Economics and Organization of Logistics
10 (4), 2025, 99–113

DOI: 10.22630/EIOL.2025.10.4.27

Luiza Ochnio✉**, Michał Gostkowski, Grzegorz Koszela**
Warsaw University of Life Sciences – SGGW

Application of forecasting methods to passenger traffic volume and load factor prediction: Wizz Air airlines case study

Zastosowanie metod prognozowania do przewidywania ruchu pasażerskiego i wskaźnika wykorzystania miejsc (load factor) na przykładzie linii lotniczych Wizz Air

Abstract. For passenger traffic forecasting on the testing dataset, the best performance is achieved by the additive Winters model, followed closely by the naïve methods. In contrast, the weakest results are observed for the trend model. In the training phase, the lowest errors were obtained for My_ARIMA, Auto_ARIMA, and the additive Winters model, while the naïve approach performed worst. When averaging results across training and testing datasets, the additive Winters model demonstrates the overall lowest MAPE, indicating strong and stable forecasting performance. Forecasting results for the load factor differ slightly from those for passenger volumes. Overall, all models achieve improved accuracy in load factor prediction, suggesting that load factor exhibits more regular and learnable patterns over time. The most accurate forecasts for load factor are produced by Auto_ARIMA and the additive Winters model. Conversely, the trend model yields the highest forecasting errors.

Keywords: air transport, Wizz Air, time series analysis, ARIMA class model, forecasting

Synopsis. Celem przeprowadzonych badań jest ocena i porównanie dokładności prognozowania wybranych metod szeregów czasowych w przewidywaniu ruchu pasażerskiego i współczynnika wykorzystania miejsc na przykładzie linii lotniczej Wizz Air w latach 2015–2025. Analiza oparta jest na publicznie dostępnych historycznych danych linii lotniczych, w tym miesięcznych i kwartalnych raportach

✉ **Luiza Ochnio** – Department of Logistics, Institute of Economics and Finance, Warsaw University of Life Sciences – SGGW; e-mail: luiza_ochnio@sggw.edu.pl; https://orcid.org/0000-0001-8875-7945

Michał Gostkowski – Department of Econometrics and Statistics, Institute of Economics and Finance, Warsaw University of Life Sciences – SGGW; e-mail: michal_gostkowski@sggw.edu.pl; https://orcid.org/0000-0003-3606-1182

Grzegorz Koszela – Department of Logistics, Institute of Economics and Finance, Warsaw University of Life Sciences – SGGW; e-mail: Grzegorz_koszela@sggw.edu.pl; https://orcid.org/0000-0003-4175-4450

publikowanych na stronie relations.wizzair.com. W badaniu zastosowano kilka klasycznych metod prognozowania szeregów czasowych, w tym metody naiwne, model trendu, modele klasy ARIMA (Auto_ARIMA i My_ARIMA), model liniowy Holta oraz model Wintersa w formie addytywnej i multiplikatywnej. W przypadku prognozowania ruchu pasażerskiego na zbiorze danych testowych najlepszą wydajność osiąga addytywny model Wintersa, a zaraz za nim metody naiwne. Z kolei najsłabsze wyniki obserwuje się dla modelu trendu. Wyniki prognozowania współczynnika wykorzystania miejsc różnią się nieznacznie od wyników prognozowania wolumenu pasażerów. Wszystkie modele osiągają większą dokładność w przewidywaniu współczynnika wykorzystania miejsc, co sugeruje, że współczynnik ten wykazuje bardziej regularne i łatwe do zapamiętania wzorce w czasie. Najdokładniejsze prognozy współczynnika wykorzystania miejsca generują Auto_ARIMA i addytywny model Wintersa, natomiast model trendu obarczony jest największymi błędami prognozowania.

Słowa kluczowe: transport lotniczy, Wizz Air, analiza szeregów czasowych, model ARIMA, prognozowanie

JEL codes: C01, C32, R42

Introduction

The dynamic development of low-cost carriers in Central and Eastern Europe over the past two decades has fundamentally transformed the structure of the air transport market. One of the leaders in this segment is the Hungarian airline Wizz Air, which, since its establishment in 2003, has become one of the fastest-growing airlines in Europe. In the 2024/2025 financial year, it carried over 60 million passengers and operated a fleet of more than 200 Airbus A320/A321 family aircraft (www.wizzair.com). Such rapid expansion makes accurate forecasting of passenger demand and load factor a critical component of revenue management, route network optimization, and capacity planning.

Wizz Air is a Hungarian low-cost airline that operates extensive passenger transportation services primarily across Europe, with additional long-haul routes to the Middle East, North Africa, and Central Asia. Established in 2003, the airline has grown to become one of the largest low-cost carriers (LCCs) in the region, known for its cost-efficient operations, high aircraft utilization, and focus on point-to-point services rather than hub-and-spoke networks.

