Journal of Economics and Administrative Sciences, 26(1): 167-182, 2025
DOI: 10.37880/cumuiibf.1560514

Journal of Economics and Administrative Sciences

| esjournal.cumhuriyet.edu.tr | Founded: 2000 Available online, ISSN: 1303-1279 Publisher: Sivas Cumhuriyet Üniversitesi

Impact of Operational and Financial Efficiency on Aviation Stock Prices: A Machine Learning Model with SHAP Interpretability

Ahmet Akusta[1,a,*]

[1] *Rectorate, Konya Technical University, Konya, Türkiye*
**Corresponding author*

Research Article

History

Received: 03/10/2024
Accepted: 29/12/2024

JEL Codes: G12, G32, C45

ABSTRACT

Using a machine learning approach, this study examines how operational and financial efficiency metrics influence stock prices in the aviation industry. A CatBoost regression model enhanced with SHapley Additive exPlanations (SHAP) was developed using data from 65 global aviation companies collected between 2015 and 2023. The model predicts stock prices based on various operational and financial indicators, including Total Revenue per Available Seat Mile (ASM), Passenger Load Factor, liquidity ratios, and debt-to-assets ratios. The findings suggest that operational efficiency metrics, particularly Total Revenue per ASM and Passenger Load Factor, play a significant role in predicting stock prices within the aviation sector. Financial metrics, such as the Quick Ratio and Debt-to-Assets Ratio, also contribute to the model but appear to have a secondary influence compared to operational factors. SHAP values provided interpretable insights into the model's predictions, allowing for a better understanding of the relative importance of different features. Furthermore, the study's findings offer support for the semi-strong form of the Efficient Market Hypothesis (EMH), demonstrating that operational and financial metrics are reflected in stock prices. These results indicate that aviation companies demonstrating higher operational efficiency may be better positioned for favorable stock market performance, although financial health remains important. This study contributes to the existing literature by integrating operational and financial metrics into a machine learning framework, offering a comprehensive and interpretable model for stock price prediction in the aviation industry.

Keywords: Aviation stock prices, machine learning, SHAP values, operational efficiency, CatBoost

Operasyonel ve Finansal Verimliliğin Havacılık Hisse Senedi Fiyatları Üzerindeki Etkisi: SHAP Yorumlanabilirliğine Sahip Bir Makine Öğrenme Modeli

Süreç

Geliş: 03/10/2024
Kabul: 29/12/2024

Jel Kodları: G12, G32, C45

ÖZ

Bu çalışma, bir makine öğrenimi yaklaşımı kullanarak, operasyonel ve finansal verimlilik ölçütlerinin havacılık sektöründeki hisse senedi fiyatlarını nasıl etkilediğini incelemektedir. SHapley Additive exPlanations (SHAP) ile geliştirilmiş bir CatBoost regresyon modeli, 2015-2023 yılları arasında 65 küresel havacılık şirketinden toplanan veriler kullanılarak geliştirilmiştir. Model, Mevcut Koltuk Kilometre Başına Toplam Gelir (ASM), Yolcu Yük Faktörü, likidite oranları ve borç-varlık oranları dahil olmak üzere çeşitli operasyonel ve finansal göstergelere dayalı olarak hisse senedi fiyatlarını tahmin etmektedir. Bulgular, özellikle ASM başına Toplam Gelir ve Yolcu Yük Faktörü gibi operasyonel verimlilik ölçütlerinin havacılık sektöründeki hisse senedi fiyatlarının tahmininde önemli bir rol oynadığını göstermektedir. Hızlı oran ve borç varlık oranı gibi finansal ölçütler de modele katkıda bulunmakta ancak operasyonel faktörlere kıyasla ikincil bir etkiye sahip görünmektedir. SHAP değerleri, modelin tahminleri hakkında yorumlanabilir bilgiler sağlayarak farklı özelliklerin göreceli öneminin daha iyi anlaşılmasına olanak tanımıştır. Ayrıca çalışmanın bulguları, operasyonel ve finansal metriklerin hisse senedi fiyatlarına yansıdığını göstererek, Etkin Piyasa Hipotezi'nin (EPH) yarı-güçlü formunu desteklemektedir. Bu sonuçlar, finansal sağlık önemini korusa da, daha yüksek operasyonel verimlilik gösteren havacılık şirketlerinin olumlu borsa performansı için daha iyi konumlandırılabileceğini göstermektedir. Bu çalışma, operasyonel ve finansal ölçütleri bir makine öğrenimi çerçevesine entegre ederek havacılık sektöründe hisse senedi fiyat tahmini için kapsamlı ve yorumlanabilir bir model sunarak mevcut literatüre katkıda bulunmaktadır.

Anahtar Kelimeler: Havacılık hisse senedi fiyatları, makine öğrenimi, SHAP değerleri, operasyonel verimlilik, CatBoost

[a] ahmetakusta@hotmail.com 0000-0002-5160-3210

How to Cite: Akusta A (2025) Impact of Operational and Financial Efficiency on Aviation Stock Prices: A Machine Learning Model with SHAP Interpretability, Journal of Economics and Administrative Sciences, 26(1): 167-182, DOI: 10.37880/cumuiibf.1560514

Introduction

The aviation industry, highly sensitive to both internal operational dynamics and external macroeconomic forces, provides a unique setting for examining how efficiency impacts stock performance. Key performance indicators, such as revenue passenger kilometers (RPK), cost per available seat kilometer (CASK), and load factor (LF), have been extensively studied in this context. Increases in RPK are generally associated with stock price growth, while higher CASK and LF reflect inefficiencies in cost management and capacity utilization, adversely affecting stock values (Alici and Sevil, 2022). These findings align with broader discussions in air transport economics, where operational efficiency encompasses aerodynamic, fuel, and cost efficiency, each influencing an airline's ability to maximize revenue within specific market conditions (McLean, 2005).

Operational efficiency's broader implications on stock prices are supported by empirical research. Metrics such as the BOPO ratio negatively influence stock prices, indicating that operational inefficiencies directly affect financial performance (Wulansari et al., 2024). Similarly, profitability and cost efficiency are critical determinants of stock values, with profitability exhibiting a stronger positive impact (Putra et al., 2024). Technological interventions, such as blockchain, have been proposed to enhance operational transparency and efficiency, though market responses to such advancements have been mixed (Gimerská et al., 2023). Additionally, macroeconomic indicators such as inflation and GDP interact with operational metrics to shape stock prices, underscoring the multifaceted relationship between internal efficiency and external economic conditions (Aldabbas et al., 2023).

Financial efficiency also plays a critical role in determining stock prices within the aviation industry. Effective financial management of assets, operating income, and market capitalization enhances financial health and stock performance (Slater, 2023). Hedging strategies, mainly through financial derivatives, have been shown to mitigate operational cost fluctuations, stabilize profitability, and boost stock prices during market instability (Dewikristi Siallagan & Prijadi, 2020). These tools help airlines manage external shocks, such as fuel price volatility and exchange rate fluctuations, critical factors influencing profitability (Yilmaz & Köse, 2023).

Operational cost management is another critical factor influencing stock performance. Research on the Brazilian and Indian aviation sectors highlights the importance of managing labor, aircraft acquisition, and fuel costs, all of which dominate the cost structure in this industry. Efficiently controlling these factors positively impacts profitability and stock valuation (Lopes & Beuren, 2017; Singh et al., 2019). Moreover, operational metrics such as stage length and seats per kilometer are crucial for determining operational efficiency and its effect on stock performance (Singh et al., 2019). These findings align with the Efficient Market Hypothesis (EMH) principles, which asserts that financial asset prices reflect all available information, thereby reducing opportunities for abnormal returns (Okur & Gurbuz, 2015; Rossi, 2016).

The Efficient Market Hypothesis (EMH), initially proposed by Eugene Fama (1970), serves as a foundational theory in finance. EMH posits that stock prices fully incorporate all relevant information, making it impossible to achieve consistent abnormal returns. Fama classified market efficiency into three forms: weak, semi-strong, and strong. Weak-form efficiency suggests that historical prices do not predict future movements, semi-strong form posits that all publicly available information is reflected in prices, and strong form extends this to include insider information (Fama, 1970; Yongxin, 2009). However, anomalies such as speculative bubbles and excess volatility challenge the validity of EMH, particularly during periods of market stress (Sachdeva, 2020). Studies on market anomalies, including calendar effects and return predictability, further question EMH's applicability across different contexts (Rossi, 2016; Woo et al., 2020).

Macroeconomic factors like oil prices, exchange rates, and GDP growth further complicate the relationship between efficiency and stock prices. Depending on global markets, airlines are especially vulnerable to these external forces. Research demonstrates that external factors, such as oil price volatility and GDP growth, often outweigh internal operational metrics, such as load factor or fleet size, in determining profitability (Yilmaz & Köse, 2023). Moreover, regulatory and political environments significantly impact airline performance, with stable conditions fostering better operational efficiency and stock performance (Yadav & Goriet, 2022). Strategic financial tools, such as hedging and diversification, are essential for managing these external pressures and stabilizing financial outcomes (Woo et al., 2020).

Corporate governance also plays a vital role in influencing stock prices within the aviation industry. Airlines with robust governance frameworks, including larger boards and frequent committee meetings, demonstrate stronger financial performance, which translates into higher stock market valuations (Wang et al., 2011). Effective governance enables airlines to manage risks and navigate external shocks effectively, enhancing shareholder value (Lee & Park, 2013). Similarly, strategic alliances and partnerships allow airlines to optimize operational efficiency, expand market reach, and improve profitability, contributing to stronger financial outcomes and stock valuations (Bissessur & Alamdari, 1998).

