A close up of a logo

Description automatically generated

“Optimising Airline Fleet Planning: A Universal Model Using Advanced Analytics”

By

**Seyed Roohollah Mousavi**

19530842

Supervised by

**Prof. Deepika Deepika**

Submitted in partial fulfilment of the requirements of

the Dept. of Computer Science

of the University of Westminster

for award of the Master of Science

JANUARY 2024

DECLARATION

I, *Seyed Roohollah Mousavi* declare that I am the sole author of this Project; that all references cited have been consulted; that I have conducted all work of which this is a record, and that the finished work lies within the prescribed word limits.

This has not previously been accepted as part of any other degree submission.

Signed: Seyed Roohollah Mousavi

Date: 17 January 2024

FORM OF CONSENT

I, *Seyed Roohollah Mousavi* hereby consent that this Project, submitted in partial fulfilment of the requirements for the award of the MSc degree, if successful, may be made available in paper or electronic format for inter-library loan or photocopying (subject to the law of copyright), and that the title and abstract may be made available to outside organisations.

Signed: Seyed Roohollah Mousavi

Date: 17 January 2024

Abstract

This research presents an adaptive model for airline fleet planning, leveraging advanced analytics and domain knowledge. The study addresses the challenge of determining optimal fleet size and composition, considering factors such as acquisition, retrofitting, seating configuration, and route profitability. The research employs a manual compilation of Low-Cost Carriers (LCCs) and Cargo carriers operating in the United States from 1990 to 2023, revealing distinct market demand patterns for LCCs and Full-Service Carriers (FSCs). The study utilizes the Mean Absolute Percentage Error (MAPE) and R-squared for model evaluation, highlighting the effectiveness of Long Short-Term Memory (LSTM) and Autoregressive Integrated Moving Average with exogenous variables (ARIMAX) in different market segments. The research also integrates fleet planning models with the Hub and Spoke (HS) network, introducing managerial decision-making options that impact operational costs and extend aircraft lifespan. The study concludes by emphasizing the need for individual route forecasting and proposes avenues for future research, particularly in demand segmentation by class. The research contributes valuable insights into forecasting methodologies and decision variables shaping optimal fleet composition, paving the way for continued advancements in aviation research.

Acknowledgements

First and foremost, I would like to thank Mona, my wife, whose unwavering support has been a pillar of strength during this challenging period. Her understanding and patience, especially when household chores were left undone due to my commitment to this research, have been invaluable. I appreciate her letting me skip washing-up!

I would also like to extend my appreciation to my academic advisors and colleagues, whose guidance and feedback have been crucial in shaping this research.

Word Count: 13622

This project was done using R, Python, and PowerBI.

The codes for R and Python have been pushed to GitHub and can be accessed through this link:

https://github.com/Hessamous/Dissertation

Table of Contents

[1 Introduction 1](#_Toc156403371)

[1.1 Background and Motivation 1](#_Toc156403372)

[1.2 Financial Challenges Faced by Airlines 2](#_Toc156403373)

[1.3 Airline Planning Process and Fleet Planning Problem 2](#_Toc156403374)

[1.4 Objectives and Aim of this Thesis 3](#_Toc156403375)

[1.5 Variables Affecting Optimal Fleet Composition 4](#_Toc156403376)

[1.6 Methodology Overview 4](#_Toc156403377)

[1.7 Chapters Overview 5](#_Toc156403378)

[2 Literature review 5](#_Toc156403379)

[2.1 Pioneering Researchers in Airline Fleet Planning 5](#_Toc156403380)

[2.2 Fleet Capacity Management and Leasing 6](#_Toc156403381)

[2.3 Environmental objectives in fleet planning 6](#_Toc156403382)

[2.4 Fleet Assignment Models 7](#_Toc156403383)

[2.5 Decision Theory 7](#_Toc156403384)

[2.6 Recent Works 8](#_Toc156403385)

[2.7 Conclusion 8](#_Toc156403386)

[3 Research Question and Aim 9](#_Toc156403387)

[3.1 Research Question: 9](#_Toc156403388)

[3.1.1 Sub-question A: 9](#_Toc156403389)

[3.1.2 Sub-question B: 9](#_Toc156403390)

[3.1.3 Sub-question C: 9](#_Toc156403391)

[3.2 Advanced Demand Forecasting 9](#_Toc156403392)

[3.3 Comprehensive Business Optimization for Fleet Planning 10](#_Toc156403393)

[4 Data and EDA 10](#_Toc156403394)

[4.1 Introduction 10](#_Toc156403395)

[4.2 Data Collection 11](#_Toc156403396)

[4.2.1 T-100 11](#_Toc156403397)

[4.2.2 Feature-specific Data 11](#_Toc156403398)

[4.2.3 Fuel Price Data 12](#_Toc156403399)

[4.3 Pre-processing and Feature Engineering 13](#_Toc156403400)

[4.3.1 Preliminary Cleaning 13](#_Toc156403401)

[4.3.2 Added Features 13](#_Toc156403402)

[4.3.3 Aggregation for EDA 15](#_Toc156403403)

[4.3.4 Aggregation for Forecast 15](#_Toc156403404)

[4.4 Exploratory Data Analysis 16](#_Toc156403405)

[4.4.1 Overview 16](#_Toc156403406)

[4.4.2 The Pandemic Effect 19](#_Toc156403407)

[4.4.3 Conclusion 22](#_Toc156403408)

[5 Forecast methodology and analysis 23](#_Toc156403409)

[5.1 ARIMA 23](#_Toc156403410)

[5.2 LSTM 24](#_Toc156403411)

[5.3 Exogenous Variables 25](#_Toc156403412)

[5.4 Fitness Metrics 26](#_Toc156403413)

[5.4.1 Common Fitness Metrics 26](#_Toc156403414)

[5.4.2 Rationale for Choosing MAPE 27](#_Toc156403415)

[5.4.3 R-Squared 27](#_Toc156403416)

[5.5 Results 27](#_Toc156403417)

[5.5.1 ARIMA 27](#_Toc156403418)

[5.5.2 LSTM 31](#_Toc156403419)

[5.6 Comparison and Conclusion 38](#_Toc156403420)

[5.6.1 Results on Aggregated Routes Data 38](#_Toc156403421)

[5.6.2 Best Predictors for Each Focus Category 39](#_Toc156403422)

[5.6.3 Key Insights and Future Directions 39](#_Toc156403423)

[6 Business optimisation model description and analysis 40](#_Toc156403424)

[6.1 Underlying Concepts 40](#_Toc156403425)

[6.1.1 Hub and Spoke Network 40](#_Toc156403426)

[6.1.2 Variable Routing Cost and DOC 41](#_Toc156403427)

[6.1.3 Specific Range 41](#_Toc156403428)

[6.2 Our Contributions 42](#_Toc156403429)

[6.2.1 Retrofit 42](#_Toc156403430)

[6.2.2 Reconfiguration 43](#_Toc156403431)

[6.3 Model 43](#_Toc156403432)

[6.3.1 Nomenclature 44](#_Toc156403433)

[6.3.2 Mathematical Formulation 45](#_Toc156403434)

[6.3.3 Model Description 46](#_Toc156403435)

[7 Conclusion 47](#_Toc156403436)

[8 Bibliography 48](#_Toc156403437)

[9 Appendices 53](#_Toc156403438)

# Introduction

## Background and Motivation

The aviation industry, characterized by its volatile nature, orchestrates a complex interplay of elements wherein a single misjudgement can disrupt the entire operational symphony. Within this intricate ballet, the strategic management of airline fleets assumes a pivotal role, resonating across dimensions of operational efficiency, financial sustainability, and customer contentment. This master's dissertation plunges into this domain equipped with advanced analytics, aiming to elucidate the intricacies and decision-making processes that delineate the aviation landscape.

