated in Figure 1, a decision tree is a type of binary tree where each node represents a test of an attribute, and the leaves indicate the predicted value. When the target variable has a continuous real value (typically represented by real numbers), the decision tree is referred to as a regression tree.

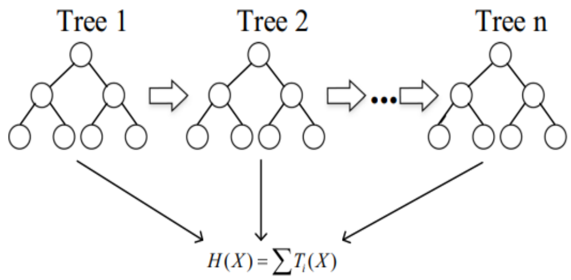

Figure 1. A GBRT (Gradient Boosting Regression Tree) model. Source: Huang et al. 2019

Each tree within the model is a decision tree, and the t_i tree is built sequentially after the t_(i-1) tree. The predicted value of the GBRT model is the sum of the values predicted by each individual tree. The target input value for the t_i tree is the residual between the current predicted value and the true target value, defined as:

$$input_i = \sum_{j=1}^{i-1} T_j - y_{true} \quad (6)$$

Where T_j denotes the prediction result of the jjj-th decision tree $1 \leq j \leq i - 1$and y_{true} is the true target value of the examples.

3.2.6. Evaluating the Predictive Power of the Model

After identifying the appropriate forecasting model, it is essential to conduct forecast evaluation tests to assess the model's predictive capability for future projections. If the predictive accuracy of the selected model does not meet the desired level of statistical significance in these tests, it should not be used for future forecasting. In the study, forecasted and actual passenger numbers, along with the error margins for each of the various models, are presented for evaluation.
To identify outliers, residual distribution graphs, bias, and covariance values are utilized. To measure the accuracy of the forecasts and to determine the predictive power of the model, Mean Absolute Percentage Error (MAPE) and the Theil Inequality Coefficient are employed. According to Lewis (1982), models with a MAPE value below 10% are classified as very good, those between 10–20% as good, those between 20–50% as acceptable, and those above 50% as inaccurate or erroneous. Additionally, the Theil coefficient is expected to be close to zero (Vergil & Özkan, 2007). This coefficient is divided into three components: the "Bias" proportion, which represents systematic error, with values closer to zero indicating higher reliability in forecast results. The second component, the covariance proportion, reflects unsystematic error; a larger value compared to other components indicates the error is unsystematic (Bozkurt, 2013:186). The covariance proportion represents the variability in the model that arises beyond our control, helping to explain external influences on forecast error.

3.2.7. Structuring the Forecasting Strategy

Static and dynamic forecasting are utilized in the comprehensive evaluation of statistical and machine learning models. Static forecasting refers to cases where the model's structure and parameters remain fixed once the training and testing datasets are defined. In contrast, dynamic forecasting is more complex, as the model's structure and parameters are recalibrated whenever new observations are introduced. Therefore, in dynamic forecasting, the training data are updated after each forecast by adding the most recent observation.

In this study, the independent variables used include GDP, exchange rate, and Brent crude oil prices, all of which are factors influencing airline cargo tonnage (Tuncer & Aydoğan, 2019; Totamane et al., 2009).

4. Findings

4.1. Descriptive Statistics

The graphical representation of data for the relevant series is provided in Figure 2. As illustrated in Figure 2, airline cargo tonnage increased from 2012 until 2020, experienced a decline in 2020 and 2021 due to the pandemic, and then resumed an upward trend in 2022. Additionally, demand demonstrates seasonal fluctuations. Crude oil purchase prices exhibit a volatile pattern, while GDP has shown a rising trend following a decline observed from 2012. Furthermore, the USD/TRY exchange rate has been on an upward trajectory since 2018. Descriptive statistics for the series are presented below in Table 1.

4.2. Descriptive Statistics

As shown in Table 1, the Jarque-Bera test results indicate that, at the 95% significance level, the probability values for Airline Cargo Tonnage (Ton) and Brent are greater than 0.05 ($p > 0.05$). Therefore, these series can be considered normally distributed. However, the probability values for Gross Domestic Product (GDP) and Exchange Rate are 0.00002 and 0.0, respectively, indicating statistical significance at the 95% level ($p < 0.05$). Thus, the null hypothesis H_0 is rejected for these series, suggesting that they do not follow a normal distribution. To address this, logarithmic transformations were applied to these series, and the corresponding results are presented in Table 2.