In its passenger transportation model, Wizz Air emphasizes low fares, ancillary revenue, and operational efficiency. The airline's route network connects major, secondary, and underserved airports, particularly catalyzing air connectivity for emerging markets in Eastern Europe. Over the years, Wizz Air has demonstrated strong passenger volume growth, driven by network expansion, increased flight frequencies, and entry into new markets. Its passenger transportation figures have regularly grown as the airline expands both intra-European and select international routes. The airline places strategic emphasis on high load factors (percentage of occupied seats), a key performance indicator in passenger transportation. Higher load factors improve yield and spread fixed operational costs over more passengers.

In choosing this airline, the authors also took into account that despite disruptions such as the COVID-19 pandemic, Wizz Air maintained flexibility through capacity adjustments, route realignment, and disciplined cost controls. This resilience supported recovery in passenger transportation volumes post-pandemic.

Forecasting passenger traffic and load factor is a highly complex task due to the influence of numerous exogenous factors (fluctuations in fuel prices, exchange rates, macroeconomic conditions, geopolitical events, pandemics, and tourism seasonality) as well as endogenous factors (the carrier's pricing policy, marketing activities, schedule changes, and network effects).

The objective of this paper is to analyze and compare the effectiveness of selected forecasting methods in predicting monthly passenger traffic and load factor using Wizz Air as a case study over the period 2015–2025. The analysis employs publicly available historical data from the airline (monthly and quarterly reports published on relations.wizzair.com) and macroeconomic indicators. The methods used omit other indicators; the focus is only on the historical passenger and load factor data.

The paper is structured as follows: a literature review on forecasting in air transport, a description of the data and research methodology, presentation of empirical results (including error metrics such as MAPE), discussion of the findings, and practical conclusions for revenue management in the low-cost segment.

The primary research question is: which forecasting methods perform best under the high volatility typical of low-cost carriers, and to what extent is it possible to achieve predictive advantage using only publicly available data? The results may have direct applications not only for Wizz Air but also for other LCCs operating under a similar business model.

Forecasting in air transport

Accurate demand forecasting is a critical component of air transport planning and management, influencing decisions related to fleet sizing, route development, crew scheduling, infrastructure investment, and revenue management. Time series forecasting methods are widely applied in this domain due to the availability of historical data on passenger volumes, air traffic movements, and cargo flows. Among these methods, the Autoregressive Integrated Moving Average (ARIMA), Holt's exponential smoothing, and Naïve models are frequently used because of their conceptual simplicity, interpretability, and relatively strong performance in short- to medium-term forecasting horizons.

The ARIMA model, introduced by Box and Jenkins [Box et al. 2015], is one of the most established statistical approaches for time series forecasting in air transport. ARIMA combines autoregressive (AR) terms, differencing to achieve stationarity (I), and moving average (MA) components to capture temporal dependencies in the data. In air transport applications, ARIMA models are commonly used to forecast passenger demand, aircraft movements, and air cargo volumes [Tsui et al. 2014]. Seasonal extensions, such as SARIMA, are particularly relevant due to pronounced seasonal patterns in air travel caused by tourism cycles, holidays, and business travel trends. While ARIMA models are powerful, they require careful identification of model orders and assumptions of linearity and stationarity, which may limit their performance during structural breaks such as economic crises or pandemics [Hyndman & Athanasopoulos 2021].

Holt's exponential smoothing method, especially the Holt linear trend model, provides an alternative approach that explicitly captures level and trend components in time series data. This method assigns exponentially decreasing weights to past observations, allowing it to adapt relatively quickly to changes in growth patterns. In air transport forecasting, Holt's method has been applied to passenger traffic and airport throughput data where trends are present but seasonal effects are weak or modeled separately [Dantas et al. 2017]. Compared to ARIMA, Holt's method is computationally simpler and easier to implement, making it attractive for operational forecasting. However, its reliance on smooth trend extrapolation may reduce accuracy when demand exhibits abrupt changes or complex seasonal behavior [Givoni & Banister 2009].

Naïve forecasting models serve as important benchmarks in air transport studies. The simplest Naïve model assumes that future demand equals the most recent observed value, while seasonal Naïve variants repeat observations from the same period in the previous year. Although simplistic, Naïve models often perform surprisingly well in short-term forecasting and are valuable for evaluating the incremental benefit of more complex models. In air transport research, they are commonly used as baseline comparators when assessing advanced statistical or machine learning approaches.