In 2023, the global airline industry showed signs of recovery, posting a projected net profit of $23.3 billion after a challenging 2022, according to the International Air Transport Association (IATA) (International Air Transport Association (IATA), 2023). However, when considering recent studies, including one focused on the Asia-Pacific region and historical data from North American airlines, a consistent narrative emerges: the airline industry is characterized by marginal net profits and poor financial performance.

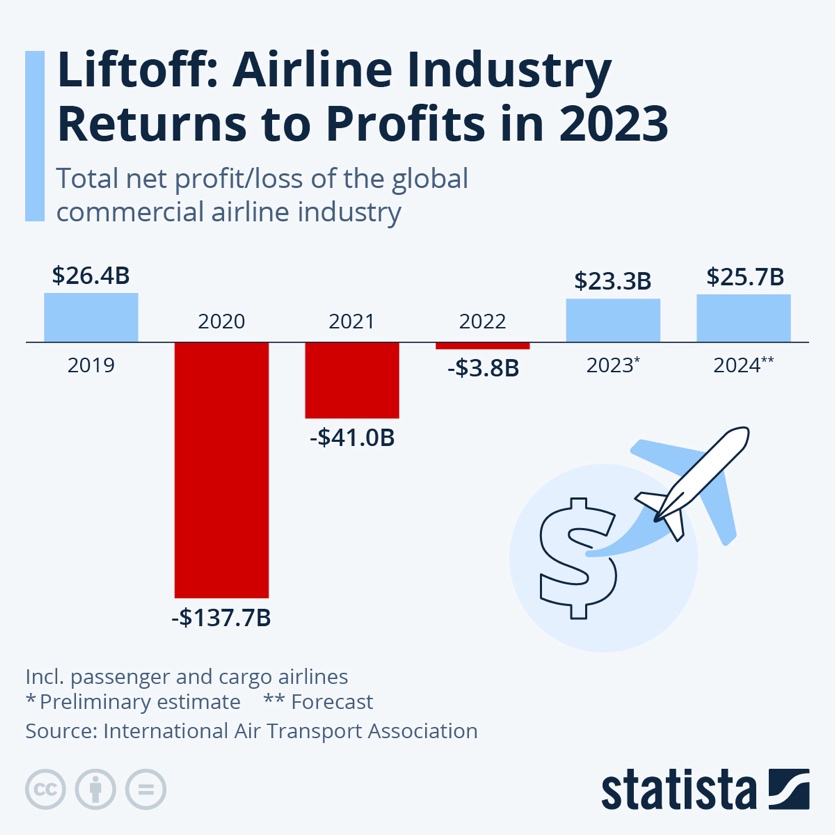


Figure Total net profit/loss of the global commercial airline industry (Richter, 2023)

The Asia-Pacific study during 2016–2019 revealed that despite rapid growth, airlines in the region faced thin and volatile profit margins. Only a few full-service carriers consistently improved financial efficiency, while concerns were raised about the excessive use of financial leverage, prompting the need for effective capital structure management and technological innovation. (Huang, 2021)

Contrasting this with historical data from 2004 to 2011, North American airlines struggled to provide satisfactory returns on invested capital, averaging 4.1% against a cost of capital of 7.5%. In 2012, the average net profit margin was a mere 1.1%. This historical perspective underscores the persistent challenges faced by airlines in delivering robust financial performance. (Sa, Santos, & Clarke, Portfolio-based airline fleet planning under stochastic demand, 2019)

## Financial Challenges Faced by Airlines

Financial considerations form a crucible of challenges for airlines, extending far beyond routine operational costs. Fuel prices, constituting up to 30% of an airline's expenditures (IATA Sustainability and Economics), exhibit notorious volatility, with recent surges significantly impacting airline profitability. Beyond fuel, the industry grapples with global economic shifts, technological disruptions, and geopolitical events, creating an intricate financial environment where long-term commitments like aircraft acquisition, often reaching billions of dollars, intersect with a perpetually changing market. This financial tapestry weaves together several challenges: capital intensity, demanding careful balance between fleet expansion and financial stability; substantial operating costs like fuel, maintenance, crew salaries, and leasing fees, requiring effective fleet planning for cost minimization and revenue maximization; and market volatility, necessitating robust strategies to navigate fluctuating demand due to economic cycles, geopolitical events, and unforeseen crises. This thesis proposes a novel fleet optimization model that leverages advanced analytics to tackle these challenges and guide airlines towards improved financial performance in this dynamic and volatile environment.

## Airline Planning Process and Fleet Planning Problem

The landscape of airline planning has undergone a metamorphosis from an art form guided by intuition and market trends to a sophisticated science orchestrated by data-driven insights. This multifaceted challenge demands careful consideration of a complex melody of factors, from the ever-shifting whispers of market demands to the rumbling pronouncements of fuel price predictions. Each note in this symphony, be it the intricate characteristics of individual aircraft or the thundering concerns of environmental sustainability, resonates across the airline's entire existence, influencing its financial health, brand image, operational flexibility, and ultimately, its growth potential.

The orchestration of this intricate symphony unfolds in meticulously interconnected stages. Network planning lays the groundwork, meticulously designing route networks that harmonize passenger demand with competitive realities and profitability imperatives. This blueprint directly influences the fleet requirements, serving as the conductor guiding the composition of the orchestra. Fleet planning, the pivotal phase, revolves around determining the optimal size, composition, and deployment of the fleet. Airlines must strike a delicate balance, harmonizing their fleet decisions with the network's demands and the financial constraints that define the stage. Finally, scheduling, ensures optimal utilization of the aircraft, seamless integration with crew availability, and passenger convenience. It seamlessly translates the fleet planning blueprint into the real-world dynamics of operation, ensuring the smooth and efficient performance of the entire airline ensemble.

Traditionally, fleet planning was perceived as an art guided by intuition and prevailing market trends. However, with the industry's ascent to unprecedented altitudes, it became evident that this facet possessed the potential to either fortify or jeopardize an airline's standing. The selection of appropriate aircraft, in optimal quantities, and directed towards suitable destinations, evolved into a strategic imperative. This decision not only wielded influence over profitability but also moulded operational adaptability and responsiveness to the dynamic currents of the market. This meticulous calculus unfolds amid the backdrop of fluctuating fuel prices, dynamic market forces, and geopolitical considerations, necessitating a comprehensive understanding of intricacies and adept optimization of fleet composition.

## Objectives and Aim of this Thesis

This work is dedicated to the development of a comprehensive model for airline fleet planning, fortified by advanced analytics. Our objectives encompass three key areas:

Firstly, we aim to enhance robustness by creating a model that thrives amidst demand uncertainties, operational disruptions, and external shocks. This involves fortifying the planning framework to withstand various challenges that may arise in the dynamic aviation industry.