Table 1. Descriptive Statistics

Statistic	Ton	GDP	Exchange Rate	Brent
Mean	281071.2	7.27E+10	6.764236	75.04694
Median	275582.0	7.23E+10	3.815000	71.71500
Maximum	484194.0	9.42E+10	29.07000	125.4500
Minimum	87953.00	6.00E+10	1.760000	18.38000
Std.Dev.	83161.89	8.69E+09	6.489477	25.56991
Skewness	0.335005	0.929005	1.768739	0.177831
Kurtosis	2.758974	3.934021	5.387848	1.933500
Jarque-Bera	3.042042	25.94757	109.2934	7.583511
Probabilty	0.218489	0.00002	0.00000	0.052556
Sum	40474253	1.05E+13	974.0500	10806.76
Sum Sq.Dev.	9.89E+11	1.06E+22	6022.203	93693.85
Observation	144	144	144	144

According to the Jarque-Bera test results following the logarithmic transformation of the series, the probability values for the LGDP and Exchange Rate series are 0.07 and 0.06, respectively ($p > 0.05$). This indicates that these series have also been adjusted to conform to a normal distribution.

Table 2. Log (Exchangre Rate, GDP)

Statistic	Exchange Rate	GDP
Mean	1.561783	2.500.257
Median	1.338899	2.500.455
Maximum	3.369707	2.526.833
Minimum	0.565314	2.481.803
Std.Dev.	0.79861	0.114053
Skewness	0.615611	0.581266
Kurtosis	2.327863	3.352372
Jarque-Bera	11.80607	8.853872
Probabilty	0.072132	0.067322
Sum	974.0500	3.600.371
Sum Sq.Dev.	6022.203	1.860.155
Observation	144	144

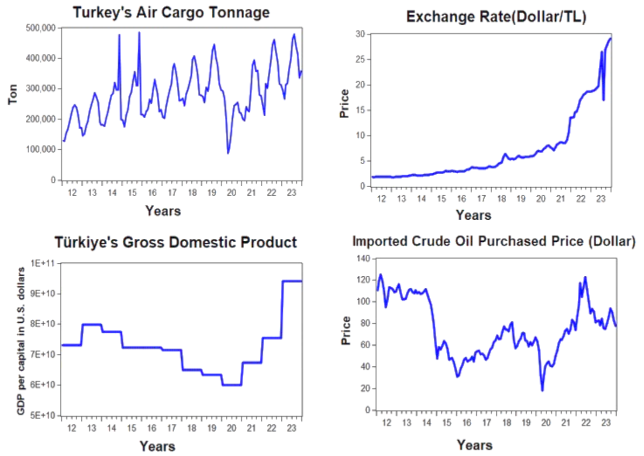

Figure 2. Graphical Representation of the Series (Monthly Periods for the 2012-2023 Period)

4.2. Forecasting performance of statistical models

The forecast performance of the two statistical models is presented in Table 3. Prediction performance has been evaluated using seven metrics, with the best results highlighted in bold.

The results according to forecasting strategies are divided into six sections and presented in Table 3. Each section consists of two statistical models and seven evaluation metrics. The forecast performances of the static and dynamic strategies have been analyzed and compared.

In the static analysis, examining short-term forecasts (1-step and 5-step), the SARIMA model provides more accurate predictions with lower MAE and MAPE values, indicating better absolute error rates. This finding suggests that the SARIMA model performs better in the short term. In contrast, for long-term forecasts (10-step), the ARIMAX model demonstrates a more balanced and consistent performance with lower RMSE, MAE, and Theil's U values, indicating a lower overall error rate. This result shows that the ARIMAX model outperforms the SARIMA model in long-term forecasting, providing more reliable outcomes. In conclusion, the SARIMA model produces more accurate predictions for short-term forecasts, while the ARIMAX model performs better in terms of overall error rates for long-term forecasts. This finding suggests that the SARIMA model should be preferred for short-term projections, whereas the ARIMAX model is more suitable for long-term projections.