The aviation industry exemplifies the intricate relationships between operational and financial efficiencies and stock performance. By analyzing these factors within the Efficient Market Hypothesis (EMH) framework, as established by Fama (1970), this study seeks to provide insights into the determinants of stock price movements. This comprehensive exploration highlights how internal metrics and external forces collectively shape financial outcomes in this dynamic and competitive sector.

This paper is structured as follows: The next section reviews existing literature on operational and financial efficiency in the aviation sector and their implications for stock performance. The third section details the data collection process and methodology, emphasizing the CatBoost model and SHAP interpretability framework. Section four presents the findings, focusing on feature importance and the role of operational and financial metrics. Section five's discussion contextualizes existing research results, highlighting practical implications and limitations. Finally, the conclusion synthesizes the key insights and suggests avenues for future research.

Literature Review

Studies have consistently examined the relationship between operational metrics and their implications for airline profitability and stock valuation. Alici and Sevil (2022) provide evidence that key performance indicators such as revenue passenger kilometers (RPK) and load factor (LF) correlate positively with airline stock prices. In contrast, cost per available seat kilometer (CASK) shows a negative association. Similarly, Labantová and Begera (2014) emphasize the importance of operational indicators such as seat occupancy and fuel efficiency in enhancing profitability and influencing stock market performance. These analyses suggest a measurable link between operational data and financial outcomes, providing a basis for further exploration of efficiency drivers in the aviation sector.

The increasing reliance on data-driven methods across industries highlights the transformative impact of machine learning in optimizing business processes. In the food and beverage sector, Şahinbaş (2022) illustrates how machine learning models such as XGBoost and CatBoost can significantly enhance price prediction accuracy, enabling restaurant owners and entrepreneurs to set competitive prices aligned with market expectations. This approach ensures better customer satisfaction and improves profitability by identifying optimal pricing strategies based on specific business features. Similarly, Yüksel (2023) underscores the pivotal role of demand forecasting in the retail sector, demonstrating the superior performance of CatBoost in predicting consumer demand with high accuracy. These findings emphasize the transition from heuristic to quantitative approaches, showcasing the power of advanced algorithms in achieving operational efficiency, reducing costs, and minimizing environmental impacts through waste reduction. Together, these studies reflect the growing necessity for businesses to adopt machine learning to remain competitive, leveraging data science to make informed, actionable decisions.

Studies on flight price prediction and passenger satisfaction have employed advanced machine learning models to uncover nuanced relationships between dependent variables such as flight prices and satisfaction levels and independent factors like time, space, and density. Choudhary et al. (2023) utilize a Random Forest model to predict flight ticket prices, achieving 95% accuracy by integrating variables such as temporal and spatial data. Similarly, Kumar et al. (2024) corroborate the effectiveness of Random Forest, highlighting its generalizability and superior performance compared to baseline models. Both studies emphasize the pivotal role of time and spatial metrics in driving prediction accuracy, suggesting the broader applicability of these factors in aviation pricing strategies. Nagesh et al. (2023) extend this analysis using Decision Tree Regression, demonstrating 85% accuracy in predicting ticket prices, thereby underscoring tree-based models' effectiveness for numerical and categorical data.

In contrast, focusing on passenger satisfaction introduces new dimensions to the analysis. Hong et al. (2023) explore the relationship between satisfaction levels and operational metrics using RF-RFE-LR, revealing time and density as significant predictors. Bhargav and Prabu (2023) further refine this approach by comparing KNN and a Hybrid Random Forest model, identifying the latter as superior in analyzing satisfaction survey data. These studies align in their methodological reliance on machine learning but diverge in application focus—pricing versus satisfaction. While flight price studies emphasize quantitative accuracy in prediction, satisfaction studies delve deeper into subjective metrics, offering insights into consumer experience. Together, they reflect the aviation sector's shift from heuristic approaches to data-driven methods, harnessing advanced algorithms to optimize operational and customer-centric outcomes.

Operational scale and workforce management have also been investigated for their impact on cost behavior. Lopes and Beuren (2017), focusing on Brazilian airlines, highlight cost asymmetry as a significant factor where workforce and fleet size variations directly affect operational expenses. Singh et al. (2019) similarly underline the importance of optimizing variables such as fuel prices, payload, and stage length in reducing costs. These findings illustrate the nuanced relationships between operational strategies and their financial effects, though further studies could refine these conclusions across different market contexts.

The impact of external economic factors on airline financial outcomes has been extensively studied, with varying conclusions about the extent of these influences. Yilmaz and Köse (2023) identify GDP, oil prices, and exchange rates as significant variables affecting profitability and stock prices, often to a greater extent than internal metrics. Complementing this, Yadav and Goriet (2022) address the effects of political and legal conditions on airline performance, offering insights into how external pressures shape operational strategies. Dewikristi Siallagan and Prijadi (2020) explore the role of hedging strategies, particularly fuel price management, in stabilizing financial outcomes, reinforcing that external variables are integral to financial resilience.

The role of financial management in improving airline operations, particularly during structural transitions such as IPOs, has received attention. Lee and Park (2013) note significant improvements in operational metrics like RPK and LF among low-cost carriers (LCCs) following public listings, suggesting that financial restructuring can influence operational efficiency. Amankwah-Amoah (2018) identifies financial constraints as a barrier to operational improvement for African airlines, highlighting regional variations in how financial structuring affects performance. These studies indicate that financial management practices can affect operational outcomes, though the specific mechanisms warrant further examination.

Collaboration among airlines, mainly through alliances, has been identified as a potential driver of operational and financial efficiency. Bissessur and Alamdari (1998) focus on network size and service frequency, suggesting these factors can enhance profitability through improved resource utilization. Zhang et al. (2021) expand on this perspective by integrating operational and financial efficiency metrics through a network DEA approach. While these findings highlight collaborative strategies as a possible avenue for efficiency gains, the effectiveness of such measures may vary depending on market conditions and alliance structures.

Environmental performance has been increasingly recognized for its potential effects on financial and operational outcomes. Seufert et al. (2017) propose an efficiency measure called Luenberger-Hicks-Moorsteen by Briec and Kristiaan Kerstens (2009) that accounts for CO2 emissions, showing that airlines adopting stricter environmental practices achieve improved financial performance. Lee et al. (2017) report similar findings, with higher environmental efficiency scores linked to better profitability and stock performance. These observations align with Saranga and Nagpal (2016), who highlight cost reductions from environmentally focused operational improvements in Indian low-cost carriers. While these studies present compelling evidence of a relationship between environmental strategies and financial outcomes, their applicability across different regulatory contexts remains an open question.

Research has also investigated how airlines respond to macroeconomic volatility and operational disruptions. Piranti (2021) examines the influence of fuel price fluctuations and exchange rates, observing significant effects on stock performance. This perspective aligns with Mantin et al. (2012), who analyze the impact of external shocks like the 9/11 attacks, noting that macroeconomic instability exacerbates operational challenges. Evans and Schäfer (2014) simulate responses to airport capacity constraints, suggesting that scheduling and aircraft size adjustments can mitigate financial impacts. Lee (2023) further identifies network flexibility as a critical factor in absorbing external shocks. These studies collectively underline the importance of strategic planning in managing disruptions, though their implications may vary depending on specific market or operational conditions.

The relationship between operational and financial efficiency and airline performance continues to be a subject of rigorous study, with diverse approaches highlighting both internal and external influences. While operational metrics such as RPK, LF, and CASK are frequently linked to profitability and stock valuations, macroeconomic and regulatory factors also significantly shape outcomes. Financial management, collaborative efforts, and environmental considerations offer additional pathways for improving performance, though the effectiveness of these strategies may differ across markets and regulatory environments. The interconnected nature of these findings underscores the complexity of achieving sustained efficiency and financial stability in the aviation industry. Further research addressing regional and contextual variations could provide deeper insights into these dynamics.

Previous studies have primarily used traditional methods to explore operational and financial metrics' impact on aviation stock prices, often treating these factors separately. This study differs by integrating both metrics into a machine learning framework using CatBoost and SHAP values, providing a unified, interpretable model. This approach offers more profound insights into the relative importance of these factors, bridging the gap between traditional methods and modern data-driven analysis.

Methodology

Data Collection and Preprocessing

The data employed in this study were obtained from the Eikon platform (retrieved on 2 June 2024), encompassing the period from 2015 to 2023 and comprising 65 aviation companies, shown in Table 1, across the globe. Eikon provided comprehensive financial and operational data for each company, which is essential for analyzing the factors influencing stock prices in the aviation industry.

The time frame of 2015-2023 and the selection of 65 companies for this study represent the most comprehensive and reliable dataset available, emphasizing temporal coverage and data integrity. These 65 companies were explicitly chosen for their consistent and high-quality reporting, ensuring robust and valid analysis. This focus on data integrity may result in the exclusion of smaller companies with incomplete records, prioritizing the inclusion of firms with the most accurate and comprehensive data.

Data collection was finalized in 2024, with the latest available data extending to 2023. Significant gaps and inconsistencies were observed in records before 2015, which would have compromised the analysis's robustness and comparability. By contrast, the period from 2015 onwards provides the most extensive and complete records for operational and financial variables, ensuring consistency across the selected companies.