Secondly, we seek to leverage analytics by employing data-driven approaches to optimize fleet decisions. Our goal is to align these decisions with operational realities, ensuring that the fleet planning process is not only efficient but also adaptable to the ever-evolving demands of the airline industry.

Lastly, our approach integrates fleet planning with the financial health of airlines, emphasizing the importance of sustainable and profitable operations. By addressing the financial realities, our model aims to contribute to the overall success and longevity of airline operations.

Drawing from previous experience in the air navigation industry and fortified by strong analytical skills developed in academia, the author's personal motivation is rooted in leveraging advanced data manipulation and analysis. This commitment is driven by a desire to bring depth and value to the analysis dimension of airline fleet optimization models, ultimately contributing to the advancement of the field.

## Variables Affecting Optimal Fleet Composition

Achieving optimal fleet composition involves a nuanced understanding of a diverse set of variables that collectively shape fleet planning decisions. Key factors influencing this complex problem include market demand patterns, competition levels, aircraft characteristics, and financial considerations. Pioneering research by (New, 1975) and (Schick & Stroup, 1981) introduced linear programming models to address these considerations. More contemporary scholars, such as (Khoo & Teoh, 2014) and (Müller, Kieckhäfer, & Spengler, 2018), have expanded the scope by incorporating environmental objectives and retrofit decisions into the optimization framework. The depth of research underscores the multifaceted nature of the optimal fleet composition problem.

The fleet planning challenge can be likened to assembling a complex jigsaw puzzle, with airlines grappling with various variables. Key considerations include the selection of aircraft types, which directly impacts capacity, range, and operational flexibility. This work delves into the nuanced characteristics of aircraft in general, considering categorical specifications for short, medium, and long-haul aircraft. Additionally, route characteristics play a crucial role, with fleet decisions needing to account for factors such as route lengths, passenger volumes, and seasonal variations. Robust fleet planning also considers market demand dynamics, including passenger preferences, market growth trajectories, and competitive dynamics. Furthermore, the perpetual challenge of aging fleet management involves balancing maintenance costs, reliability, and replacement strategies for aging aircraft. The intricate interplay of these variables highlights the multifaceted nature of the optimal fleet composition problem.

## Methodology Overview

This thesis adopts a dual-part approach to unravel the intricacies of airline fleet planning. The first section delves into meticulous data collection, feature engineering, and the implementation and comparison of forecasting algorithms such as ARIMA and LSTM models. The second section introduces a comprehensive business model integrating seating configuration, forecasted data, aircraft characteristics, airline financial situations, and retrofit options. Together, these sections contribute a holistic perspective to tackle the multifaceted challenges posed by fleet planning.

The methodology encompassing a thorough data collection process, feature engineering, exploratory data analysis, and the implementation and comparison of various forecasting algorithms serves as the foundation. Advanced analytics techniques, such as ARIMA, ARIMAX with different exogenous variables, and LSTM models, are employed for forecasting, demonstrating a commitment to cutting-edge methodologies. The second section of the thesis involves the development of a business optimization model, incorporating forecasted data and decision variables to guide fleet planning decisions. This dual-methodology approach ensures a robust exploration of the complex landscape of airline fleet planning.

## Chapters Overview

This dissertation dives into the intricate world of airline fleet planning, armed with the arsenal of advanced analytics. Chapter 2 embarks on a historical and contemporary journey through existing literature, laying the groundwork for comprehending the field's multifaceted challenges. From New's pioneering work in 1975 to Müller et al.'s recent ground-breaking studies, the review sets the stage for the innovative methodologies this dissertation proposes. Chapter 3 meticulously maps the research journey, delineating the specific aims and objectives that guide the analysis. It establishes a roadmap for the subsequent chapters, ensuring each step aligns with the overarching research vision. Chapter 4 then peels back the curtain on the data, shedding light on the intricate collection process, its diverse sources, and the challenges encountered. Ensuring transparency and reliability, it reveals the methodologies employed to secure robust and dependable data for analysis. With the foundation laid, Chapter 5 delves deep into the intricacies of forecasting methodologies. It embarks on a comprehensive examination of implementing and comparing various algorithms, from ARIMA to LSTM, positioning the thesis at the forefront of predictive analytics. Chapter 6 bridges the gap between theoretical frameworks and practical implications by introducing the comprehensive business optimization model developed for fleet planning. It meticulously details the model's components and analyses its application in diverse scenarios, grounding the research in real-world applicability. Finally, Chapter 7 synthesizes the key findings and their implications, drawing together the diverse threads explored throughout the thesis. It offers a comprehensive conclusion that serves as the culmination of the research journey, reflecting on the significant contributions made to the field. The remaining chapters meticulously document the references and appendices, anchoring the thesis in existing scholarship and providing a valuable resource for further exploration.

This thesis aims to contribute substantive insights to the field of airline fleet planning, leveraging advanced analytics to optimize decision-making processes in the face of complex financial and operational challenges. Through a comprehensive exploration of historical foundations, meticulous data analysis, and innovative business modelling, the research seeks to provide a universal model for optimizing airline fleet planning.

# Literature review

## **Pioneering Researchers in Airline Fleet Planning**

(New, 1975) and (Shube & Stroup, 1975) are among the pioneering researchers in airline fleet planning. They developed and applied linear programming models aiming to minimize the net present value of cash flows associated with aircraft operation. The primary decision variables in these models are the timing of aircraft investment and disposal, offering a chronological perspective on fleet management. Early deterministic models in the '80s evolved into considering demand uncertainty in multi-year fleet planning (Schick & Stroup, 1981).

(Subramanian, Scheff, Quillinan, Wiper, & Marsten, 1994) further solidified the model's practical value by demonstrating its successful application at Delta Air Lines. Here, it optimized aircraft assignment to new routes while minimizing costs and adhering to operational constraints, showcasing its real-world effectiveness. (Powell & Carvalho, 1997) on the other hand, tackled a core challenge in fleet planning models: the trade-off between realistic, integer-valued solutions (representing whole aircraft) and computationally efficient optimization. While integer solutions are intuitive and practical, they often significantly increase computation time. To address this, they proposed modelling the multi-commodity network flow problem as a dynamic control problem.

## Fleet Capacity Management and Leasing

Leasing aircraft can offer flexibility in managing fleet capacity, especially in an environment characterized by demand fluctuations. Faced with fluctuating demand, airlines increasingly turn to aircraft leasing for its unmatched flexibility in managing fleet capacity. In this way, they can scale up or down quickly without making hefty upfront investments, as evidenced by (Kaplan, 2017)'s report of operating lessors' fleet share skyrocketing from 2% to 40% in just 40 years, already sitting at 15% in the 1990s.

Recognizing the cyclical and stochastic nature of demand, studies like (Oum, Zhang, & Zhang, 2000) propose a cost trade-off formulation to optimize the lease/own mix for airlines. This approach pinpoints the ideal proportion of leased aircraft, maximizing efficiency while maintaining operational flexibility. (Listes & Dekker, 2005) propose a new method for robust airline fleet planning called "scenario aggregation" that explicitly considers passenger demand uncertainty. This is crucial for dynamic capacity allocation systems where airlines need to adapt their fleet to constantly changing market conditions. Holloway's contribution in 2008 introduced the concept of Demand-Driven Fleet Management (DFM) (Holloway, 2008). It emphasizes the flexibility that airlines must have to switch aircraft based on capacity requirements on or close to the day of operation, reflecting the dynamic nature of airline operations. (Bazargan & Hartman, 2012) developed an extended fleet planning model, integrating aircraft leasing as a decision variable. On a related note, (Hsu, Li,, Liu, & Chao, 2011) explored the ramifications of demand volatility on the proportion of leased aircraft in an airline's fleet.