ARIMA, Holt, and Naïve models each play a distinct role in air transport forecasting. ARIMA offers flexibility and strong statistical foundations, Holt's method provides simplicity and adaptability to trends, and Naïve models establish essential reference points [Adiatma et al. 2024]. The choice of model depends on data characteristics, forecast horizon, and operational objectives, and in practice, comparative evaluation is essential to ensure robust forecasting performance. Zachariah et al. [2023], in their systematic review, analyzed passenger demand forecasting techniques ranging from econometric models to deep neural networks and hybrid approaches, considering factors affecting load factor (e.g., seasonality, fuel prices). The authors discussed advantages and limitations of different methods, including applications in low-cost carrier networks, and the findings highlight that hybrid models offer the highest accuracy. As many authors noted [Do et al. 2020; Ramadhani et al. 2020; Al-Sultan et al. 2021; Muros Anguita & Díaz Olariaga 2022; Şimşek et al. 2024], more advanced, multistep models, such as neural networks or hybrid models, improve prediction accuracy compared to the classical methods; however, for preliminary estimation, simpler models can provide initial information, and their results are often comparable to more advanced methods. As it was proved, these classical approaches are foundational for LCC forecasting studies [Zachariah et al. 2023]. In planning passenger traffic and LCC cargo transport, various methods are used that can somehow approximate final forecasts, although this can be difficult, as many unexpected events can affect the actual outcomes. It is impossible to predict various causes of air traffic closures or cases such as a pandemic, but still forecasting is a base for planning and achieving financial profitability. It is particularly important to estimate these parameters for LCCs, as any delay can result in a series of delays; nevertheless, some airlines intentionally follow an overbooking policy because one of the most important business indicators is the load factor, which is the occupancy of seats on a given flight.

Research methodology and analysis

Data collected by Wizz Air were used for the analysis. The data refer to the number of passengers and load factor for the period from April 2014 to April 2025. The original data set has been divided into two parts: train and test. The train dataset contained records up to April 2024, and records from May 2024 (last 12 months) were classified as the test set to assess the accuracy of forecasts. The basic descriptive statistics for passengers and load factor are presented in Table 1. The average passenger volume is 2,884,551.91 (min. value 78,389 and max. value 6,203,673, so the range is 6,125,284), and the average load factor is 86.49% (minimum: 52.20% and maximum 96.30%, so the range is 44.10%). In that case, the better descriptive statistic is the median, which here is equal to 90%, a very high result for the airline.

Table 1. Basic descriptive statistics of analyzed variables
Tabela 1. Podstawowe statystyki opisowe analizowanych zmiennych

Statistic/variable	**Passengers**	**Load factor**
Average	2884551.91	86.49%
Median	2569229.00	90.00%
Standard deviation	1518770.79	9.32%
Kurtosis	–0.79	2.90
Skewness	0.38	–1.85
Range	6125284	44.10%
Minimum	78389	52.20%
Maximum	6203673	96.30%

Source: own preparation
Źródło: opracowanie własne

In the comparative study, selected time series analysis models were used: naïve methods, trend model, ARIMA class models (Auto_ARIMA and My_ARIMA), Holt's linear model, Winters model (additive and multiplicative versions).

Naïve methods

The first method used in the analysis was naïve methods. It is a very simple method and can be used to construct only short-term forecasts (i.e., one period ahead). The hidden assumption is no significant changes in the analyzed process. This method is quite easy to implement and can be used successfully when small random fluctuations are observed in the analyzed variable, but at the same time, the quality of forecasts is rather low.

In the literature, several versions of the naïve method are present. In the study, the following version was used [Cieślak 2005]:

$$y_t^* = y_{t-1} \tag{1}$$

where:
y_t^* – the forecast of the variable Y for the period t,
y_{t-1} – the value of the forecast variable Y in the period t–1.

Trend model

The second group of methods used in the paper was the trend model, where the time variable is the explanatory variable. The time variable represents the influence of unknown factors on the analyzed variable and provides the possibility to describe those changes using the time variable. The basic trend model looks like:

$$y_t^* = f(t) + \varepsilon_t,\ t = 1, ..., \mathrm{n} \tag{2}$$

where:
$f(t)$ – time (trend) function,
ε_t – random variable.

The first step is to determine the function form (i.e., linear, quadratic), and the second is to determine its parameters. The most popular and, at the same time, most used form of trend function is a linear function:

$$y_t^* = \alpha + \beta t \tag{3}$$

Using linear approximation in many situations can be considered too much of a simplification; the quadratic form should be used:

$$y_t^* = \alpha_0 + \alpha_1 t + \alpha_2 t^2, \alpha_2 > 0 \tag{4}$$

The quadratic form has three parameters, thanks to which the model has higher flexibility and can better reflect different nonlinear phenomena. Similar to the naïve method, the trend models can be used to construct short-term forecasts. Using a trend model to construct long-term forecasts may produce large errors [Cieślak 2005].

ARIMA class model

The next model class used in the paper was the ARIMA model. ARIMA models are a very numerous group of time series models and are constructed based on the autocorrelation function [Borkowski and Marcinkowski 1999; Borkowski and Krawiec 2009; Osińska 2006]. They can be used to model stationary or non-stationary processes. In ARIMA, there are three basic components: autoregressive model (AR), moving average model (MA), and autoregressive and moving average models (ARMA). The item "I" in the model name refers to the differencing method [Chrabołowska and Nazarko 2003; Zeliaś et al. 2013]. The ARIMA (p, d, q) means that the final model contains an autoregressive process (order p), a moving average (order q), and is integrated to degree d and can be written as:

$$\Delta^d Y_t = c + \alpha_1 \Delta^d Y_{t-1} + \cdots + \Delta^d Y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q} \tag{5}$$