The dataset included operational and financial variables used as predictors in the model. As seen in Table 2, the critical variables selected for this study were:

Table 1. Aviation Companies

#	Identifier	Company Name	#	Identifier	Company Name
1	LHAG.DE	Deutsche Lufthansa AG	34	CAPI.KL	Capital A Berhad
2	0293.HK	Cathay Pacific Airways Ltd	35	JET.NS	Jet Airways (India) Ltd
3	ALK.N	Alaska Air Group Inc	36	REX.AX	Regional Express Holdings
4	AIRF.PA	Air France KLM SA	37	CPA.N	Copa Holdings SA
5	9201.T	Japan Airlines Co Ltd	38	CHR.TO	Chorus Aviation Inc
6	PAL.PS	PAL Holdings Inc	39	UAL.OQ	United Airlines Holdings Inc
7	600115.SS	China Eastern Airlines Corp	40	ALGT.OQ	Allegiant Travel Co
8	SKYW.OQ	SkyWest Inc	41	601111.SS	Air China Ltd
9	LUV.N	Southwest Airlines Co	42	AC.TO	Air Canada
10	THYAO.IS	Turk Hava Yollari AO	43	ICEAIR.IC	Icelandair Group hf
11	HA.OQ	Hawaiian Holdings Inc	44	AIRA.DU	Air Arabia PJSC
12	SIAL.SI	Singapore Airlines Ltd	45	DAL.N	Delta Air Lines Inc
13	THAI.BK	Thai Airways International	46	AGNr.AT	Aegean Airlines SA
14	AIR.NZ	Air New Zealand Ltd	47	CEB.PS	Cebu Air Inc
15	FIA1S.HE	Finnair Oyj	48	JAZK.KW	Jazeera Airways Co KSCP
16	9202.T	ANA Holdings Inc	49	PGSUS.IS	Pegasus Hava Tasimaciligi
17	2610.TW	China Airlines Ltd	50	AIRX.KL	AirAsia X Bhd
18	003490.KS	Korean Air Lines Co Ltd	51	SAVE.N	Spirit Airlines Inc
19	AFLT.MM	Aeroflot-Rossiyskiye Avialinii PAO	52	ICAG.L	International Consolidated Airlines Group SA
20	SPJT.BO	Spicejet Ltd	53	601021.SS	Spring Airlines Co Ltd
21	QAN.AX	Qantas Airways Ltd	54	GIAA.JK	Garuda Indonesia (Persero)
22	LTM.SN	LATAM Airlines Group SA	55	AVT_p.CN	Avianca Holdings SA
23	RYA.I	Ryanair Holdings PLC	56	BA.BK	Bangkok Airways PCL
24	UTAR.MM	Aviakompaniya UTair PAO	57	AQZ.AX	Alliance Aviation Services
25	2618.TW	Eva Airways Corp	58	AAV.BK	Asia Aviation PCL
26	EZJ.L	Easyjet PLC	59	603885.SS	JUNEYAO AIRLINES Co Ltd
27	600221.SS	Hainan Airlines Holding Co Ltd	60	VOLARA.MX	Controladora Vuela Compania de Aviacion SAB de
28	020560.KS	Asiana Airlines Inc	61	AAL.OQ	American Airlines Group Inc
29	JBLU.OQ	JetBlue Airways Corp	62	WIZZ.L	Wizz Air Holdings PLC
30	ELAL.TA	El Al Israel Airlines Ltd	63	AZUL.N	Azul SA
31	600029.SS	China Southern Airlines Co	64	INGL.NS	Interglobe Aviation Ltd
32	NAS.OL	Norwegian Air Shuttle ASA	65	089590.KS	JejuAir Co Ltd
33	GOLL4.SA	Gol Linhas Aereas Inteligentes SA			

Table 2. Variables

Variables	Definiton	Literature Reference
Total Revenue per ASM (USD)	Refers to the total operating revenue (which may include passenger, cargo, and other revenues) divided by the available seat miles, providing a measure of revenue generation relative to capacity provided (Zou et al., 2015).	Lee (2019) explored the effect of operational performance on financial performance, focusing on revenue per available seat mile (ASM) as a significant factor. This variable captures the operational revenue efficiency of airlines.
Passenger Load Factor	Calculated as the total Revenue Passenger Miles (RPMs) divided by the Available Seat Miles (ASMs), expressed as a percentage. It indicates the efficiency of an airline in filling available seating capacity (Zou et al., 2015).	Alici & Sevil (2022) found that the passenger load factor significantly influences stock prices, indicating that operational efficiency impacts financial performance. Yılmaz & Köse (2023) also highlighted the passenger load factor as a critical driver of profitability in air transportation.
Quick Ratio	The Quick Ratio is a stringent test of liquidity that considers assets most easily convertible to cash, excluding inventories. The formula: Quick Ratio = (Cash + Marketable Securities + Accounts Receivable) / Current Liabilities. It measures the ability to meet short-term obligations without relying on inventory (Subramanyam and Wild, 2014).	Tanrıverdi *et al*. (2023) analyzed financial performance, including the quick ratio, as part of sustainability evaluations in airlines, showing its role in liquidity assessment. Sumerli Sarıgül *et al*. (2023) used the quick ratio to assess the financial stability of European airlines.
Current Ratio	Measures a company's ability to meet short-term liabilities with short-term assets. Formula: Current Ratio = Current Assets / Current Liabilities. It provides a safety margin for covering potential shortfalls in non-cash current assets (Subramanyam and Wild, 2014).	Tanrıverdi *et al*. (2023) emphasized the importance of the current ratio in assessing airline liquidity and operational sustainability. Sumerli Sarıgül *et al*. (2023) incorporated the current ratio in evaluating European airlines' financial health.
Tot Debt/Tot Assets, %	Indicates the proportion of a company's assets financed by debt. Formula: Debt to Total Assets Ratio = Total Debt / Total Assets. A higher ratio reflects greater financial leverage and potential insolvency risk (Subramanyam and Wild, 2014).	Tanrıverdi *et al*. (2023) utilized the debt-to-asset ratio to assess airlines' financial leverage and its impact on sustainability during the COVID-19 pandemic. Sumerli Sarıgül *et al.,* (2023) included this variable in their financial performance analysis of European airlines.
Return On Assets	Measures the efficiency of using assets to generate profits. Formula: ROA = Net Profit/ Total Assets. ROA provides insight into profitability relative to the company's asset base (Subramanyam and Wild, 2014).	Alici & Sevil (2022) demonstrated that return on assets (ROA) directly influences stock prices, making it a critical variable in operational and financial performance evaluations. Lopes *et al*., (2016) examined the role of ROA as a profitability driver among the top 30 global airlines.
Asset Turnover	Reflects the efficiency of utilizing assets to generate sales. Formula: Asset Turnover = Sales / Average Total Assets. High turnover indicates the effective use of assets in generating revenue (Subramanyam and Wild, 2014).	Sumerli Sarıgül *et al*., (2023) considered asset turnover a vital indicator of operational efficiency in European airlines' financial performance. Lee *et al*. (2019) also included asset turnover in their operational and financial performance analysis post-IPO.
Beta	Beta measures a stock's volatility or systematic risk relative to the overall market. A beta of 1 indicates that the stock moves in line with the market; greater than 1 indicates higher volatility and less than 1 indicates lower volatility (Fama and French, 1992).	Alici & Sevil (2022) analyzed the Beta of airline stocks as a measure of volatility and its relation to operational efficiency, further linking it to stock price movement.
Net Profit Margin, %	Measures the profit generated per dollar of sales (operating revenues). Formula: Net Profit Margin = Net Profit/ Net Sales (Operating Revenues). A higher ratio indicates greater efficiency in generating profits from revenues and controlling expenses (Subramanyam and Wild, 2014).	Yılmaz & Köse (2023) identified net profit margin as a significant determinant of profitability in the airline industry, highlighting the influence of internal and external factors. Pamungkas & Suhadak (2017) explored how jet fuel prices and macroeconomic variables impact net profit margins in the Asian airline sector.
Number of Planes, Prd-Prd Diff	This refers to the difference in the number of planes (fleet size) operated by an airline between two consecutive periods (Zou et al., 2015).	Lopes & Beuren (2017) examined the influence of fleet size changes (number of planes, Prd-Prd Diff) on operational efficiency, linking it to cost behavior in Brazilian airline companies.

Following Table 2, the selected variables were chosen to provide a balanced assessment of operational and financial factors crucial to airline performance. Return on Assets (ROA) and Net Profit Margin were key profitability indicators, highlighting how effectively airlines utilize assets and control costs to generate income. ROA is particularly relevant in the asset-intensive aviation sector, offering a clear measure of asset management efficiency. Liquidity ratios, including the Quick Ratio and Current Ratio, were selected to evaluate the ability to meet short-term financial obligations, a critical factor in maintaining stability during economic fluctuations. The Debt-to-Assets Ratio was included to assess financial leverage and long-term solvency, while Asset Turnover was chosen to measure how effectively revenue is generated relative to assets. Although other indicators, such as Return on Equity (ROE), were considered, ROA was prioritized for its direct relevance to operational efficiency. The variables were selected based on their relevance to airline operations, prevalence in prior literature, and ability to collectively provide a comprehensive view of financial health and operational efficiency.

The preservation of data integrity is of paramount importance. The application of interpolation techniques maintains the continuity and integrity of datasets by estimating missing values based on known data points, thus preventing abrupt changes that could mislead analyses (Arp et al., 2022; Huang, 2021). In this study, linear interpolation was performed using the 'interpolate' method from the pandas library (McKinney, 2010). This method estimates missing values by fitting a straight line between adjacent data points, ensuring a seamless integration of missing values into the dataset while maintaining data trends and consistency. This process enhances the precision of statistical models and analytical procedures by incorporating missing values. For example, linear interpolation has been demonstrated to optimize the performance of models such as ARIMA and multivariate regression, resulting in more accurate predictions and reduced errors (Xu & E, 2020).