## Environmental objectives in fleet planning

With rising environmental concerns and the airline industry's significant carbon footprint, researchers started to incorporate environmental objectives into fleet planning. (Khoo & Teoh, 2014) and (Rosskopf, Lehner, & Gollnick, 2014) are notable for presenting models that give weight to both economic and environmental objectives, marking a shift towards sustainability in airline fleet management.

(Müller, Kieckhäfer, & Spengler, 2018) ventured into the intricate relationship between emission thresholds, retrofit options, and their subsequent pricing and savings. This work provides invaluable insights into retrofit decisions and their influence on fleet composition. Alongside fleet planning, associated decisions like fleet assignment and crew pairing have also been tackled.

## Fleet Assignment Models

The fleet assignment model (FAM) plays a critical role in optimizing airline operations by bridging the gap between available aircraft (supply) and outbound passengers (demand). Given a fixed fleet composition, FAM performs the intricate task of assigning specific aircraft types to scheduled flights, ensuring optimal resource utilization. Its core objectives can be twofold: maximizing operating profit through strategic revenue generation or minimizing operating costs for enhanced efficiency. Ultimately, FAM empowers airlines to make data-driven decisions that maximize their financial performance.

Researchers have brought to the fore enhanced revenue models in the context of fleet assignments. Cynthia Barnhart's work in 2009 titled "Airline Fleet Assignment with Enhanced Revenue Modeling" is a testament to this effort (Berhart, Farahat, & Lohatepanont, 2009). Moreover, (Clarke, Hane, Johnson, & Nemhauser, 1996) widened the scope of fleet assignments by incorporating maintenance and crew considerations, adding another layer of complexity to fleet planning models. (Özener, Matoglu, Günes, Haouari, & Sözer, 2017) devised a comprehensive model that integrates fleet assignment with crew pairing, addressing both aspects in a unified framework.

A robust fleet planning framework is introduced by (Sa, Robust fleet planning under stochastic demand (Master's Thesis), 2016) that proposes a multi-year optimization model that combines demand forecasting with fleet assignment, generating a portfolio of potential fleets with varying sizes and compositions. This research innovates by explicitly considering the uncertainty in future air travel demand, capturing its impact on fleet performance. Same author has Proposed three-step methodology for identifying robust fleets based on profit performance under various demand scenarios (Sa, Santos, & Clarke, Portfolio-based airline fleet planning under stochastic demand, 2019).

## Decision Theory

Multiple-criteria decision-making (MCDM) methods have been used extensively throughout the literature since 2011. A cursory review of the extant literature reveals some noteworthy trends and points of contention. For instance, (Ozdemir, Basligil, & Karaca, 2011)utilized the Analytic Network Process in their comparison of aircrafts like A319, A320, and B737, considering factors like maintenance cost, reliability, and delivery time. This study, like others, centres its focus on cost, time, and physical attributes. Similarly, (Gomes, Fernandes, & Soares de Mello, 2012) took a novel approach with NAIADE Method for assessing aircraft types such as Cessna 208 and Beechcraft 1900, underlining financial, logistic, and quality factors. Notably, many of the studies, ranging from those of (Teoh & Khoo, 2015) to (Ardil, 2020), employ the Analytic Hierarchy Process among other methodologies to weigh the merits and demerits of various aircraft types against a set of established criteria.

## Recent Works

In recent years, Oliveira's works in 2021 and 2022 shed light on the link between energy intensity reduction and fleet modernization. He found that spikes in fuel prices can expedite the fleet rollover and modernization by approximately 3–4 years (Oliveira, Narcizo, Caliari, Morales, & Prado, 2021). Moreover, his 2022 study indicates that increasing energy costs might encourage greater fleet modernization in the long run, with airlines potentially aiming for more eco-efficient operations up to two years post a surge in fuel prices (Oliveira, Caliari, & Narcizo, An empirical model of fleet modernization: On the relationship between market concentration and innovation adoption by airlines, 2022).

Multiple contributions in recent years have revolved around combining network design problems and fleet planning problems arguing that solving them simultaneously is the key to a successful airline strategy. Including (Wu, Zhang, Wang, & Shi, 2022) where the objective of their model is to minimize the total system cost, including both hub establishment and aircraft-related expenses. To overcome the complexities of the integrated model, the researchers develop a heuristic solution algorithm based on a Genetic Algorithm framework. This allows for efficient optimization even with large and complex datasets. (Mohri, Nasrollahi, Pirayesh, & Mohammadi, 2022) proposed a similar model integrating a Hub Location Problem (HLP) with airline fleet planning. Their work focuses specifically on optimizing global hub networks for international flights and recognizes the crucial role of fleet size and diversity in hub network design. Their HLP models are tested on several network instances, including real-world international flight data, demonstrating the effectiveness of integrated fleet planning models in generating practical and efficient hub network configurations.

## Conclusion

A critical observation across these studies suggests some limitations. Primarily, the aircraft types processed are frequently restricted to a narrow range, often just a handful of models. This potentially limits the breadth of their findings. Furthermore, the historical data upon which these evaluations are based appears rather simplistic, not delving into intricate factors that might affect aircraft performance or long-term viability. Additionally, route considerations within these models are rather general, lacking specificity to routes or flight patterns.

There emerges a clear opportunity to utilize more advanced analytics, machine learning, and data science principles to bridge these gaps. Such an approach promises a more comprehensive evaluation of aircraft, accounting for intricate historical data, specific route considerations, and other nuanced factors that might influence aircraft selection. In conclusion, while the present body of literature offers valuable insights, there remains ample space for richer, more detailed investigations that tap into the potential of contemporary analytical tools.

# Research Question and Aim

Leveraging the author’s extensive air navigation industry experience and academic foundation in advanced analytics, he is motivated by the challenge of manipulating and analysing complex data sets. This commitment stems from a deep desire to enhance the analytical capabilities of airline fleet optimization models, ultimately contributing to the refinement and advancement of fleet planning methodologies within this domain.

## Research Question:

At the core of this research lies the challenge of constructing an adaptive model for airlines that brings together the various components of fleet management: acquisition, retrofitting, seating configuration and route profitability, under the lens of advanced analytics. From this central theme, the research seeks to address the following question:

*“How can advanced analytics coupled with domain knowledge aid in creating a robust model to guide airlines in determining the optimal fleet size and composition?”*

To unravel the complexities of this overarching question, we outline the following sub-questions:

### Sub-question A:

*What variables affect the optimal fleet composition for an airline?*

### Sub-question B:

*Which decisional variables can offer the best pragmatic utilizations to the operator?*

### Sub-question C:

*Which of the variables affecting the optimal fleet composition are to be forecasted, and what is the best approach to its inherent uncertainty?*

Dedicated to the meticulous development of an advanced and all-encompassing model to address these questions, this work is predicated on two distinct stages: Forecasting and Modelling.