The findings from the dynamic analysis indicate that the ARIMAX model provides lower values for critical metrics such as RMSE, MAE, and Theil's U in short-term (1-step), medium-term (5-step), and long-term (10-step) forecasts, which measure error rates. These results demonstrate that the overall predictive accuracy of the ARIMAX model is higher than that of the SARIMA model. Specifically, ARIMAX exhibits lower RMSE and MAE values across all forecast steps, indicating superior error performance. Additionally, the ARIMAX model's superiority in terms of MAPE and Theil's U values suggests that it offers not only better absolute error rates but also a more balanced and reliable forecast in terms of relative error performance. Consequently, the data indicate that the ARIMAX model is a stronger choice for dynamic forecasting, suitable for both short- and long-term predictions. In conclusion, the comparison of static and dynamic analyses reveals that the lower error rates of the static SARIMA model indicate its reliability and consistency in short-term forecasting. However, in long-term (10-step) forecasts, the ARIMAX model demonstrates lower RMSE and Theil's U values in both static and dynamic forecasts, providing a more balanced forecast in terms of overall error performance. Notably, static ARIMAX forecasts yield the best performance over the long term, marked by lower error rates. This finding suggests that the static ARIMAX model is more suitable for long-term forecasting.

4.3. Forecasting performance of statistical models

In the machine learning comparison in Table 4, the Random Forest model exhibited the lowest RMSE, MAE, and MAPE values in both static and dynamic forecasts, along with the lowest bias (BIAS) rate. This result indicates that the Random Forest model offers greater reliability in both short- and long-term forecasts compared to other models. On the other hand, The Support Vector Regression (SVR) model demonstrated weaker predictive performance compared to other machine learning models, particularly in long-term forecasts. This outcome can be attributed to SVR's sensitivity to noise in time series data and its limitations in capturing nonlinear trends effectively. Unlike tree-based models such as Random Forest and Gradient Boosting, which can handle complex interactions between variables, SVR relies on a kernel-based transformation, making it less robust for highly volatile air cargo data. The Gradient Boosting Regression Tree (GBRT) model showed moderate performance, generally demonstrating higher error rates and thus offering limited reliability.

Overall, the Random Forest model stands out as the most suitable machine learning method in terms of forecast reliability. Notably, for 5-month forecast periods, the GBRT model produced the best results. Consequently, the superior performance of Random Forest in both short- and long-term forecasts makes it the most preferred model among forecasting methods. Furthermore, when comparing static and dynamic forecasts of the Random Forest model, the dynamic method yielded the best results across 1-step, 5-step, and 10-step forecasts.

4.4. Comparison of Machine Learning and Statistical Methods

In machine learning methods, the best results were achieved with the dynamic approach. In statistical methods, the SARIMA model provided the most accurate results for short-term forecasts (1-step and 5-step), while the ARIMAX model performed best for long-term forecasts (10-step). A comparison of these models revealed that Random Forest yielded the highest accuracy in 1-step, 5-step, and 10-step forecasts.

Given the superior performance of the Random Forest analysis in both short- and long-term forecasts, a projection graph for 2024-2025 is presented in Figure 3. The model's forecast for total airline cargo in 2024 is 4,675,000 tons, while the estimate for 2025 is 4,690,000 tons. Considering the airline cargo tonnage in 2023 was 4,163,142 tons, a 12% increase is projected for 2024, followed by a 1% increase for 2025 compared to the previous year.