To ensure the integrity and continuity of the time-series data, interpolation was selected as the preferred method for handling missing values. Interpolation is particularly effective in maintaining the consistency of trends over time, which is essential for time-series analysis. After cleaning, the individual datasets for each of the 65 airlines were aggregated into a single unified dataset. A "Company" identifier was introduced as a categorical variable for company-specific factors. This addition ensures that the model can distinguish between different airlines and capture any variations in performance or stock price trends unique to each company. This aggregation allowed for constructing a panel dataset, combining cross-sectional and time-series data, facilitating more comprehensive analyses across both dimensions.

Descriptive statistics for the dataset are shown in Table 3, highlighting the robust structure and integrity of the data used in this study.

Table 3. Descriptive Statistics

Variable	Mean	Std Dev	Min	25%	50%	75%	Max
Beta	1,226	0,597	-0,192	0,798	1,169	1,563	3,398
Asset Turnover	0,668	0,329	0,055	0,438	0,642	0,841	2,214
Return On Assets	0,012	0,086	-0,531	-0,019	0,030	0,063	0,196
Net Profit Margin, (%)	-0,04	2,112	-29,73	-0,060	0,030	0,078	40,30
Tot Debt/Tot Assets	0,456	0,260	0,000	0,311	0,433	0,578	3,480
Current Ratio	0,850	0,463	0,010	0,516	0,781	1,078	2,975
Quick Ratio	0,790	0,443	0,010	0,470	0,713	1,013	2,676
Combined Alpha Model Rank	53,41	32,22	1,000	23,000	57,00	82,00	100,00
Passenger Load Factor	0,787	0,105	0,133	0,764	0,816	0,843	0,960
Number of Planes, Prd-Prd Diff	8,348	33,06	-250,0	0,000	6,000	17,00	317,00
Total Revenue per ASM (USD)	0,135	0,097	0,000	0,097	0,121	0,148	1,915
Price Close (USD)	20,24	68,02	0,004	0,722	3,603	18,98	986,99

Catboost Model

CatBoost CatBoost is an advanced implementation of gradient boosting on decision trees designed to handle categorical features efficiently and mitigate prediction biases. Unlike traditional gradient boosting models, CatBoost introduces innovative mechanisms like ordered boosting and specialized categorical feature handling to address issues of target leakage and prediction shift (Prokhorenkova et al., 2019)

CatBoost improves efficiency and reduces sensitivity to noise in target variables by discretizing continuous targets into quantized categories. This approach enhances the model's stability during training and minimizes overfitting risks (Prokhorenkova et al., 2019). Similar mechanisms are employed in related frameworks like LightGBM, which utilizes gradient-based one-side sampling (GOSS) to prioritize instances with large gradients, achieving comparable improvements in computational efficiency (Ke et al., 2017)

The model's automated handling of categorical features eliminates the need for extensive preprocessing, streamlining the modeling pipeline. CatBoost uses ordered target encoding, which assigns unique numerical representations to categorical variables based on historical data without introducing target leakage. This approach aligns with methods like target statistics in other gradient boosting frameworks but incorporates permutations to ensure data consistency (Prokhorenkova et al., 2019; Dorogush et al., 2018)

CatBoost incorporates intrinsic feature prioritization through its tree-based structure, automatically selecting the most relevant features during training. Similarly, LightGBM employs a gradient-based split criterion to dynamically adjust feature relevance across iterations, showcasing a shared emphasis on optimizing feature utilization (Ke et al., 2017; Prokhorenkova et al., 2019)

CatBoost's automation of preprocessing tasks, such as encoding categorical variables and handling missing data, is another key feature. CatBoost's tree-based structure inherently handles numerical data magnitudes, unlike methods that require normalization or scaling. This characteristic simplifies preprocessing while preserving model performance. LightGBM similarly accelerates training processes through feature bundling techniques, such as exclusive feature bundling (EFB), which groups sparse features into bins to optimize computational efficiency (Ke et al., 2017)

By leveraging these techniques, CatBoost strives to balance performance and interpretability, making it a strong candidate for handling heterogeneous and noisy datasets. Its focus on automating preprocessing, reducing prediction bias, and improving computational efficiency positions it as a promising solution for addressing complex machine learning challenges.

Modeling Process

In building the predictive model, CatBoost was selected due to its superior handling of categorical features without requiring extensive preprocessing. CatBoost employs an innovative algorithm that prevents prediction shifts caused by target leakage, making it highly effective in dealing with categorical data (Panigrahi *et al*., 2022). One of its main advantages is its ability to automatically manage categorical features and missing values during training, rather than relying on preprocessing, which streamlines the overall modeling process (Zhao *et al*., 2021).

The initial CatBoost model was trained with the default hyperparameters. The results from the baseline model are summarized in Table 4, which reports the mean squared error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). These metrics are critical for evaluating the model's overall performance.

The MSE measures the quality of a model by calculating the average squared difference between predicted and actual values. It evaluates how close predictions are to the observed outcomes and is minimized by the regression function, which provides the best fit for squared error. The MAE measures the average absolute difference between predictions and actual values, offering a more straightforward and interpretable error assessment by focusing on the magnitude of discrepancies without squaring them. The MAPE evaluates the average percentage difference between predicted and actual values, making it particularly useful for comparing errors across different scales or datasets. MAPE normalizes errors relative to the actual values, emphasizing relative accuracy rather than absolute differences (De Myttenaere et al., 2016).

The current model configuration, summarized in Table 2, characterized by a depth of 6, L2 leaf regularization of 3, and a learning rate of 0.06, demonstrates a balance between model complexity, regularization to mitigate overfitting, and stability in learning. MSE of 285.79 suggests that the squared prediction errors are notable, indicating some variability in model accuracy. MAE of 9.42 points to an average prediction deviation of 9.42 units, which provides a more interpretable measure of performance but also signals room for potential improvement. MAPE of 4.7% suggests that, on average, predictions deviate by 4.7% from actual values, indicating relatively strong performance in terms of proportional accuracy. While these results demonstrate reasonable effectiveness, particularly in percentage-based metrics, the magnitude of absolute errors suggests opportunities for further optimization.

Table 4. Baseline Model Results

depth	l2 leaf reg	Learning rate	MSE	MAE	MAPE
6	3	0,06	285,788057	9,42244375	4,7034661

Table 5. Hyperparameter Search Space

Hyperparameter	Values
Depth	[4, 6, 8]
Learning Rate	[0,01 0,02 0,03 0,04 0,05 0,06 0,07 0,08 0,09 0,1]
L2 Leaf Reg	[1, 3, 5]

After performing hyperparameter tuning, the following hyperparameters were found to yield better results:

Table 6. Optimized Model Results

depth	l2 leaf reg	Learning rate	MSE	MAE	MAPE
8	5	0,02	281,7407	9,383539	3,107654

Hyperparameter Tuning

Hyperparameter tuning was employed to explore multiple combinations of key parameters efficiently and enhance the predictive capabilities of the CatBoost model. The primary hyperparameters tuned included the learning rate, L2 leaf regularization, and tree depth. Each parameter is critical in balancing model complexity, generalization, and convergence.

The learning rate determines the magnitude of model parameter updates during training. Proper tuning is essential, as a meager learning rate can lead to slow convergence, while a high learning rate risks divergence from the optimal solution (Konar et al., 2020). A smaller learning rate allows the model to converge more gradually, exploring the solution space in finer increments, thereby improving generalization. However, this approach often requires an increase in the number of iterations to achieve satisfactory performance. In contrast, a larger learning rate accelerates convergence but may overshoot the optimal solution, degrading model performance. CatBoost recommends a smaller learning rate combined with more iterations for most tasks to achieve a stable and accurate model (Dorogush et al., 2017).

L2 leaf regularization, controlled by the l2_leaf_reg parameter, is applied to penalize large leaf weights in the decision trees. This constraint smooths predictions by encouraging the model to focus on general patterns rather than dataset-specific nuances, thereby reducing the risk of overfitting (Satter et al., 2023). Larger values of L2 regularization strengthen this effect, improving generalization on noisy datasets, but excessively high values may lead to underfitting. Optimal tuning of this parameter involves a trade-off between bias and variance, typically achieved through experimentation and cross-validation (Dorogush et al., 2017). Additionally, CatBoost's use of "oblivious trees" (where all nodes at the same level test the same feature) enhances regularization effects, improving efficiency and robustness.

The depth of the decision trees, determined by the depth parameter, controls the model's complexity and ability to capture feature interactions. A greater depth allows the model to represent more intricate patterns in the training data, which can improve performance on complex datasets. However, deeper trees also increase the risk of overfitting, mainly when training data contains noise or has limited samples (Astrup et al., 2008). Conversely, a smaller depth leads to simpler trees that may generalize better but risk underfitting the training data. Striking the right balance in tree depth is crucial, as it depends on the dataset's complexity. CatBoost defaults to a depth of 6, often serving as a practical starting point for many tasks (Dorogush et al., 2017).