## Advanced Demand Forecasting

In the first part, we focus on demand forecasting specific to the route, business model of the airline, class level, load type (passengers, mail, or freight), competition level, and haul category of the route. We collect and compile data, perform feature engineering, and reorganize the data for dashboarding, exploratory data analysis, and different forecasting requirements. We implement and compare multiple forecasting algorithms, including ARIMA, ARIMAX with the number of airlines operating on a route as an exogenous variable, ARIMAX with competition level as an exogenous variable, sequential LSTM with airlines operating on a category of routes as an exogenous feature, and LSTM with competition level as an exogenous feature. We also perform grid search hyperparameter tuning to optimize these models.

## Comprehensive Business Optimization for Fleet Planning

In the second part, we create a comprehensive business model for fleet planning. This model considers forecasts from the previous section, seating configuration, aircraft-specific characteristics, the airline's financial situation, and retrofit or reconfiguration options. The decision variables in this model include buying, wet leasing, dry leasing, selling, leasing out, reconfiguration, and retrofit. Different scenarios for a hypothetical airline are created and tested for evaluation. We also perform sensitivity analysis on the financial situation, existing aircraft, retrofit costs, fuel costs, and other variables.

# Data and EDA

## Introduction

This chapter serves as a critical foundation, providing an overview of the datasets integral to our research and the subsequent feature engineering and exploratory data analysis processes. The datasets under scrutiny include the T100 dataset, capturing comprehensive information on aviation operations, as well as datasets specific to Low-Cost Carriers (LCCs) and Cargo carriers. Additionally, auxiliary datasets derived from the T100 framework, fuel price datasets, and aircraft characteristics datasets contribute essential dimensions to our analytical endeavours.

A succinct description of the origin, features, and governing authority of each dataset will be provided, setting the stage for a comprehensive understanding of the underlying data infrastructure. Following this, a detailed exploration of the feature engineering steps undertaken will unfold, elucidating the meticulous processes employed to enhance the datasets' predictive capabilities. Measures to ensure data integrity, the aggregation of data for distinct analytical purposes, and the strategic addition of features tailored for forecasting will be expounded upon.

Furthermore, this chapter will delve into a thorough exploratory data analysis, unravelling the evolving landscape of commercial aviation across the years. By illuminating patterns, trends, and anomalies within the datasets, this analysis lays the groundwork for subsequent in-depth investigations and model development. In essence, this chapter serves as a gateway to the intricate world of data that propels our research forward.

## Data Collection

### T-100

The main dataset assembled for this research, spanning over three decades from 1990 to 2023, encapsulates a comprehensive record of loads carried by all carriers on every Origin-Destination (OD) route within the United States. This monumental compilation, amounting to approximately 2 gigabytes of data, emanates from the Office of Airline Information within the Bureau of Transportation Statistics (BTS) of the U.S. Department of Transportation in a database called T-100. (US Department of Transportation, 2023) Accessing and aggregating this wealth of information involved navigating the vast landscape of aviation data, sourcing it from T100 records for each year, and integrating them into a cohesive, pivoted dataset primed for data cleaning and pre-processing.

The T-100 Domestic Market (U.S. Carriers) database, also known as the Air Carrier Statistics database, is a comprehensive dataset that contains information reported by U.S. carriers operating between airports within the United States and its territories. This data, which has been collected since 1990, includes details on passengers, freight, and mail enplaned at the origin airport and deplaned at the destination airport. These data are often referred to as "market" or on-flight origin and destination records. (US Department of Transportation, 2023)

The data is collected by the Office of Airline Information using Form T-100, which U.S. air carriers are required to fill out monthly. (US Department of Transportation, 2024) The form collects summarized flight stage data and on-flight market data for revenue flights. All traffic statistics are compiled in terms of the operating carrier, regardless of any code-sharing or joint-service agreements. The data is then processed and made publicly available, except for military data. (Cornell Law School, 2019)

The T-100 data is significant for several reasons. It provides valuable insights into domestic air travel patterns, including passenger and freight traffic. This information is crucial for researchers, analysts, and industry professionals who use it for market analysis, trend identification, and policy development. The data also plays a vital role in safety monitoring, as it allows for the comparison of enplanement data among carriers with similar operating characteristics. Carriers that expand operations at a high rate are monitored more closely for safety reasons. Furthermore, the United States is obligated to report certain air carrier data to the International Civil Aviation Organization (ICAO), and the traffic data supplied to ICAO are extracted from the U.S. air carriers' Schedule T-100 submissions. (Chadwick, 2023)

### Feature-specific Data

In preparation for the subsequent feature engineering phase, an essential prerequisite was the compilation of a comprehensive list of United States Low-Cost Carriers (LCCs) and Cargo carriers operating between 1990 and the present. Surprisingly, no official historical records of LCCs or Cargo carriers were readily available from authoritative sources, adding complexity to this task. Even though there is a broad consensus on what defines an airline as a Low-Cost Carrier, the absence of an official historical list necessitated the author's manual compilation. (List of Low-Cost-Carriers (LCCs) based on ICAO definition, 2017)

This undertaking was further complicated by the dynamic nature of the aviation industry, where airlines undergo mergers, consolidations, and alliances, resulting in the inheritance of trade names, callsigns, or codes from former entities. To address this challenge, the author meticulously curated a list by scouring literature for discussions or analyses pertaining to LCCs or Cargo carriers. This involved a comprehensive review of both new and historical papers. (Bitzan & Peoples, 2016) (Chowdhury, 2007) (Daraban, 2012) (Spiewanowski, 2015) Additionally, online sources detailing both operational and defunct LCCs and Cargo carriers were consulted and integrated into the evolving list. (Wikipedia, 2023)

It is worth noting here that, despite the presence of a cargo carrier flag in the T-100 dataset, our observation revealed that almost none of the carriers register themselves under that category. (Suissa, 2010) The reasons for this phenomenon, though beyond the scope of this work, underscore the importance of the manual curation undertaken to ensure a comprehensive and accurate representation of cargo carriers in our dataset.

An additional tactic involved investigating carriers within the data exhibiting high freight values and low passenger usage. Through a meticulous examination of these distinctive carriers, a notable number of predominantly Cargo carriers were identified. After considerable effort, the author successfully compiled a definitive list comprising 43 Low-Cost Carriers and 53 Cargo carriers that operated in the United States from 1990 to 2023. Their unique identifiers were then matched against T-100 database and stored for subsequent steps. This scrupulous curation of carrier data marks a pivotal achievement, setting the stage for the ensuing phase of feature engineering, wherein these carriers and their attributes will wield substantial influence.

A group of letters and numbers

Description automatically generated

Figure Unique identifiers of LCCs and Cargo carriers

### Fuel Price Data

The fuel price forecast is a critical component of this work, particularly for business optimization modelling, as it directly impacts operational costs and pricing strategies. In the context of aviation, the forecast of kerosene prices, which is a type of jet fuel, is especially significant. The data for these forecasts is sourced from the U.S. Energy Information Administration (EIA), which has been tracking kerosene prices since 1990. The EIA provides detailed historical data on U.S. Gulf Coast Kerosene-Type Jet Fuel Spot Price FOB (Free on Board), which reflects the price of jet fuel available for immediate delivery and excludes the cost of shipping. (US Energy Information and Administration, 2023)

This data is essential for creating accurate and reliable fuel price forecasts. The EIA's data is meticulously collected and published, offering a comprehensive view of fuel price trends over time. The spot prices are updated regularly, ensuring that the data reflects the most current market conditions.