In the paper, one ARIMA model was created based on ACF and PACF values. To create the second one, the Hyndman–Khandakar algorithm was used. In the first step of the algorithm, the number of differences is determined using repeated KPSS tests. In the second

step, the values of p and q are determined in an iterative approach trying to minimize the AICc value. The Hyndman–Khandakar algorithm can be described as follows:
Step 1. The number of differences $0 \leq d \leq 2$ is determined using repeated KPSS tests.
Step 2. The values of p and q are then chosen by minimizing the AICc after differencing the data d times.

a) Four initial models are fitted:
   ARIMA(0,d,0), ARIMA(2,d,2), ARIMA(1,d,0), ARIMA(0,d,1).
   A constant is included unless $d = 2$. If $d \leq 1$, an additional model is also fitted: ARIMA(0,d,0) without a constant.
b) The best model (with the smallest AICc value) fitted in step (a) is set to be the "current model".
c) Variations on the current model are considered:
   – vary p and/or q from the current model by ±1,
   – include/exclude c from the current model.
   The best model considered so far (either the current model or one of these variations) becomes the new current model.
d) Repeat Step 2(c) until no lower AICc can be found.

Holt's linear model

In the Holt model, the first-order approximation is used in building the forecasts. The Holt model can be used to predict the time series with a trend factor:

$$F_{t-1} = \alpha y_{t-1} + (1-\alpha)(F_{t-2} + S_{t-2}) \tag{6}$$

$$S_{t-1} = \beta(F_{t-1} + F_{t-2}) + (1+\beta)S_{t-2} \tag{7}$$

where:
F_{t-1} – the smoothed value of the predicted variable t–1,
S_{t-1} – the smoothed value of the trend increase at t–1,
α, β – the model parameters with values between 0 and 1.

The values of α, *β* are determined using several experiment of the potential parameters and the best combination is selected to produce the forecasts based on the average error of expired forecasts [Witkowska et al. 2012]. To produce forecasts, the following formula is used:

$$y_t^* = F_n + (t-n)S_n,\ t > n \tag{8}$$

where:
y_t^* – forecast value of variable Y,
F_n – the smoothed value of the predicted variable,
S_n – the smoothed value of the trend increment.

Winters model

The last model group in the paper was the Winters model. The Winters model can be used successfully to analyze processes with a trend defined as a long-term increase or decrease in the data, seasonal fluctuation, and random fluctuation. In the literature, two versions of the model are available. For the additive version, the formula is as follows:

$$F_{t-1} = \alpha\left(y_{t-1} - C_{t-1-r}\right) + \left(1-\alpha\right)\left(F_{t-2} + S_{t-2}\right) \tag{9}$$

$$S_{t-1} = \beta\left(F_{t-1} - F_{t-2}\right) + \left(1+\beta\right)S_{t-2} \tag{10}$$

$$C_{t-1} = \gamma\left(y_{t-1} - F_{t-1}\right) + \left(1-\gamma\right)C_{t-1-r} \tag{11}$$

For multiplicative version:

$$F_{t-1} = \alpha\left(\frac{y_{t-1}}{C_{t-1-r}}\right) + \left(1-\alpha\right)\left(F_{t-2} + S_{t-2}\right) \tag{12}$$

$$S_{t-1} = \beta\left(F_{t-1} - F_{t-2}\right) + \left(1+\beta\right)S_{t-2} \tag{13}$$

$$C_{t-1} = \gamma\left(\frac{y_{t-1}}{F_{t-1}}\right) + \left(1-\gamma\right)C_{t-1-r} \tag{14}$$

where:
F_{t-1} – smoothed value after eliminating seasonal fluctuations,
S_{t-1} – smoothed value of the trend increase,
C_{t-1} – assessment of the seasonality index,
α, β, γ – the parameters for the trend level, trend changes and seasonal fluctuations, respectively (0–1).

After setting the model parameters, to produce forecasts in the additive version when $t > n$, the forecasts are calculated using the formula:

$$y_t^* = F_n + S_n\left(t-n\right) + C_{t-r} \tag{15}$$

$$y_t^* = \left[F_n + S_n\left(t-n\right)\right]C_{t-r} \tag{16}$$

The problematic item is to select which version of the Winters model should be used (i.e., additive or multiplicative) in the analysis. It can be done based on the nature of the seasonality components. If the seasonal component remains at a similar level, it could be treated as a hint to select the additive model. Otherwise, when the seasonal component is not stable (increased or decreased), then the multiplicative version should be considered in the analysis.

Forecasts accuracy

In the paper, to determine model accuracy, the mean absolute percentage error (MAPE) was used:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_t - y_t^*}{y_t}\right|100\% \tag{17}$$

MAPE is the most popular metric used to assess error because the variable's units are scaled to percentage units, making it easier to understand.

Results and discussion

Based on the calculations of each predictive model and using data obtained directly from the Wizz Air airline database, the final results are presented in comparative charts for all applied methods in order to visualize the differences between the forecasts. They were presented separately for both variables: the passenger number forecasts and load factor predictions.