Table 3. Forecasting performance of two statistical models

				Static				
	1-step ahead			**5-step ahead**			**10-step ahead**	
SARIMA	RMSE	**45532.3**	**SARIMA**	RMSE	**45532.3**	**SARIMA**	RMSE	45532.3
	MAE	**28712.19**		MAE	**28712.19**		MAE	32121.2
	MAPE	**11.31**		MAPE	**11.314**		MAPE	11.314
	BIAS	**0.001105**		BIAS	**0.0011**		BIAS	0.001
	VARIANCE	**0.08**		VARIANCE	**0.0804**		VARIANCE	0.08
	COVARIANCE	**0.91**		COVARIANCE	**0.9184**		COVARIANCE	0.918
	THEIL	**0.07**		THEIL	**0.077**		THEIL	0.077
ARIMAX	RMSE	44837.5	**ARIMAX**	RMSE	44837.5	**ARIMAX**	RMSE	**44837.5**
	MAE	30319.34		MAE	30319.4		MAE	**30319.34**
	MAPE	12.24		MAPE	12.248		MAPE	**12.482**
	BIAS	0		BIAS	0		BIAS	**0**
	VARIANCE	0.09		VARIANCE	0.096		VARIANCE	**0.0963**
	COVARIANCE	0.9		COVARIANCE	0.9		COVARIANCE	**0.903**
	THEIL	0.07		THEIL	0.076		THEIL	**0.077**
				Dynamic				
	1-step ahead			**5-step ahead**			**10-step ahead**	
SARIMA	RMSE	79556.89	**SARIMA**	RMSE	79556	**SARIMA**	RMSE	79556.89
	MAE	62985.49		MAE	62985.49		MAE	62985.49
	MAPE	2.460.376		MAPE	2.460.376		MAPE	2.460.376
	BIAS	0.004		BIAS	0.004		BIAS	0.004
	VARIANCE	0.808		VARIANCE	0.808		VARIANCE	0.808
	COVARIANCE	0.18		COVARIANCE	0.186		COVARIANCE	0.186
	THEIL	0.13		THEIL	0.13		THEIL	0.138
ARIMAX	RMSE	**69490.7**	**ARIMAX**	RMSE	**69490.7**	**ARIMAX**	RMSE	**44837.5**
	MAE	**54202.93**		MAE	**54202.93**		MAE	**30319.34**
	MAPE	**212.932**		MAPE	**21.293**		MAPE	**12.24**
	BIAS	**0.0016**		BIAS	**0.001**		BIAS	**0**
	VARIANCE	**0.4155**		VARIANCE	**0.4155**		VARIANCE	**0.096**
	COVARIANCE	**0.582**		COVARIANCE	**0.5827**		COVARIANCE	**0.903**
	THEIL	**0.12**		THEIL	**0.1203**		THEIL	**0.076**

Table.4 Forecasting performance of two statistical models

				Static				
1-step ahead			**5-step ahead**			**10-step ahead**		
RF	RMSE	**8553.4**	**RF**	RMSE	**24293.66**	**RF**	RMSE	**38184.05**
	MAE	**8553.4**		MAE	**22698.45**		MAE	**34345.93**
	MAPE	**2.35**		MAPE	**7.88**		MAPE	**14.02**
	BIAS	**0.0005**		BIAS	**4.61**		BIAS	**0.000**
	VARIANCE			VARIANCE	**0.99**		VARIANCE	**0.65**
	COVARIANCE			COVARIANCE	**0.58**		COVARIANCE	**0.47**
	THEIL	**0.012**		THEIL	**0.04**		THEIL	**0.09**
SVR	RMSE	86951.8	**SVR**	RMSE	60506.75	**SVR**	RMSE	60506.75
	MAE	86951.88		MAE	42024.14		MAE	44024.14
	MAPE	24.11		MAPE	9.016		MAPE	18.031
	BIAS	0.058		BIAS	0.008		BIAS	0.992
	VARIANCE			VARIANCE	1.914		VARIANCE	7.1964
	COVARIANCE			COVARIANCE	1.773		COVARIANCE	1.1935
	THEIL	0.137		THEIL	0.074		THEIL	0.10636
GBR	RMSE	20574.42	**GBR**	RMSE	26052.112	**GBR**	RMSE	37652.818
	MAE	20574.42		MAE	25444.459		MAE	33477.928
	MAPE	5.70		MAPE	8.523		MAPE	135.319
	BIAS	0.032		BIAS	0.0007		BIAS	0.0001
	VARIANCE			VARIANCE	0.714		VARIANCE	0.465
	COVARIANCE			COVARIANCE	0.648		COVARIANCE	0.064
	THEIL	0.029		THEIL	0.043		THEIL	0.0644
				Dynamic				
1-step ahead			**5-step ahead**			**10-step ahead**		
RF	RMSE	**8553.4**	**RF**	RMSE	20860.00	**RF**	RMSE	**38464.43**
	MAE	**8553.4**		MAE	18626.868		MAE	**30979.42**
	MAPE	**2.37**		MAPE	6.482		MAPE	**13.318**
	BIAS	**0.000**		BIAS	0.000		BIAS	**0.000**
	VARIANCE			VARIANCE	0.89		VARIANCE	**0.458**
	COVARIANCE			COVARIANCE	0.40		COVARIANCE	**0.470**
	THEIL	**0.012**		THEIL	0.034		THEIL	**0.064**
SVR	RMSE	86951.885	**SVR**	RMSE	42957.747	**SVR**	RMSE	60542.591
	MAE	86951.885		MAE	29828.977		MAE	44049.188
	MAPE	24.119		MAPE	8.988		MAPE	18.049
	BIAS	0.058		BIAS	0.008		BIAS	0.002
	VARIANCE			VARIANCE	1.966		VARIANCE	3.480
	COVARIANCE			COVARIANCE	0.193		COVARIANCE	1.189
	THEIL	0.1371		THEIL	0.074		THEIL	0.106
GBR	RMSE	20574.427	**GBR**	RMSE	**16703.041**	**GBR**	RMSE	44917.485
	MAE	20574.427		MAE	**15484.229**		MAE	32583.513
	MAPE	5.707		MAPE	**5.156**		MAPE	14.368
	BIAS	0.003		BIAS	**3.360**		BIAS	0.0002
	VARIANCE			VARIANCE	**0.333**		VARIANCE	0.209
	COVARIANCE			COVARIANCE	**0.136**		COVARIANCE	0.636
	THEIL	0.029		THEIL	**0.027**		THEIL	0.076