Analysts and forecasters use this data to identify patterns, seasonal variations, and long-term trends in fuel prices. By applying various forecasting models, such as multivariate regression, analysts have predicted future fuel prices with a degree of confidence. (Haire & Machemehl, 2009) These forecasts were then used by airlines, shipping companies, and other stakeholders to make informed decisions about budgeting, fuel hedging, and pricing of services. The significance of accurate fuel price forecasts cannot be overstated. For airlines, fuel costs constitute a substantial portion of operating expenses. Therefore, the ability to anticipate changes in fuel prices can lead to more effective financial planning and competitive pricing. Moreover, the volatility of fuel prices, influenced by factors such as crude oil prices, geopolitical events, and supply-demand imbalances, adds complexity to the forecasting process. (Procurement Resource, 2022)

By leveraging historical data from the EIA, we incorporate these variables into our model, providing a more nuanced and robust forecast that accounts for a range of possible scenarios. In summary, the EIA's historical data on kerosene prices is an invaluable resource for forecasting fuel costs. The data's accessibility and granularity enable detailed analysis and the development of sophisticated forecasting models. These forecasts play a pivotal role in the strategic planning and financial management of companies reliant on kerosene-type jet fuel, ultimately affecting the broader economy and the efficiency of the transportation sector.

## Pre-processing and Feature Engineering

### Preliminary Cleaning

As previously outlined, the amassed dataset is of substantial magnitude, totalling around 2 gigabytes. The initial phase of data cleaning focused on the removal of nonsensical entries, specifically those with zero distance, and the exclusion of records reporting zero passengers, zero freight, and zero mail in their monthly data. These straightforward yet crucial steps resulted in the elimination of nearly 300,000 entries, reducing the dataset to 6.9 million entries.

### Added Features

In the feature addition phase, a new attribute called "Business Model" was introduced to the dataset, which is crucial for the analysis conducted in this work. Three primary business models were identified within the airline industry: Low-Cost Carriers (LCC), Full-Service Carriers (FSC), and Cargo carriers.

LCCs are airlines that offer highly competitive fares and fewer comforts by charging for extras like food, priority boarding, seat allocating, and baggage. FSCs, on the other hand, are traditional airlines that offer more amenities to passengers, such as in-flight entertainment, meals, and baggage allowances included in the ticket price. Cargo carriers are airlines mainly dedicated to the transport of cargo by air. (R, 2023) (E, 2020)

While the T-100 database provides its own classification of airlines based on size (as illustrated in Figure 3), this categorization was deemed irrelevant for the current study, which focuses on market demand forecasts specific to LCCs, FSCs, and Cargo carriers. A central assumption of this work is that the markets for LCCs and FSCs are relatively distinct from each other.

A screenshot of a report

Description automatically generated

Figure T-100's Classification of Airlines

To refine the classification of airlines, a multi-conditional approach was adopted, leveraging the T-100's size-based classification. Airlines were categorized as FSC, LCC, or Cargo only if they were large enough to be classified within T-100's categories 1 through 4. Smaller airlines, which did not fit the criteria for FSC, were labelled as "Small Certified Carriers" (SCC) or "Air Taxi," based on T-100's descriptions. These smaller categories, while not the primary focus of the study, provide additional data that can be analysed as needed. Categories 0 and 9 were excluded from the analysis to maintain the focus on the main business models under consideration.

The "Haul" feature, a practical addition to the dataset, replaces the "Distance" attribute by categorizing the distance of a flight into four distinct categories: "Very Short," "Short," "Medium," and "Long." The definitions for flight hauls can vary, with some based on distance, others on duration, and some locally centred definitions based on the region of the destination. For this particular dataset, the definitions provided by Eurocontrol have been adopted.

Eurocontrol, the European Organisation for the Safety of Air Navigation, defines flight haul categories based on the actual distance flown. "Very Short" haul flights typically cover distances up to 500 kilometres, "Short" haul flights range from 500 to 1,500 kilometres, "Medium" haul flights span distances between 1,500 and 4,000 kilometres, and "Long" haul flights cover distances exceeding 4,000 kilometres. (Eurocontrol, 2011) These categorizations are widely used in the aviation industry and provide a practical framework for understanding and analysing flight distances.

This meticulous approach to data cleaning and feature engineering is a testament to the importance of preparing a dataset for in-depth analysis. By removing irrelevant and erroneous data and by adding meaningful features that reflect industry-specific knowledge, the dataset is now primed for sophisticated analytical techniques that can yield insights into market dynamics and business strategies within the airline industry.

### Aggregation for EDA

The first outcome of the data cleaning and feature engineering process served the specific purpose of conducting exploratory data analysis (EDA). This dataset is tailored to individual airlines, meaning that for each carrier, on each route, at each point in time, an entry is recorded if it operated on that specific route. The variables of interest, namely passengers, freight, and mail (collectively referred to as "load" in this work), are aggregated through summation.

A noteworthy and immensely valuable characteristic of the T-100 data is its ability to be grouped based on the market areas of origin and destination. Leveraging the "OriginCityMarketID" feature, all airports operating within a city can be grouped into a singular origin or destination. This proves particularly advantageous when comparing different business models, given that Low-Cost Carriers (LCCs) are known to operate from less expensive airports, often situated slightly outside the city in neighbouring towns. Figure 4 provides a visual excerpt of market areas and their associated airports.

A screenshot of a computer

Description automatically generated

Figure A few Market Areas and their Associated Airports

### Aggregation for Forecast

In preparation for forecasting, where market or route demand transcends individual air carriers, our focus shifted to aggregating data based on business models, adhering to the earlier assumption that the Low-Cost Carrier (LCC) market is distinct from Full-Service Carriers (FSC). (Daft & Albers, 2015) Thus, the aggregation grouped airlines together based on their business model and operating routes.

This aggregation process introduced a noteworthy feature: the count of airlines operating on a specific route at a particular point in time. This feature, aptly named "Airlines," enables the capture of competition levels for a route within each business model at distinct points in time. As elucidated later, this feature assumes the role of an exogenous variable crucial for forecasting market demand.

Derived from the "Airlines" feature, another valuable variable emerged— "Competition." While predicting the exact number of airlines on a specific route might challenge even experts, assessing the degree of competition proves to be a more feasible endeavour. The "Competition" feature is categorized “Autonomy” if only one airline operates in the route, categorized as "Low" for 2 to 5 airlines, "Moderate" for 6 to 10 airlines, and "High" for more than 10 airlines on a route. This pragmatic approach ensures a more realistic assessment of competition, laying the foundation for improved forecasting accuracy.

The significance of these features is underscored by their relevance to industry research and regulatory assessments. For instance, the U.S. Government Accountability Office (GAO) has utilized T-100 enplanement data to evaluate market structure characteristics, including market concentration and the number of effective competitors. These indicators are crucial for assessing the potential degree of competition within the airline industry and its impact on consumers. (United States Government Accountability Office , 2014)

## Exploratory Data Analysis

### Overview

The evolution of business models within the air carrier spectrum becomes evident when examining the market share and the number of airlines in 1990, as depicted in Figure 5. Despite Low-Cost Carrier (LCC) airlines comprising 11% of the market share, the number of LCC airlines (13%) exceeded those in the Full-Service Carrier (FSC) category, revealing a heightened competition for unit LCC market demand compared to FSC. In 2023 (Figure 6), LCCs and FSCs dominate the market share, with LCCs experiencing a threefold increase.