Wizz Air passengers transportation

The final results of Wizz Air passenger transportation in the examined period are presented in Figure 1. The best performance, as the best-fitted line, is for the Winters additive model. It seems that the Winters multiplicative model also largely matches the trend of changes in the number of passengers in reality, but the MAPE is slightly higher in comparison to Winters additive.

Using MAPE within a broader monitoring strategy allows for understanding average error sizes. This assessment is crucial for determining the reliability of forecasts. MAPE is commonly used because it's easy to interpret, and calculating it provides insight into predictive performance – the lower the error value, the better the predictive performance. In Table 2, values of MAPE for train and test datasets are presented. Regarding the Mean Absolute Percentage Error, the best performance for Wizz Air passenger transportation for the testing dataset was for the Winters additive forecasting method (2.358), followed closely by the naïve methods (3.353), while the worst were the trend model and My_ARIMA (10.844 and 12.773, respectively). For training, the best results were for My_ARIMA, Auto_ARIMA, and Winters additive methods, and the worst for Naïve. On average, for training and testing, the lowest MAPE is for the Winters additive model.

In the literature, it is rare to find such a wide comparison of the results of so many models at once on one dataset. Researchers usually compare two or three methods in various environmental conditions or focus on just one. According to other authors, e.g., Dantas et al. [2017], Holt and Winters models have the best performance in forecasting passenger demand, which partly confirms the authors' findings. Those models are in some cases extremely effective, especially when combined. On the other hand, Verma [2025], in his research project, evaluated classic time-series forecasting models including ARIMA/SARIMA, Holt's linear exponential smoothing, and triple exponential smooth-

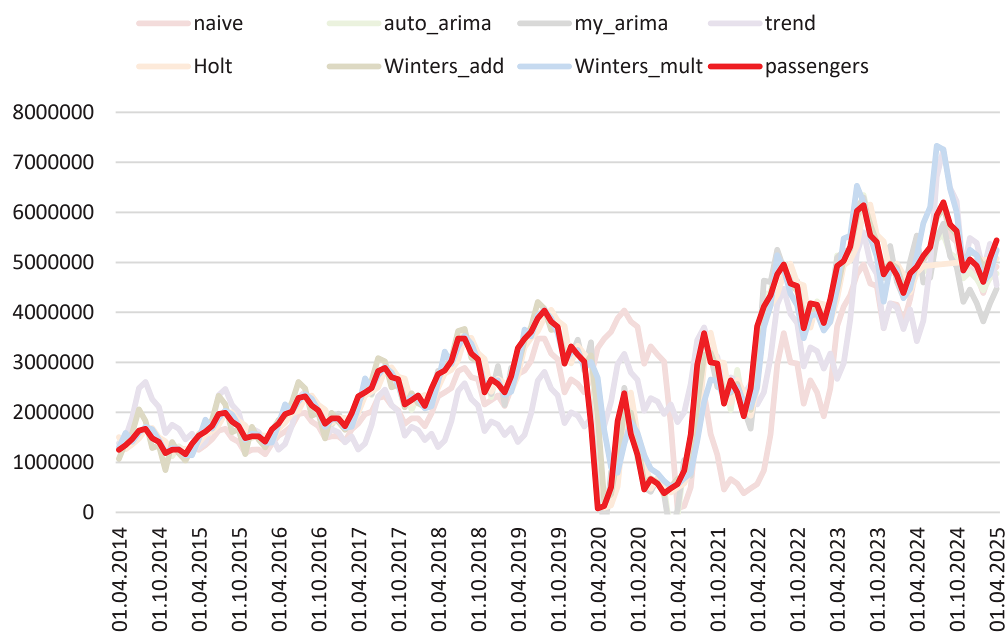

Figure 1. Forecasts created by analyzed models for passenger variable

Rysunek 1. Prognozy wygenerowane przez analizowane modele dla zmiennej: liczba pasażerów

Source: own preparation

Źródło: opracowanie własne

Table 2. MAPE values for testing and training data for passenger variable

Tabela 2. Wartości MAPE dla zestawu danych uczących i testowych dla zmiennej: liczba pasażerów

Method	**Training [%]**	**Testing [%]**
Auto_Arima	28.14656	7.347795
Holts linear	31.985986	7.943642
My_Arima	24.365581	12.77354
Naïve	118.377747	3.35356
Trend model	69.816673	10.844342
Winters additive	28.757749	2.35882
Winters multiplicative	48.306531	8.793606

Source: own preparation

Źródło: opracowanie własne

ing (Holt–Winters) for both passenger and freight traffic in aviation. He finds that Holt's linear model performs very well for passenger demand forecasting relative to alternatives, providing insight into trend-oriented and seasonal forecasting in airline contexts. The very often cited paper written by Tsui et al. [2014] focused on using ARIMA models as the main forecasting method to predict monthly passenger volumes and demonstrated how traditional statistical time series methods perform in capturing historical patterns and predicting future traffic flows. Also, Jafari in 2022, in his study (reviewed in broader literature summaries), discussed the application of both Holt–Winters and ARIMA to air passenger demand data and examined forecasting performance and challenges in volatile airline markets. It includes comparison with baseline approaches (e.g., naïve seasonal forecasts) within aviation contexts, illustrating how classical statistical models fare against structural changes in demand patterns.