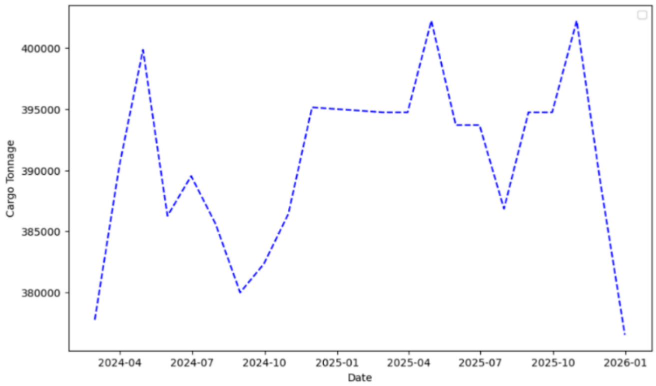

Figure 3. Random Forest Forecast of Airline Cargo for 2024-2025

5. Conclusion

In order to provide reliable forecasts of future air cargo tonnage in air cargo logistics, five different forecasting models were used to evaluate the advantages and disadvantages of statistical and machine learning-based forecasting approaches. To enable a more comprehensive comparison, various scenarios were developed by applying two distinct forecasting strategies—static and dynamic—across three forecasting periods.

Stationarity tests were conducted on the airline cargo tonnage series, used as the dependent variable, and on the independent variables: Gross Domestic Product (GDP), Brent oil prices, and exchange rate (Currency). According to the Jarque-Bera test results, the probability values for the Airline Cargo Tonnage (Ton) and Brent series were found to be significant at the 95% confidence level ($p > 0.05$), indicating that these series are normally distributed. For the Currency and GDP series, a logarithmic transformation was applied, after which the Jarque-Bera test confirmed that these series also conformed to a normal distribution.

Overall, the results indicate that machine learning-based models generally perform better with time series data. Specifically, the Random Forest model achieved the highest predictive accuracy across 1- to 10-month periods, as evidenced by its lower RMSE, MAE, and MAPE values compared to other models. The Gradient Boosting Regressor (GBR) outperformed other methods in 5-month forecasts, whereas SARIMA demonstrated strong competitiveness in short-term predictions. These findings suggest that Random Forest's ability to handle nonlinear patterns and interactions between multiple influencing factors makes it particularly suitable for air cargo forecasting. Furthermore, the superior performance of dynamic modeling strategies highlights the potential of machine learning methods in improving forecasting reliability over traditional statistical approaches.

The Random Forest forecast for the next two years predicts a 12% increase in 2024 and a 1% increase in 2025 compared to the previous year. Considering that air cargo volumes are influenced by numerous factors and long-term variables, future studies will focus on developing more advanced and sophisticated forecasting techniques that incorporate additional variables cited in the literature alongside those used in this model.

Ethical approval
Not applicable.