A screenshot of a graph

Description automatically generated

Figure Market Share and Number of Airlines in 1990

A screenshot of a graph

Description automatically generated

Figure Market Share and Number of Airlines in 2023

Figures 7 to 11 illustrate the top 100 routes operated by different business models in 2023, based on passengers enplaned for airlines and freight for Cargo carriers. Notable insights include FSCs focusing on the east coast and long-haul flights, LCCs operating diverse, shorter-range routes, Small Certified Carriers (SCCs) in Alaska, Air Taxis in regional short routes, and Cargo carriers primarily serving Alaska.

A map of the united states

Description automatically generated

Figure Top 100 routes of FSC airlines in 2023

A map with lines and dots

Description automatically generated

Figure Top 100 routes of LCC airlines in 2023

A map of the united states

Description automatically generated

Figure Top 100 routes of SCC airlines in 2023

A map of the united states

Description automatically generated

Figure Top 100 routes of Air Taxis in 2023

A map of the united states

Description automatically generated

Figure Top 100 routes of Cargo carriers in 2023

These diagrams indicate a few important insights:

1. FSCs are more focused on the east coast and the long-haul flights.
2. LCC routes are more diverse, shorter in range and more focused on the mainland.
3. SCCs operate predominantly in Alaska.
4. Air Taxis operate regional often very short or short routes.
5. Cargo carriers offer service mainly to and from Alaska.

### The Pandemic Effect

Examining the aftermath of the pandemic reveals distinctive trends in the aviation industry. In 2020, there was a significant decline in passengers enplaned, as illustrated in Figure 12. Over the subsequent three years, the recovery in passenger numbers was gradual. Intriguingly, freight enplaned demonstrated a contrasting pattern, experiencing growth during the lockdown and subsequently returning to pre-pandemic levels.

A graph of blue and red lines

Description automatically generated Figure Passengers, Freight and Mail Enplaned Pre- and Post-Pandemic

The dynamics of market share, depicted in Figure 13, exhibit a temporary shift during the lockdown. Low-Cost Carriers (LCCs) temporarily gained a larger share, but this was short-lived, quickly reverting to previous levels.

A graph of blue and pink bars

Description automatically generated

Figure Market Share of Different Business Models Pre- and Post-Pandemic

The average number of airlines operating on each route, outlined in Figure 14, reveals a pre-pandemic gradual decline. However, during the lockdown, the number of airlines plummeted with the number of passengers enplaned, followed by only a partial recovery post-pandemic. Notably, the unique trajectory of LCCs, as isolated in Figure 15, showcases an upward trend pre-pandemic, which accelerated after the pandemic.

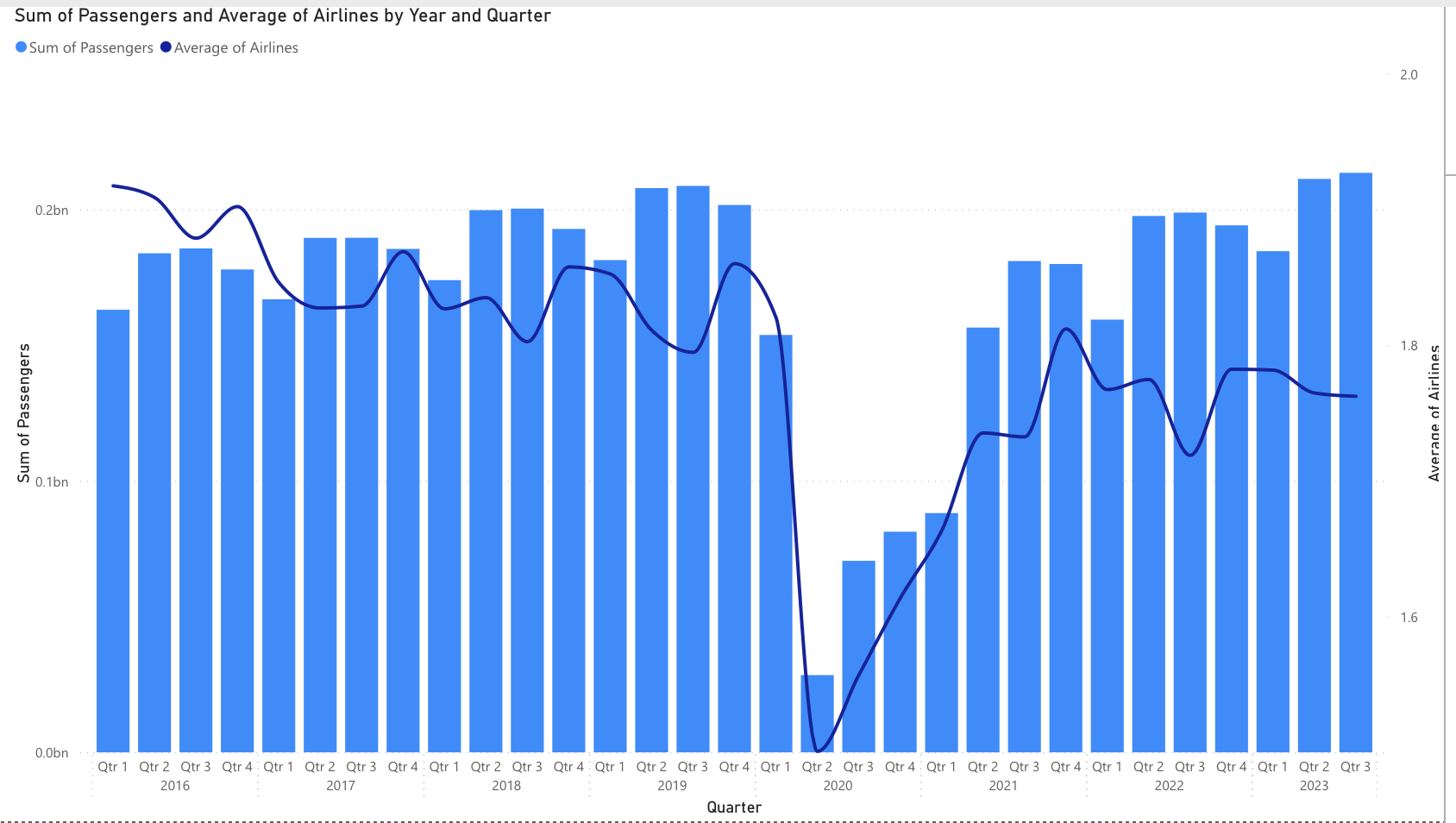


Figure Average Number of Airlines Operating at Each Route

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated

Figure Competition in the LCC and FSC Markets

Figure 16 provides insights into the number of routes operated by both LCCs and Full-Service Carriers (FSCs). It becomes apparent that LCCs demonstrated a continuous increase in the number of routes both pre- and post-pandemic. This observation suggests not only the expansion of the LCC market but also an increasing level of competition, as more LCC airlines vie for space on an average route.