The hyperparameter tuning search space was defined in Table 5:

As seen in Table 6, the optimized model configuration, with a depth of 8, L2 leaf regularization of 5, and a reduced learning rate of 0.02, reflects an approach aimed at enhancing the model's performance through more complex representation, stronger regularization, and finer-grained updates during training. MSE of 281.74 indicates a slight improvement in the average squared prediction error compared to the initial configuration, suggesting better precision in some predictions. MAE of 9.38 represents a modest reduction in the average absolute error, implying an incremental improvement in overall prediction accuracy. Notably, MAPE has decreased to 3.11%, reflecting a substantial enhancement in the model's proportional accuracy and effectiveness in capturing relative differences. These results suggest that the hyperparameter adjustments have improved predictive performance, particularly in percentage-based evaluations while maintaining reasonable absolute error metrics. This highlights the model's increased reliability in predicting stock prices across diverse companies.

SHAP (SHapley Additive exPlanations)

SHAP (SHapley Additive exPlanations) enhances the interpretability of machine learning models, addressing the need for transparency in AI systems. SHAP is a robust framework for interpreting machine learning models, grounded in game theory. It applies to the SHAP value concept, a measure from cooperative game theory, to attribute the contribution of individual features to model predictions in a mathematically sound and interpretable manner (Lundberg & Lee, 2017).

The SHAP value quantifies a feature's contribution by evaluating its marginal impact on the prediction. This involves calculating the difference in the model's output with and without the feature across all possible subsets of features that exclude the one being analyzed. This comprehensive approach ensures that the feature's influence is contextualized relative to all potential feature combinations (Lundberg & Lee, 2017).

To achieve a fair allocation of contributions, the computation incorporates:

- Subset Evaluation: Marginal contributions are determined by treating subsets of features as coalitions and comparing model outputs with and without the feature for each subset.
- Permutational Averaging: Contributions are averaged over all permutations of feature orderings, guaranteeing that the evaluation sequence does not influence the results.

- Combinatorial Weighting: Each subset's influence is scaled based on its size relative to the total feature set, ensuring a proportional representation of the feature's impact.

While this approach is rigorous, calculating exact SHAP values involves factorial computational complexity, which limits scalability for datasets with many features. To address this, SHAP incorporates model-specific optimizations, such as tailored algorithms for tree-based models and neural networks, to enhance efficiency. For general models, techniques like Kernel SHAP use local approximations to estimate contributions while preserving theoretical integrity (Lundberg & Lee, 2017).

SHAP's foundation in fairness and consistency makes it a cornerstone of explainable AI, particularly valued for its ability to provide transparent feature attributions. However, practical challenges remain, particularly in managing computational demands for high-dimensional data (Lundberg & Lee, 2017).

Feature Importance

Feature importance helps interpret machine learning models by quantifying the contribution of each feature to the model's predictions. However, traditional methods like gain, commonly used in tree-based models, must be more consistent and reliable in specific contexts (Nohara *et al.*, 2021). This inconsistency can lead to misleading interpretations of the model's behavior, especially when dealing with complex datasets.

To address these issues, Lundberg *et al*. (2018) proposed SHAP values, which offer a more robust and consistent approach to feature importance. SHAP values are grounded in game theory and provide a unified framework for local explanations by considering all possible feature combinations and their marginal contributions to the outcome. This method ensures that feature importance is calculated fairly and consistently, making it easier to interpret the impact of each variable on the model's predictions. SHAP is particularly valuable for high-stakes industries, such as aviation, where accurate interpretability of models is essential for decision-making.

The SHAP decision plot provides an in-depth visualization of how individual feature values contribute to a single prediction, offering valuable insights into the model's behavior for specific companies. It effectively allows analysts to explore the influence of different variables on the model's predictions, enhancing interpretability.

In Figure 1, the Y-axis lists the features, ranked by their importance in the model. Features higher on the axis impact the model's predictions more. For instance, the company variable is the most significant feature, followed by Total Revenue per ASM, Passenger Load Factor, and other features. On the X-axis, SHAP values represent the influence of each feature on the model's output, such as stock price predictions. Positive SHAP values indicate that a feature increases the prediction, pushing it higher, while negative SHAP values suggest that the feature lowers the prediction.

Another critical aspect of the decision plot is the color bar, which visually represents the feature values. Red points indicate higher feature values, while blue points represent lower values. This color scheme helps quickly identify feature values' impact on the model's predictions. For example, high values of Total Revenue per ASM, represented in red, generally push the model's output higher, while lower values (in blue) tend to decrease it. Similarly, higher values increase the predicted output for the Passenger Load Factor, and lower values reduce it.

Some features, such as Total Debt to Total Assets and Current Ratio, exhibit a more complex relationship with the model's target variable. Their high or low values can influence the predictions in either direction, indicating that their impact is more complex. In contrast, the Net Profit Margin demonstrates a more precise pattern, where higher values (red) push the prediction upward, while lower values (blue) decrease the predicted outcome.

The SHAP summary, shown in Table 7, plot shows how each feature contributes to the model's predictions. Features at the top, like Company and Total Revenue per ASM (USD), are the most influential. The colors indicate how feature values interact with SHAP values and help understand whether high or low feature values drive higher or lower predictions.

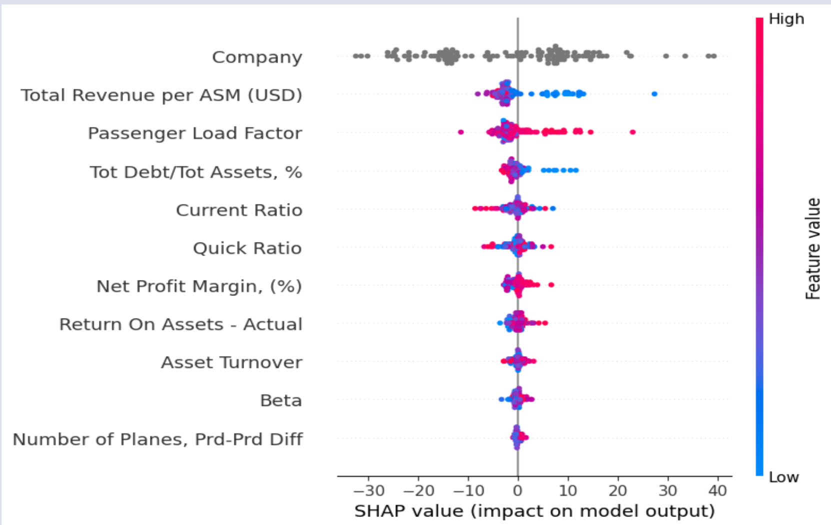

Figure 1. SHAP Values

Table 7. Model Importance Values

Feature	Importance %
Company	30,899796
Total Revenue per ASM (USD)	28,454733
Passenger Load Factor	15,601517
Quick Ratio	5,530049
Current Ratio	4,610544
Tot Debt/Tot Assets, %	4,498126
Return On Assets	2,944894
Asset Turnover	2,659982
Beta	2,306306
Net Profit Margin, (%)	1,358261
Number of Planes, Prd-Prd Diff	1,135791

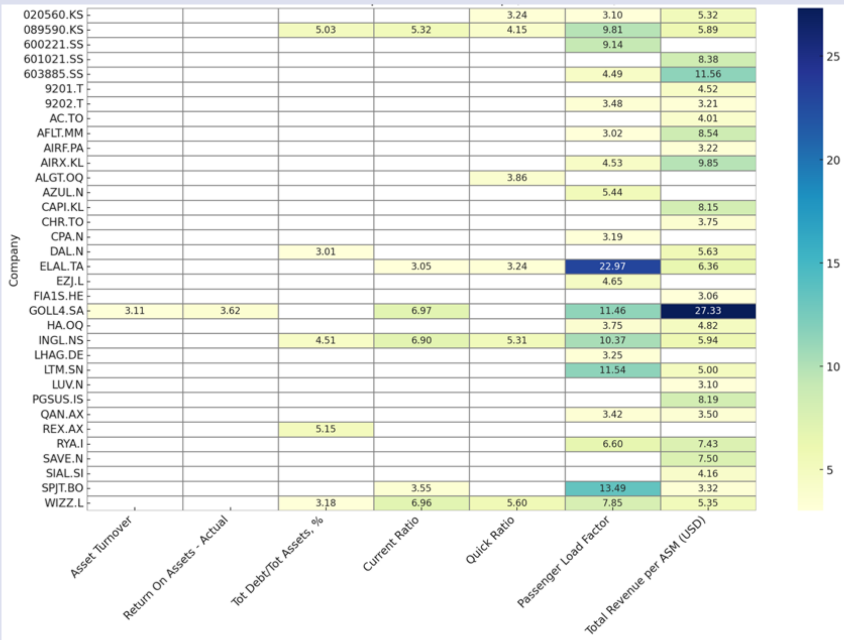

Figure 2. Model Importance Across Companies

This section examines the importance of features derived from the machine learning model in predicting airline stock prices. By leveraging SHAP values, we can quantify each feature's contribution to the model's predictions, offering a transparent view of how operational and financial metrics influence stock prices in the aviation industry. Additionally, a feature importance heatmap provides a granular breakdown of how these variables impact individual airlines.

The Company variable is the most influential feature, with an importance value of 30.90%. This highlights substantial airline stock price behavior variations driven by management strategies, market conditions, and operational practices. These company-specific differences underscore the need for granular analysis, as internal and external factors unique to each airline significantly impact stock performance.

Total revenue per ASM and passenger load factor are critical indicators of revenue generation and operational efficiency, with values of 28.45% and 15.60%, respectively. Revenue per ASM reflects an airline's ability to maximize earnings relative to its seating capacity, signaling strong financial and operational health. Similarly, high passenger load factors indicate efficient fleet utilization, directly contributing to profitability and robust stock performance. These metrics collectively emphasize that operational success is pivotal in determining stock prices.