Conflicts of Interest
The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Adetunji, A. B., Akande, O. N., Ajala, F. A., Oyewo, O., Akande, Y. F., & Oluwadara, G. (2022). House price prediction using random forest machine learning technique. Procedia Computer Science, 199, 806-813.

Anggraeni, W., et al. (2017). The performance of ARIMAX model and Vector Autoregressive (VAR) model in forecasting strategic commodity price in Indonesia. Procedia Computer Science, 124, 189-196.

Bakırcı, M. (2013). Ulaşım Coğrafyası Açısından Türkiye'de Havayolu Ulaşımının Tarihsel Gelişimi ve Yapısı [The Historical Development and Structure of Air Transportation in Turkey from the Perspective of Transport Geography]. Marmara Coğrafya Dergisi, (27). URL:https://dergipark.org.tr/en/pub/marucog/issue/472/3869

Bierens, H. J. (1987). ARMAX model specification testing, with an application to unemployment in the Netherlands. Journal of Econometrics, 35(1), 161-190.

Bowerman, B. L., & O'Connell, R. T. (1993). Forecasting and time series: An applied approach (3rd ed.). Duxbury Press.

Box, G. E. P., & Jenkins, G. M. (1970). Time series analysis: Forecasting and control. Holden-Day.

Bozkurt, H. Y. (2013). Zaman serileri analizi [Time series analysis]. Ekin Yayınevi.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. Wadsworth & Brooks/Cole.

Burger, C., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide to time-series methods for tourism demand forecasting—A case study of Durban, South Africa. Tourism Management, 22, 403–409.

Çubukcuoğlu, B., et al. (2013). OECD Ülkeleri İçin Genişbant Abone Sayısını Etkileyen Faktörlerin Çoklu Regresyon Modeli ile Analizi [Analysis of Factors Affecting Broadband Subscriber Numbers in OECD Countries with Multiple Regression Model]. Sosyal Bilimler Araştırmaları Dergisi, 8(2), 26-41.

Ekinler, F. (2022). Sürdürülebilirlik ile İşletme Maliyeti Arasındaki İlişki [The Relationship Between Sustainability and Business Cost]. International Social Mentality and Researcher Thinkers Journal, 8(61), 1224-1230.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

Friedman, J. H. (2002). Stochastic gradient boosting. Computational Statistics & Data Analysis, 38(4), 367–378.

Gujarati, D. N. (2014). Econometrics by example (2nd ed.). Palgrave Macmillan.

Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. Computers & Operations Research, 32(10), 2513–2522.

Kılıç, S. (2013). Doğrusal regresyon analizi [Linear Regression Analysis]. Journal of Mood Disorders, 3(2), 90-92.

Kongcharoen, C., & Kruangpradit, T. (2013). Autoregressive integrated moving average with explanatory variable

(ARIMAX) model for Thailand export. Proceedings of the 33rd International Symposium on Forecasting. South Korea.

Lee, Y. S., & Tong, L. I. (2011). Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. Knowledge-Based Systems, 24, 66–72.

Lewis, C. D. (1982). Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. Butterworth-Heinemann.

Majkovic, D., O'Kiely, P., Kramberger, B., Vracko, M., Turk, J., Pazek, K., & Rozman, C. (2016). Comparison of using regression modeling and an artificial neural network for herbage dry matter yield forecasting. Journal of Chemometrics, 30, 203–209.

Manna, S., Biswas, S., Kundu, R., Rakshit, S., Gupta, P., & Barman, S. (2017). A statistical approach to predict flight delay using gradient boosted decision tree. 2017 International Conference on Computational Intelligence in Data Science (ICCIDS) (pp. 1–5). IEEE.

Meltzer, R. (2023). What is Random Forest? [Beginner's Guide Examples]. CareerFoundry. https://careerfoundry.com/en/blog/data-analytics/what-is-random-forest/

Moslemi, Z., Clark, L., Kernal, S., Rehome, S., Sprengel, S., Tamizifar, A., & Hodgett, M. (2024). Comprehensive forecasting of California's energy consumption: A multi-source and sectoral analysis using ARIMA and ARIMAX models. arXiv preprint arXiv:2402.04432.

Nacar, E. N., & Erdebilli, B. (2021). Makine Öğrenmesi Algoritmaları ile Satış Tahmini [Sales Forecasting with Machine Learning Algorithms]. Endüstri Mühendisliği, 32(2), 307-320.