Wizz Air load factor

Taking into account the variable load factor, the results turned out to be slightly different than for the number of passengers over the studied years. Generally, all models performed better in this case, as presented in Figure 2. Given the critical importance of maintaining high load factors for low-cost carriers, this finding is particularly relevant.

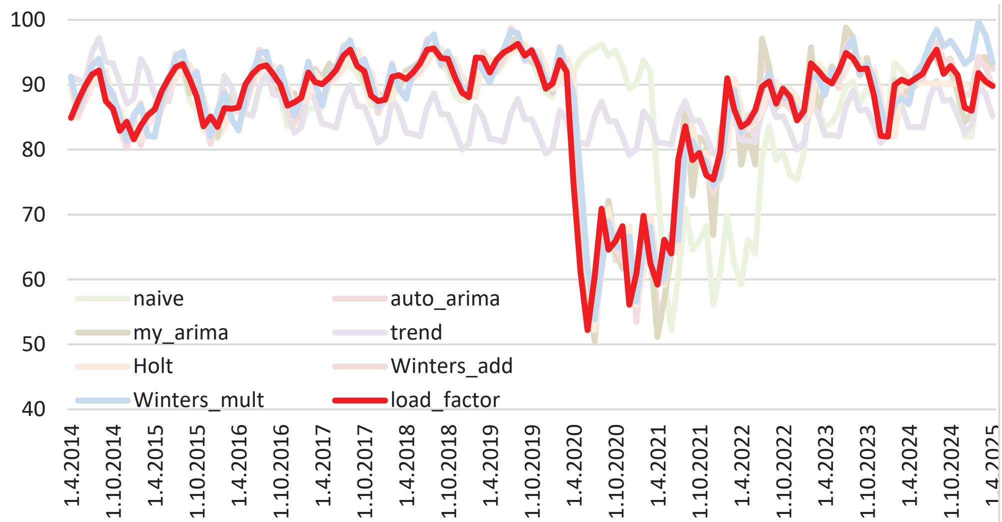

Figure 2. Forecasts created by analyzed models for load factor

Rysunek 2. Prognozy w oparciu o badane modele dla zmiennej: współczynnik wykorzystania miejsc (obłożenia)

Source: own preparation

Źródło: opracowanie własne

The best results were for Auto_ARIMA and Winters additive (for the testing dataset 1.484 and 1.497 respectively, and for training less than 3 in both cases). According to data shown in Table 3, MAPE is the highest for the trend model (testing dataset 5.148 and training 9.120).

Anyway, all methods this time gave quite good results – MAPE less than 5% is considered an indication that the forecast is acceptably accurate; MAPE greater than 10% but less than 25% indicates low but acceptable accuracy; and MAPE greater than 25% indicates very low accuracy [Hyndman & Athanasopoulos 2021]. In Table 3, almost all methods (except the trend model) did not exceed 5%.

Load factor forecasting is commonly addressed using time series methods, as LF exhibits temporal dependence, trend behavior, and often strong seasonality. Among the most widely applied approaches are naïve methods, exponential smoothing models (Holt and Holt–Winters), and ARIMA-class models. Al-Sultan et al. in 2021 compared ARIMA, ETS (Holt and Holt–Winters variants), and naïve models for long-term aviation forecasting. While the primary variable was passenger volume, the authors also discussed implications for load factor planning, showing that smoothing models often provide stable forecasts suitable for capacity utilization analysis.

Table 3. MAPE values for testing and training dataset for load factor variable
Tabela 3. Wartości MAPE dla danych uczących i testowych dla współczynnika wykorzystania miejsc

Method	Training [%]	Testing [%]
Auto_Arima	2.925928	1.484631
Holts linear	3.740248	2.44208
My_Arima	2.898434	1.704027
Naïve	10.232627	1.760718
Trend model	9.120238	5.148948
Winters additive	2.954084	1.497741
Winters multiplicative	3.144627	4.917345

Source: own preparation
Źródło: opracowanie własne

Other researchers [Hyndman & Athanasopoulos 2021] applied trend models, Holt's linear method, Holt–Winters seasonal models, ARIMA, and naïve forecasts. Load factor was calculated as an example of a bounded time series (0–100%), where naïve and exponential smoothing methods performed competitively for short-term airline planning, which is also consistent with our findings. According to Muros Anguita & Díaz Olariaga [2022], who evaluated ARIMA/SARIMA, Holt–Winters, and naïve models as benchmarks against neural networks, forecast accuracy is linked to load factor optimization, and they noted that classical models capture recurring utilization patterns effectively in stable low-cost carrier operations.