Financial health metrics assess short-term liquidity and stability, particularly the Quick Ratio (5.53%) and Current Ratio (4.61%). The Quick Ratio's focus on immediate financial solvency is especially valuable during economic uncertainty, while the Current Ratio offers a broader perspective on financial resilience. These indicators highlight the importance of liquidity management in an industry prone to cash flow volatility and external disruptions.

Debt-to-Assets Ratio (4.50%) reflects that debt management plays a moderate role. While debt can fuel growth and expansion, excessive reliance on borrowing introduces financial risk, affecting investor confidence. This suggests a nuanced impact, where the context of debt levels and operational success determines its effect on stock prices.

Supporting metrics such as ROA (2.94%), Asset Turnover (2.66%), and Net Profit Margin (1.36%) provide insights into profitability and asset utilization. Although these metrics contribute to understanding financial health, their lower importance than liquidity and operational efficiency suggests that investors prioritize revenue generation and financial stability over short-term profitability.

Market risk, represented by Beta (2.31%), holds a minor yet notable role. While Beta informs on stock price volatility, its relatively low importance indicates that aviation investors focus more on operational and financial fundamentals than market risk.

Finally, the Number of Planes (Prd-Prd Diff) has the least impact (1.14%). While fleet size changes reflect strategic decisions, they are overshadowed by more immediate factors like revenue generation, operational efficiency, and financial stability.

In Figure 2, the feature importances for individual companies reveal insightful patterns regarding the predictive power of various metrics. Among these, operational efficiency indicators such as Total Revenue per ASM (Available Seat Mile) and Passenger Load Factor emerge as dominant features for several companies. These metrics indicate how effectively airlines manage their capacity and revenue generation, and they are essential for companies like Gol Linhas Aéreas Inteligentes and El Al Israel Airlines Ltd. The prominence of these features underscores the central role that efficient capacity utilization and revenue management play in predicting stock price movements for these firms. Such insights highlight the strong relationship between operational performance and financial outcomes in the aviation sector.

In addition to operational metrics, financial health indicators, particularly liquidity ratios like the Current Ratio and Quick Ratio, are shown to be highly influential for other companies such as InterGlobe Aviation Ltd. and Gol Linhas Aéreas Inteligentes. These features suggest that short-term financial stability is crucial for investors when assessing stock price performance. Liquidity ratios provide a snapshot of a company's ability to meet its short-term liabilities, which likely informs investor confidence and risk perception. The fact that such financial health metrics are more critical for some companies than others reflects the aviation sector's varying business models and financial structures.

The analysis also reveals a nuanced landscape where the importance of specific features varies across companies. For instance, while the Passenger Load Factor is crucial for predicting stock performance in SpiceJet Ltd, it holds significantly less weight for El Al Israel Airlines Ltd. This variation suggests that certain companies may depend more heavily on operational metrics, such as how efficiently they manage seat occupancy. In contrast, others may rely more on financial health or asset management indicators. This company-specific variation in feature importance suggests that different strategic priorities drive stock price performance across the aviation industry.

Overall, the analysis indicates that operational efficiency—mainly revenue per ASM and Passenger Load Factor—plays a critical role in stock price prediction for many companies within the aviation sector. This finding suggests that investors highly value airlines' ability to maximize asset utilization and revenue generation. At the same time, financial health metrics, particularly liquidity ratios, also hold substantial importance for specific companies, signaling that the ability to meet short-term obligations is a crucial factor influencing investor sentiment. The heterogeneity in feature importance across companies emphasizes the need for tailored strategies when analyzing stock performance, as some firms may prioritize operational efficiency. In contrast, others focus on maintaining financial stability. This complexity underscores the diverse drivers of stock price performance in the aviation sector, reflecting a blend of operational and financial considerations.

Discussion

This study aimed to explore how operational and financial efficiency metrics impact stock prices in the aviation industry. The results from the machine learning model indicate that operational efficiency, particularly Total Revenue per Available Seat Mile (ASM) and Passenger Load Factor, significantly influences stock prices. Financial metrics, such as the Quick Ratio and Debt-to-Assets Ratio, also play a role, though their influence is secondary to operational factors. The prominence of operational efficiency in predicting stock prices suggests that airlines' ability to maximize revenue generation and asset utilization is a crucial determinant of their market performance. This finding aligns with the research question, confirming that operational and financial health are essential, but operational efficiency provides a stronger, more immediate signal to the market.

The choice of the CatBoost model in this study provided a significant advantage in handling categorical variables and managing missing data without extensive preprocessing. CatBoost's inherent ability to minimize prediction shifts caused by target leakage ensured a robust training process, mainly when dealing with heterogeneous and multivariate datasets typical of the aviation industry. This feature made it well-suited for analyzing stock price determinants, where operational and financial metrics often vary in structure and scale. Moreover, the model's hyperparameter tuning—focusing on depth, L2 leaf regularization, and learning rate—improved prediction accuracy, reducing MSE and MAPE in the optimized model compared to the baseline. The tuning process demonstrated that a balanced combination of tree depth (8), regularization (5), and a learning rate (0.02) delivered the most accurate predictions with lower prediction errors, enhancing the model's reliability in identifying the critical drivers of stock price performance.

The interpretability of the CatBoost model, augmented by SHAP values, further validated its relevance in this context. SHAP analysis revealed that key operational metrics such as Total Revenue per ASM and Passenger Load Factor influenced predictions most. In contrast, financial metrics like Quick Ratio and Debt-to-Assets Ratio played complementary roles. This aligns with previous literature emphasizing operational efficiency as a major driver of stock prices but adds an essential layer of transparency to the model's predictions. The visualization of SHAP values clarified individual feature contributions and facilitated company-specific insights, demonstrating the flexibility and utility of the CatBoost model in complex, real-world scenarios.

The study also relates to the Efficient Market Hypothesis (EMH), particularly its semi-strong form, which posits that stock prices reflect all publicly available information. The findings suggest that operational and financial metrics are indeed integrated into stock prices, providing support for the semi-strong form of EMH. However, the study's ability to identify specific metrics, such as Total Revenue per ASM and Passenger Load Factor, as dominant predictors raises questions about whether all publicly available information is equally weighted or efficiently processed by the market. While the results do not explicitly challenge EMH, they suggest that machine learning models can extract nuanced patterns from operational and financial data, potentially uncovering inefficiencies or underutilized information. This perspective complements existing discussions in finance about the degree of market efficiency and highlights the need for further research to determine whether these insights represent true inefficiencies or are within the bounds of market behavior under EMH.

The findings of this study align with prior research emphasizing the critical role of operational efficiency in influencing stock price performance. Studies by Alici and Sevil (2022) and Labantová and Begera (2014) highlight a positive correlation between higher operational metrics, such as Revenue Passenger Kilometers and Load Factor, and stock prices. This study extends these findings using machine learning to demonstrate that operational metrics, mainly revenue per ASM, remain dominant predictors in a more complex, data-driven framework. However, the results diverge from those of (Yadav & Goriet, 2022; Yilmaz & Köse, 2023), who emphasized external factors like macroeconomic conditions and regulatory influences as primary drivers of stock prices. This study emphasizes company-specific operational and financial performance, suggesting that internal metrics may play a more significant role in determining stock valuation in the aviation sector, although macroeconomic factors remain relevant.

The study's findings carry important implications for both investors and airline management. For investors, operational efficiency metrics, such as Total Revenue per ASM and Passenger Load Factor, provide valuable insights into stock performance and should be prioritized when assessing aviation stocks. The significance of these metrics indicates that airlines that optimize revenue generation and fleet utilization are better positioned for favorable stock price performance, making these critical areas for investment analysis. The results suggest that maintaining operational efficiency is crucial for ensuring positive market responses for airline management. The importance of liquidity, reflected by metrics such as the Quick Ratio, underscores the need for airlines to manage short-term obligations effectively to maintain investor confidence, especially during volatile economic periods. These findings highlight the significance of focusing on internal efficiency to bolster market value and financial stability.

This study contributes to the existing literature by introducing a machine learning framework that aims to improve the accuracy of stock price prediction while offering more interpretable insights through SHAP values. Previous studies, such as those by Amankwah-Amoah (2018), & Singh et al., (2019), explored operational and financial efficiency but often relied on traditional statistical methods, which limited their predictive power and interpretability. This study addresses this gap by showing that machine learning models, combined with SHAP analysis, can offer a more nuanced understanding of how specific operational and financial metrics influence stock prices. The research advances the literature by providing a transparent, data-driven approach that investors and management can use to assess the performance of airlines, addressing a critical need for interpretability in machine learning-based financial models.

While this study provides valuable insights, it has limitations. First, the dataset covers only 65 aviation companies from 2015 to 2023, which may limit the generalizability of the findings to airlines operating in different economic environments or geographical regions. Expanding the dataset to include a broader range of airlines or extending the time frame could enhance the robustness of the results. Second, the model primarily focuses on internal operational and financial metrics, leaving external factors such as fuel prices, geopolitical risks, and regulatory changes that must be explored. Future research could incorporate these external variables to develop a more comprehensive model that captures both internal and external influences on stock prices. This study does not examine potential interactions between operational and financial metrics. Investigating these interactions in future research could provide deeper insights into how operational performance and financial health jointly influence stock prices.

Lastly, while the machine learning model offers solid predictive capabilities, it inherently relies on historical data. As a result, its ability to anticipate future market disruptions, such as sudden economic shocks or industry-wide regulatory changes, may be limited. Future studies should explore incorporating real-time data or developing models that can adapt to changing market conditions to enhance predictive accuracy in the aviation sector further.