Nalçacıgil, E. (2023). Hava kargo servis ve hizmetlerinin firmaların genel ve pazarlama performansı üzerindeki etkisi [The effect of air cargo services on companies' general and marketing performance]. Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi, 25(2), 637–654.

Nikolopoulos, K., Goodwin, P., Patelis, A., & Assimakopoulos, V. (2007). Forecasting with cue information: A comparison of multiple regression with alternative forecasting approaches. European Journal of Operational Research, 180(1), 354–368.

Önen, V. (2020). Arima yöntemiyle Türkiye'nin hava yolu kargo talep tahmin modellemesi ve öngörüsü [Modeling and Forecasting of Turkey's Air Cargo Demand with ARIMA Method]. Journal of Management and Economics Research, 18(4), 29-53.

Pankratz, A. (1983). Forecasting with univariate Box-Jenkins models: Concepts and cases. Wiley.

Papatya, G., & Uygur, M. N. (2019). Stratejik Karar Verme Sürecini Etkileyen Faktörler: Uluslararası Taşımacılık Sektörü İşletmelerinde Bir Araştırma [Factors Affecting the Strategic Decision-Making Process: A Research in International Transportation Sector Businesses]. Kafkas Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 10(19), 338-358.

Pepple, S. U., & Harrison, E. E. (2017). Comparative performance of Garch and Sarima techniques in the modeling of Nigerian broad money. CARD International Journal of Social Science and Conflict Management, 2(4), 258–270.

Persson, C., Bacher, P., Shiga, T., & Madsen, H. (2017). Multi-site solar power forecasting using gradient boosted regression trees. Solar Energy, 150, 423–436.

Peter, D., & Silvia, P. (2012). ARIMA Vs. ARIMAX – Which Approach is Better to Analyze and Forecast Macroeconomic Time Series. Proceedings of 30th International Conference Mathematical Methods in Economics (pp. 136-140).

Pinheiro, D. A. M. (2021). The impact of machine learning models in the prediction of air cargo (Master's thesis). Universidade NOVA de Lisboa, Portugal.

Pinheiro, D. A. M. (2021). The impact of machine learning models in the prediction of air cargo (Master's thesis). Universidade NOVA de Lisboa, Portugal.

Quinlan, J. R. (1986). Induction of decision trees. Machine Learning, 1(1), 81–106.

Rahul, R., Kameshwari, S., & Pradip Kumar, R. (2022). In flight delay prediction using random forest classifier. Proceedings of the International Conference on [Conference Name] (pp. 67–72). Springer.

Schölkopf, B., Burges, C., & Vapnik, V. (1995). Extracting support data for a given task. Proceedings of the First International Conference on Knowledge Discovery and Data Mining (pp. 144–152). AAAI Press.

Tortum, A., et al. (2014). Türkiye'de Hava Ulaşım Talebinin ARIMA Modelleri ile Tahmin Edilmesi [Forecasting Air Transportation Demand in Turkey with ARIMA Models]. Journal of the Institute of Science and Technology, 4(2), 39-54.

Vapnik, V. (1998). Statistical learning theory. Wiley.

Vapnik, V., Golowich, S., & Smola, A. (1997). Support vector method for function approximation, regression estimation, and signal processing. Advances in Neural Information Processing Systems 9 (pp. 281–287). MIT Press.

Vergil, H., & Özkan, F. (2007). Döviz Kurları Öngörüsünde Parasal Model ve Arima Modelleri: Türkiye Örneği [Monetary Model and ARIMA Models in Exchange Rate Forecasting: The Case of Turkey]. Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (13), 211-231.

Yang, C. H., Shao, J. C., Liu, Y. H., Jou, P. H., & Lin, Y. D. (2022). Application of fuzzy-based support vector regression to forecast of international airport freight volumes. Mathematics, 10(14), 2399.

Yavuz, S. (2009). Hataları Ardışık Bağımlı (Otokorelasyonlu) Olan Regresyon Modellerinin Tahmin Edilmesi [Estimating Regression Models with Serially Dependent Errors (Autocorrelation)]. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 23(3), [Page Numbers].

Cite this article: Catuk, C. (2025). Forecasting Turkey's Air Cargo Tonnage: A Comparative Analysis of Statistical Techniques and Machine Learning Methods. Journal of Aviation, 9(1), 109-117.