Concluding remarks

The purpose of this paper was to evaluate and compare the forecasting accuracy of selected time series methods in predicting passenger traffic volume and load factor. The final modeling framework relied exclusively on historical passenger volumes and load factor data, excluding external explanatory variables. The study applied several classical time series forecasting approaches, including naïve methods, a deterministic trend model, ARIMA-class models, Holt's linear exponential smoothing model, and the Winters model in both additive and multiplicative forms. Forecasting performance was assessed using the Mean Absolute Percentage Error (MAPE) for both training and testing datasets.

For passenger traffic forecasting on the testing dataset, the best performance is achieved by the additive Winters model, followed closely by the naïve methods. In contrast, the weakest results are observed for the trend model. In the training phase, the lowest errors were obtained for My_ARIMA, Auto_ARIMA, and the additive Winters model, while the naïve approach performs worst. When averaging results across training and testing datasets, the additive Winters model demonstrates the overall lowest MAPE, indicating strong and stable forecasting performance. Forecasting results for the load factor differ slightly from those for passenger volumes. Overall, all models achieve improved accuracy in load factor prediction, suggesting that load factor exhibits more regular and learnable patterns over time. The most accurate forecasts for load factor are produced by Auto_ARIMA and the additive Winters model. Conversely, the trend model yields the highest forecasting errors.

Overall, the results indicate that the Winters model and Auto_ARIMA models are particularly well suited for forecasting passenger demand and load factor in the low-cost airline context. The study did not take into account many external factors that other, more complex models might include and also learn from; it aimed to compare the prediction quality in well-known and classical methods. In the literature, researchers have various approaches using various methods; it appears that there is no perfect forecasting method, especially under conditions of such high uncertainty and variability as in LCC [Jin et al. 2020; Firat et al. 2021; Wang et al. 2021; Muros Anguita & Díaz Olariaga 2022). Each author has their own calculations and conclusions depending on the data used, the specifics of the airline, the period under study, and the particular forecasting needs. At the same time, authors who conducted a systematic analysis of research results in recent years also emphasize the wide range of methods developed. Traditional statistical methods (e.g., ARIMA models, linear regression, exponential smoothing) are increasingly being complemented or replaced by advanced machine learning techniques (including Random Forest, XGBoost, LSTM neural networks, Prophet, and hybrid models), which better handle non-linear relationships and large numbers of explanatory variables [Firat et al. 2021; Ehsani et al. 2024]. They also handle uncertainty better in highly variable conditions and reduce error by even 20–30% [Guo et al. 2019; Jin et al. 2020; Jafari & Lewison 2024). Some authors also propose a hybrid model like ML + multiple regression for medium-term demand forecasting [Lundaeva et al. 2024], indicating effectiveness over a five-year horizon for low-cost airlines. Others [He et al. 2023] are proposing a long short-term memory network model for short-term flight booking demand forecasting, which was tested on data from Chinese airlines. The results demonstrate the superiority of deep learning in dynamic pricing and cabin control, increasing revenues in low-cost models through better load factor prediction compared with classical and ML methods. A limitation of this study is the omission of certain external factors that influence passenger behavior. More advanced methods are needed for such modeling. Therefore, the next steps in research, depending on the purpose of the forecast and the acceptable level of forecast error, should also include complex models and techniques related to neural networks.

References

Adiatma T., Mohidin R., Allagabo O., Mustafa O., Yulianti N.L.P., Irianto O., 2024: Forecasting the number of domestic airplane passenger arrivals using the ARIMA model, Global Advances in Business Studies 3(2), 70–82.

Al-Sultan A.T., Al-Rukaibi A., Alsaber A., Pan J., 2021: Forecasting Air Passenger Traffic Volume: Evaluating Time Series Models in Long-Term Forecasting of Kuwait Air Passenger Data, Advances and Applications in Statistics 70(1), 69–89, DOI:10.17654/AS070010069

Alam M.S., Deb J.B., Al Amin A., Chowdhury S., 2024: An artificial neural network for predicting air traffic demand based on socio-economic parameters, Decision Analytics Journal 10, 100382.

Andreoni A., Postorino M.N. V., 2006: A multivariate ARIMA model to forecast air transport demand, Association for European Transport and Contributors, https://www.researchgate.net/publication/267254649_Amultivariate_ARIMA_model_to_forecast_air_transport_demand [accessed: 23.10.2025].

Borkowski B., Krawiec M., 2009: Ryzyko cenowe na rynku surowców rolnych, [in:] M. Hamulczuk, S. Stańko (eds.), Zarządzanie ryzykiem cenowym a możliwości stabilizowania dochodów producentów rolnych – aspekty poznawcze i aplikacyjne, Wydawnictwo SGGW, Warszawa.

Borkowski T., Marcinkowski A., 1999: Bezrobocie w perspektywie socjologicznej, [in:] Socjologia bezrobocia, Wyd. Śląsk, Katowice.

Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M., 2015: Time Series Analysis: Forecasting and Control, Wiley, Hoboken, NJ.