Conclusion

This study aimed to examine the influence of operational and financial efficiency on aviation stock prices using a machine learning model—specifically CatBoost enhanced by SHAP interpretability. The findings suggest that operational metrics such as Total Revenue per ASM and Passenger Load Factor may be among the most influential factors in predicting stock prices, potentially surpassing financial metrics like the Quick Ratio and Debt-to-Assets Ratio. These results indicate that, in this study, aviation investors might prioritize operational efficiency over short-term financial performance when evaluating stock value. Airlines with higher revenue generation per available seat mile and optimized fleet utilization tend to see favorable stock market performance. At the same time, liquidity and debt management are also considered necessary, albeit secondary, factors. The findings from this study highlight the potential significance of operational efficiency and financial health for airline management aiming to maintain or enhance stock price performance.

The research addressed the question: How do operational and financial efficiency metrics impact airlines' stock prices? The primary objective was to develop a predictive model that could accurately forecast stock prices based on a combination of these metrics, offering insights into the most influential factors. Additionally, the study sought to interpret the model's predictions using SHAP values to provide transparent explanations of how different metrics contribute to stock price fluctuations. This research achieved these objectives by showing that operational efficiency plays a significant role in stock price performance and that machine learning, coupled with SHAP, can provide meaningful, interpretable insights into stock price prediction.

This study makes several contributions to finance and operational efficiency, particularly in the aviation sector. First, it introduces a machine learning model that aims to improve the accuracy of stock price prediction and offers transparency through SHAP values, addressing a shared concern about the opacity of machine learning models. Second, the study adds to the growing body of literature highlighting the potential importance of operational efficiency in stock market performance, providing empirical evidence that operational metrics may be more influential than financial ones in driving stock prices in the aviation industry. Third, by focusing on operational and financial metrics, this study bridges a gap in the literature, offering a more comprehensive understanding of how these two domains influence stock prices. Finally, the study offers practical insights for investors and airline management by highlighting specific operational and financial metrics that could be important for improving stock performance.

While this study provides valuable insights, several avenues for future research could enhance the understanding of stock price determinants in the aviation industry. Future studies could expand the dataset to include a broader range of airlines and extend the analysis to different economic conditions, regions, or business models, such as low-cost versus full-service carriers. This would provide a more comprehensive understanding of how these factors influence stock prices across different market contexts. Additionally, incorporating external macroeconomic factors such as oil prices, interest rates, and geopolitical risks could provide a more holistic view of the forces driving stock price movements. Future studies could also explore the interactions between operational and financial metrics, examining how these variables jointly influence stock prices. Lastly, incorporating real-time data and developing models that adapt to sudden economic shifts could enhance predictive accuracy and relevance in volatile market conditions.

Katkı Oranları ve Çıkar Çatışması / Contribution Rates and Conflicts of Interest

Etik Beyan	Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.	**Ethical Statement**	It is declared that scientific and ethical principles have been followed while carrying out and writing this study and that all the sources used have been properly cited
Yazar Katkıları	Çalışmanın Tasarlanması: AA (%100) Veri Toplanması: AA (%100) Veri Analizi: AA (%100) Makalenin Yazımı: AA (%100) Makale Gönderimi ve Revizyonu: AA (%100)	**Author Contributions**	Research Design: AA (%100) Data Collection: AA (%100) Data Analysis: AA (%100) Writing the Article: AA (%100) Article Submission and Revision: AA (%100)
Etik Bildirim	iibfdergi@cumhuriyet.edu.tr	**Complaints**	iibfdergi@cumhuriyet.edu.tr
Çıkar Çatışması	Çıkar çatışması beyan edilmemiştir.	**Conflicts of Interest**	The author(s) has no conflict of interest to declare.
Finansman	Bu araştırmayı desteklemek için dış fon kullanılmamıştır.	**Grant Support**	The author(s) acknowledge that they received no external funding in support of this research.
Telif Hakkı & Lisans	Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmaları **CC BY-NC 4.0** lisansı altında yayımlanmaktadır.	**Copyright & License**	Authors publishing with the journal retain the copyright to their work licensed under the **CC BY-NC 4.0**
Etik Kurul	Etik kurul iznine ihtiyaç bulunmamaktadır	**Ethics Committee**	Ethics committee approval is not required.

References

Aldabbas, M., Arabyat, Y. A., Al-Shawawreh, T. B., Alfalayeh, G. A., & Alqudah, M. Z. (2023). The role of information technology in raising the efficiency of Amman Stock Exchange mediated by the behavior of the stock prices. *WSEAS Transactions on Business and Economics, 20,* 1129–1143. https://doi.org/10.37394/23207.2023. 20.101

Alici, A., Alici, A., & Sevil, G. (2022). Analysis of sector-specific operational performance metrics affecting stock prices of traditional airlines. *Independent Journal of Management & Production, 13*(2), 488–506. https://doi.org/10.14807 /ijmp.v13i2.1777

Amankwah-Amoah, J. (2018). Why are so many African companies uncompetitive on the global stage? Insights from the global airline industry. *Africa's Competitiveness in the Global Economy,* 195–216. https://doi.org/ 10.1007/978-3-319-67014-0_8

Arp, L., Baratchi, M., & Hoos, H. H. (2022). VPint: Value propagation-based spatial interpolation. *Data Mining and Knowledge Discovery, 36*(5), 1647–1678. https://doi.org /10.1007/s10618-022-00843-2

Astrup, R., Coates, K. D., & Hall, E. (2008). Finding the appropriate level of complexity for a simulation model: An example with a forest growth model. *Forest Ecology and Management, 256*(10), 1659–1665. https://doi.org/10. 1016/j.foreco.2008.07.016

Babu, B. S., Venkat, K. M., Reddy, J. M., & Sai, K. R. (2022). Airline ticket price prediction. *International Journal for Research in Applied Science and Engineering Technology.*

Bhargav, B., & Prabu, R. T. (2023). Airline passenger satisfaction prediction using novel hybrid random forest model comparison with K-nearest neighbour model. *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM),* 1–6.

Bissessur, A., & Alamdari, F. (1998). Factors affecting the operational success of strategic airline alliances. *Transportation, 25*(4), 331–355. https://doi. org/ 10. 1023/A:1005081621754

Briec, W., & Kerstens, K. (2004). A Luenberger-Hicks-Moorsteen productivity indicator: Its relation to the Hicks-Moorsteen productivity index and the Luenberger productivity indicator. *Economic Theory, 23*(4), 925–939. http://www.jstor.org/stable/25055794

Choudhary, A., Jagadeesh, R., Girija, E., Madhuri, M., & Shravani, N. (2023). Flyhigh: Machine learning-based airline fare prediction model. *2023 6th International Conference on Information Systems and Computer Networks (ISCON),* 1–8.

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean absolute percentage error for regression models. *Neurocomputing, 192,* 38–48. https://doi.org/ 10. 1016/j.neucom.2015.12.114

Dewikristi Siallagan, S., & Prijadi, R. (2020). The impact of operational and financial hedging to the airline operating performance. *KnE Social Sciences.* https:// doi.org/10.18502/kss.v4i6.6635

Dorogush, A. V., Ershov, V., & Gulin, A. (2017). CatBoost: Gradient boosting with categorical features support. *Workshop on ML Systems at NIPS 2017.* Retrieved from https://github.com/

Evans, A., & Schäfer, A. W. (2014). Simulating airline operational responses to airport capacity constraints. *Transport Policy, 34,* 5–13. https://doi. org/10.1016/j.tranpol.2014.02.013

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417.

Gimerská, V., Šoltés, M., & Mirdala, R. (2023). Improving operational efficiency through Quality 4.0 tool: Blockchain implementation and subsequent market reaction. *Quality Innovation Prosperity, 27*(2), 16–32. https://doi.org/ 10.12776/qip.v27i2.1877

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. Journal of big data, 7(1), 94.

Hong, A. C. Y., Khaw, K. W., Chew, X., & Yeong, W. C. (2023). Prediction of US airline passenger satisfaction using machine learning algorithms. *Data Analytics and Applied Mathematics (DAAM).*

Huang, G. (2021). Missing data filling method based on linear interpolation and LightGBM. *Journal of Physics Conference Series, 1754*(1), 012187. https://doi.org /10.1088/1742-6596/1754/1/012187

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems, 30.

Konar, J., Khandelwal, P., & Tripathi, R. (2020). Comparison of various learning rate scheduling techniques on convolutional neural network. *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS 2020).* https://doi.org/10.1109/SCEECS48394.2020.94

Köse, Y., & Yılmaz, E. (2023). An empirical research on the factors affecting profitability in air transportation. *Sosyoekonomi, 31*(58), 43–60. https://doi.org/10.17233/sosyoekonomi. 2023.04.02

Kumar, D. S., Prem, S., Ashfaq, S. M., Rafiq, P. M., & Varma, C. N. (2023). Airline fare prediction using machine learning.

Labantová, L., & Begera, V. (2014). Financial, operational and economic indicators of airline operators. *Faculty of Aeronautics, Technical University of Košice, 16*(2).

Lee, B. L., Wilson, C., Pasurka, C. A., Fujii, H., & Managi, S. (2017). Sources of airline productivity from carbon emissions: An analysis of operational performance under good and bad outputs. *Journal of Productivity Analysis, 47*(3), 223–246. https://doi.org/10.1007/s11123-016-0480-4

Lee, J. (2019). Effects of operational performance on financial performance. *Management Science Letters, 9*(1), 25–32.

Lee, K. (2023). Airline operational disruptions and loss-reduction investment. *Transportation Research Part B: Methodological, 177,* 102817. https://doi.org/10.1016/j. trb.2023.102817

Lee, S., & Park, S.-B. (2013). A study on the association between operating leverage and risk: The case of the airline industry. *International Journal of Economics and Finance, 6*(3), 120. https://doi.org/10.5539/ijef.v6n3p120

Lee, S., Kim, H., & Lee, N. (2019). A comparative analysis of financial and operational performance pre-and post-IPO: With a focus on airline companies. *Academy of Accounting and Financial Studies Journal, 23*(3), 1–14.

Lopes, I. F., & Beuren, I. M. (2017). Comportamento dos custos e sua relação com medidas de eficiência operacional em companhias aéreas. *BASE - Revista de Administração e Contabilidade Da Unisinos, 14*(1), 30–46. https://doi.org/10.4013/base.2017.141.03

Lopes, I. T., Ferraz, D. P., & Rodrigues, A. M. G. (2016). The drivers of profitability in the top 30 major airlines worldwide. *Measuring Business Excellence, 20*(2), 26–37. https://doi.org/10.1108/mbe-09-2015-0045

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems, 30* (pp. 4765–4774). Curran Associates, Inc.

Lundberg, S., Erion, G., & Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. *ArXiv Preprint.* https://doi.org/10.48550/arxiv.1802.03888

Mantin, B., & Wang, J.-H. (2012). Determinants of profitability and recovery from system-wide shocks: The case of the airline industry. *Journal of Airline and Airport Management, 2*(1), 1–33. https://doi.org/10.3926/jairm.2

McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference,* 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

McLean, D. (2005). Operational efficiency of commercial transport aircraft. *Measurement and Control, 38*(8), 243–248. https://doi.org/10.1177/002029400503800803

Nagesh, P., Naidu, K. B. J., Kowshik, P., & Sekhar, P. C. (2023). Airline ticket price prediction model. *International Journal for Research in Applied Science and Engineering Technology.*

Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2021). Explanation of machine learning models using Shapley additive explanation and application for real data in hospitals. *ArXiv Preprint.* https://doi.org/10.48550/arxiv.2112.11071

Okur, M., & Gurbuz, A. O. (2015). Behavioral finance in theory and practice. In *Banking, Finance, and Accounting: Concepts, Methodologies, Tools, and Applications* (pp. 18). IGI Global. https://doi.org/10.4018/978-1-4666-6268-1.ch016

Pamungkas, D. P., & Suhadak, S. (2017). The effect of jet fuel price and macroeconomics variables on profitability of airline industry in Asia (Study at airline companies in Indonesia, India, and China period 2006–2015). *Jurnal Ekonomi Bisnis dan Kewirausahaan, 50*(3), 208–217.

Panigrahi, R., Patne, N. R., Pemmada, S., & Manchalwar, A. D. (2022). Regression model-based hourly aggregated electricity demand prediction. *Energy Reports, 8,* 16–24. https://doi.org/10.1016/j.egyr.2022.10.004

Piranti, M. (2021). The impact of fuel price fluctuation and macroeconomic variables to airlines performance. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(3), 5486–5494.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. Advances in neural information processing systems, 31.

Putra, R. A. K., Persada, S. F., Kumalasari, R. D., Herdina, A. M., Sekardhani, M., & Razif, M. (2024). Operational cost efficiency and profitability effects on companies' distribution stock prices during COVID-19. *Jurnal Manajemen Teknologi, 23*(1), 76–90. https://doi.org/10.12695/jmt.2024.23.1.5

Rossi, M. (2015). The efficient market hypothesis and calendar anomalies: A literature review. *International Journal of Managerial and Financial Accounting, 7*(3/4), 285. https://doi.org/10.1504/IJMFA.2015.074905

Sachdeva, T. (2020). Managing shareholders in turbulent times: Evidence from Indian stock market. *International Journal of Business and Globalisation, 25*(3), 265. https://doi.org /10.1504/IJBG.2020.109005

Şahinbaş, K. (2022). Price prediction model for restaurants in Istanbul by using machine learning algorithms. *Ekonomi İşletme ve Maliye Araştırmaları Dergisi, 4*(2), 159–171.

Saranga, H., & Nagpal, R. (2016). Drivers of operational efficiency and its impact on market performance in the Indian airline industry. *Journal of Air Transport Management, 53,* 165–176. https://doi.org/10.1016/j.jairtraman.2016.03.001

Satter, S., Kwon, T.-H., & Kim, K.-D. (2023). A comparative analysis of various machine learning algorithms to improve the accuracy of HbA1c estimation using wrist PPG data. *Sensors, 23*(16), 7231. https://doi.org/10.3390/ s23167231

Seufert, J. H., Arjomandi, A., & Dakpo, K. H. (2017). Evaluating airline operational performance: A Luenberger-Hicks-Moorsteen productivity indicator. *Transportation Research Part E: Logistics and Transportation Review, 104,* 52–68. https://doi.org/10.1016/j.tre.2017.05.006

Singh, J., Sharma, S. K., & Srivastava, R. (2019). What drives Indian airlines' operational expense: An econometric model. *Journal of Air Transport Management, 77,* 32–38. https://doi.org/10.1016/j.jairtraman.2019.03.003

Slater, P. J. (2023). Mind the GAAP: Determinants of U.S. airline operating cash flows. *Journal of Academy of Business and Economics, 23*(1), 74–84. https://iabe.org/IABE-DOI/article.aspx?DOI=JABE-23-1.6

Subramanian, R., Murali, M. S., Deepak, B., Deepak, P., Reddy, H. N., & Sudharsan, R. R. (2022). Airline fare prediction using machine learning algorithms. *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT),* 877–884.

Subramanyam, K. R., & Wild, J. J. (2016). *Financial statement analysis* (11th ed.). McGraw-Hill Education.

Sumerli Sarıgül, S., Ünlü, M., & Yaşar, E. (2023). Financial performance analysis of airlines operating in Europe: CRITIC-based MAUT and MARCOS methods. *International Journal of Business and Economic Studies, 5*(2), 76–97. https://doi.org/10.54821/uiecd.1257488

Sun, Y., Wong, A. K., & Kamel, M. S. (2009). Classification of imbalanced data: A review. International journal of pattern recognition and artificial intelligence, 23(04), 687-719.

Tanrıverdi, G., Merkert, R., Karamaşa, Ç., & Asker, V. (2023). Using multi-criteria performance measurement models to evaluate the financial, operational and environmental sustainability of airlines. *Journal of Air Transport Management, 112,* 102456. https://doi.org/10.1016/j. jairtraman.2023.102456

Wang, W.-K., Lu, W.-M., & Tsai, C.-J. (2011). The relationship between airline performance and corporate governance amongst U.S. listed companies. *Journal of Air Transport Management, 17*(2), 148–152. https://doi.org/10.1016/j.jairtraman.2010.06.005

Woo, K.-Y., Mai, C., McAleer, M., & Wong, W.-K. (2020). Review on efficiency and anomalies in stock markets. *Economies, 8*(1), 20. https://doi.org/10.3390/economies8010020

Wulansari, T., Suwardi, F., Firdaus, F., Purwanto, D., Hibatullah, Z., & Marganis, V. C. (2024). The influence of liquidity, productive asset quality and operational efficiency on company stock prices: Study of banks going public for the 2016–2020 period. *International Journal of Economics and Management Sciences, 1*(2), 19–28. https://doi.org/ 10.61132/ijems.v1i2.18

Xu, Y., & E, X. (2020). ARIMA based on linear interpolation for SO2 monitoring data's calibration. *Journal of Physics Conference Series, 1616*(1), 012044. https://doi.org/10. 1088/1742-6596/1616/1/012044

Yadav, D. K., & Goriet, M. O. (2022). An illustrative evaluation of external factors that affect performance of an airline. *Journal of Aerospace Technology and Management, 14,* 1122. https://doi.org/10.1590/jatm.v14.1253

Yilmaz, E., & Köse, Y. (2023). An empirical research on the factors affecting profitability in air transportation. *Sosyoekonomi, 31*(58), 43–60. https://doi.org/10.17233/sosyoekonomi. 2023.04.02

Yongxin, L. (2009). Discussing on trend to efficient market hypothesis of securities and futures market. *2009 International Conference on Information Management, Innovation Management and Industrial Engineering, 3,* 149–152.

Yüksel, B. (2023). Zaman serilerinde talep tahmini (Demand prediction in time series). *YBS Ansiklopedi, 11*(2), 1–18.

Zhang, Q., Koutmos, D., Chen, K., & Zhu, J. (2021). Using operational and stock analytics to measure airline performance: A network DEA approach. *Decision Sciences, 52*(3), 720–748. https://doi.org/10.1111/deci.12363

Zhao, Q., Wang, H., Luo, J.-C., Luo, M., Li, L., Yu, S.-J., Li, K., Zhang, Y., Sun, P., Tu, G.-W., & Luo, Z. (2021). Development and validation of a machine-learning model for prediction of extubation failure in intensive care units. *Frontiers in Medicine, 8.* https://doi.org/10.3389/fmed.2021.676343

Zou, L., Yu, C., & Dresner, M. (2015). Fleet standardization and airline performance. *Journal of Transport Economics and Policy, 49*(1), 149–166. https://www.jstor.org/stable/ jtranseconpoli.49.1.0149