Chen R., Wang L., 2014: Multistep-Ahead Air Passengers Traffic Prediction with Hybrid ARIMA-SVMs Models, The Scientific World Journal, 946026, https://doi.org/10.1155/2014/567246

Chrabołowska J., Nazarko J., 2003: Arima Models in Forecasting Sales Gross Income, Zagadnienia Techniczno-Ekonomiczne 48(3).

Cieślak M. (red.), 2005: Prognozowanie gospodarcze. Metody i zastosowania, PWN, Warszawa.

Dantas T.M., Oliveira F.L.C., Repolho H.M.V., 2017: Air transportation demand forecast through Bagging Holt Winters methods, Journal of Air Transport Management 59, 116–123.

Do Q.H., Lo S., Chen J., Le C., Anh L.H., 2020: Forecasting Air Passenger Demand: A Comparison of LSTM and SARIMA, Journal of Computer Science 16(7), 1063–1084.

Ehsani S., Sergeeva E., Murdy W., Fox B., 2024: Predicting the Skies: A Novel Model for Flight-Level Passenger Traffic Forecasting, arXiv preprint, arXiv:2401.03397.

Firat M., Yiltas-Kaplan D., Samli R., 2021: Forecasting Air Travel Demand for Selected Destinations Using Machine Learning Methods, Journal of Universal Computer Science 27(6), 564–581.

Givoni M., Banister D., 2009: Airline and railway integration, Transport Policy 13(5), 386–397.

Guo X., Grushka-Cockayne Y., De Reyck B., Wang Z., 2019: Forecasting Airport Transfer Passenger Flow Using Real-Time Data and Machine Learning, Harvard Business School Working Paper 19-040.

He H., Chen L., Wang S., 2023: Flight short-term booking demand forecasting based on a long short-term memory network, Computers & Industrial Engineering 186, 109707.

Hyndman R.J., Athanasopoulos G., 2021: Forecasting: Principles and Practice, OTexts, Melbourne.

Jafari N., 2022: The chaos on U.S. domestic airline passenger demand forecasting caused by COVID-19, International Journal of Business Forecasting & Market Intelligence 7(3).

Jafari N., Lewison M., 2024: Forecasting air passenger traffic and market share using deep neural networks with multiple inputs and outputs, Frontiers in Artificial Intelligence 7, 1429341, https://doi.org/10.3389/frai.2024.1429341

Jin F., Li Y., Sun S., Li H., 2020: Forecasting air passenger demand with a new hybrid ensemble approach, Journal of Air Transport Management 83, 101744.

Kanavos A., Kounelis F., Iliadis L., Makris C., 2021: Deep learning models for forecasting aviation demand time series, Neural Computing and Applications 33(21), 14481–14497.

Lundaeva K.A., Saranin Z.A., Pospelov K.N., Gintciak A.M., 2024: Demand Forecasting Model for Airline Flights Based on Historical Passenger Flow Data, Applied Sciences 14(23), 11413.

Muros Anguita J.G., Díaz Olariaga O., 2022: Air passenger demand forecast through the use of Artificial Neural Networks algorithms, International Journal of Aviation, Aeronautics, and Aerospace 9(3), 7.

Osińska M., 2006: Ekonometria finansowa, Polskie Wydawnictwo Ekonomiczne, Warszawa.

Ramadhani S., Dhini A., Laoh E., 2020: Airline Passenger Forecasting using ARIMA and Artificial Neural Networks Approaches, Proceedings of the 7th International Conference on ICT for Smart Society, 1–5.

Şimşek K., Tuğrul N.Ö.Ö., Karaçuha K., Tabatadze V., Karaçuha E. 2024: Modeling and Predicting Passenger Load Factor in Air Transportation: A Deep Assessment Methodology with Fractional Calculus Approach Utilizing Reservation Data, Fractal and Fractional 8(4), 214.

Tsui W.K.H., Balli H., Gilbey A., Gow H., 2014: Forecasting of Hong Kong airport's passenger throughput, Tourism Management 42, 62–76, https://doi.org/10.1016/j.tourman.2013.10.008

Verma N., 2025: Air Passenger and Freight Demand Forecast (Master's thesis/Academic research), National College of Ireland, Dublin.

Wang L., Mykityshyn A., Johnson C., Marple B.D., 2021: Deep Learning for Flight Demand Forecasting, arXiv preprint, arXiv:2011.04476.

Witkowska D., Matuszewska-Janica A., Kompa K., 2012: Wprowadzenie do ekonometrii dynamicznej i finansowej, Wydawnictwo SGGW, Warszawa.

Wizz Air Airlines, 2025: www.wizzair.com [accessed: 23.10.2025].

Zachariah R.A., Sharma S., Kumar V., 2023: Systematic review of passenger demand forecasting in aviation industry, Multimedia Tools and Applications 82(30), 46483–46519, doi:10.1007/s11042-023-15552-1

Zeliaś A., Pawełek B., Wanat S., 2013: Prognozowanie ekonomiczne: teoria, przykłady, zadania, PWN, Warszawa.