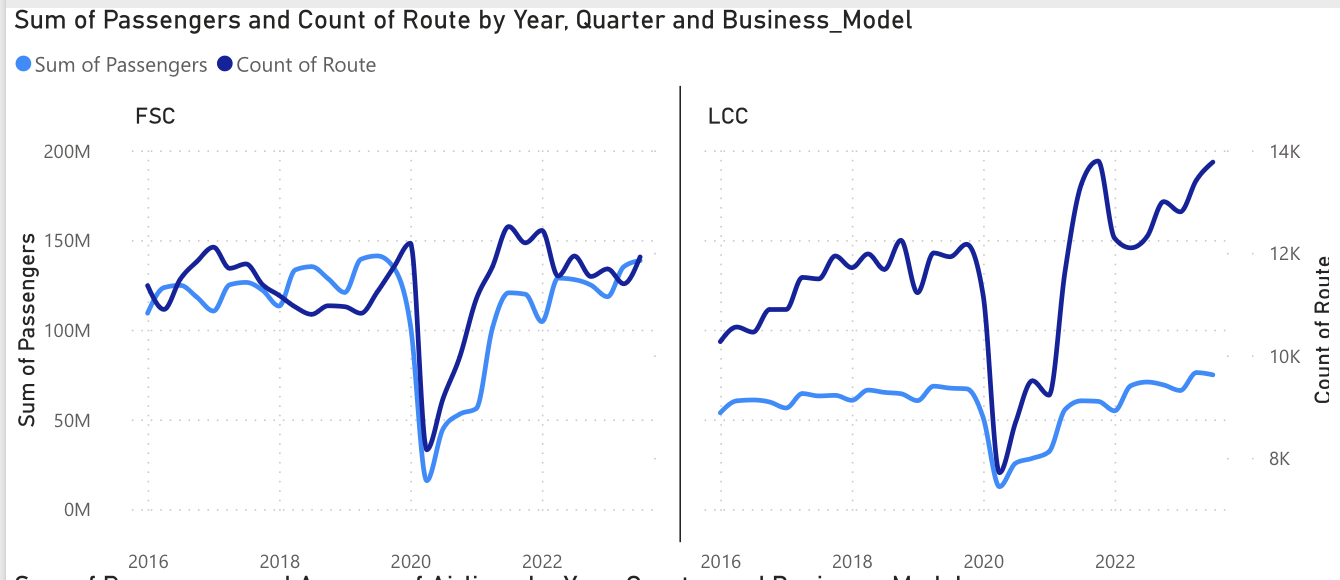


Figure Number of Routes Operated by LCCs and FSCs

And as is evident in Figure 17, this market growth is associated with Short and Medium-haul flights.

A screenshot of a graph

Description automatically generated

Figure Different Flight Hauls by Business Model over Time

### Conclusion

The extensive Exploratory Data Analysis underscores that an airline's business model significantly impacts market demand, confirming the separation of LCC and FSC market demands. Notably, the LCC market is expanding, emphasizing the dependence of an airline's proposed route on its business model. This insight is in line with addressing Sub-questions A and C of the research question.

# Forecast methodology and analysis

Forecasting future demand is a critical component in strategic planning across various industries, including aviation. There are two primary approaches to forecasting: qualitative and quantitative methods. Qualitative methods, such as the Delphi method, decision trees, and group discussions, are best suited for scenarios with no historical data or where mathematical modelling is impractical. They leverage expert opinions and intuition, allowing for a broad consideration of potential scenarios. However, these methods are subjective and can yield inconsistent results due to their susceptibility to bias and external influences.

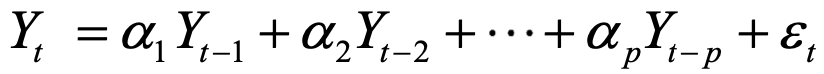
Quantitative methods, on the other hand, rely on statistical analysis of historical data. They are categorized into causal models, which explore relationships between variables, and time series analysis, which predicts future trends based on past data patterns. Time series analysis, including techniques like regression and ARIMA, is particularly favoured in aviation due to the industry's data limitations and the risk of spurious correlations in regression models with numerous independent variables. (Do, Lo, Chen, Le, & Anh, 2020) The subsequent chapters will focus on the ARIMA and LSTM models, both of which fall under time series analysis and are renowned for their predictive capabilities in time-dependent data scenarios such as airline fleet planning.

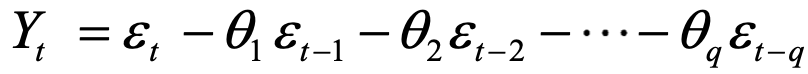
## ARIMA

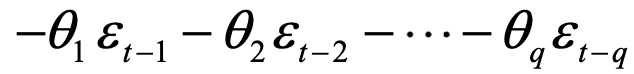
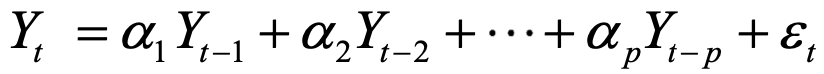
The Autoregressive Integrated Moving Average (ARIMA) model is a traditional statistical method used for time series forecasting. (Siami-Namini, Tavakoli, & Siami Namin, 2018) Developed by statisticians George Box and Gwilym Jenkins in the 1970s, ARIMA combines both autoregression (AR) and moving average (MA) models, along with a differencing pre-processing step to make the sequence stationary, known as integration (I). (Box & Jenkins, 1970)

ARIMA models are particularly useful for analysing and forecasting time-dependent data, where past values and errors influence future predictions. The model is characterized by three parameters: p (the number of autoregressive terms), d (the number of non-seasonal differences), and q (the number of moving-average terms). These parameters are used to capture the temporal structures in time series data and employ a linear regression-based forecasting approach.

ARIMA model with difference value (d = 0) is equal to the ARMA model. ARMA itself is consistent of AR(p) and MA(q):

AR(p): 

MA(q): 

ARMA(p, q): 

where, αp is autoregressive coefficient, p autoregressive order, εt the white noise, q the order of moving average, and Θq is the moving average coefficient. ARIMA can then be constructed by integration. Here, we introduce the integrator operator Δd :  
ARIMA(p, d, q): 



Adding seasonal components will result in what is sometimes called SARIMA, and adding exogenous variables results in ARIMAX. Therefore, the most general form of ARIMAX is as follows:

SARIMAX(p, d, q)(P, D, Q, s): A close up of a logo

Description automatically generated

Where P, D and Q are the seasonal lags corresponding to p, d and q, s is the number of time lags comprising one full period of seasonality, βi is the exogenous variable coefficient, and xti is the exogenous variable. (Chatfield, 2003)

## LSTM

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) that has gained popularity in recent years for its ability to effectively model and predict time series data. Unlike traditional feedforward neural networks, LSTM has feedback connections that make it a "general purpose computer" - it can process single data points (such as images) or entire sequences of data (such as speech or video).

LSTM models are particularly effective when dealing with long sequences and sequences with large gaps between relevant information. This is due to their unique cell state, which can carry information across long sequences without suffering from the vanishing or exploding gradient problem, a common issue in traditional RNNs.

A diagram of a plant

Description automatically generated

Figure The repeating module in an LSTM (Zhang, et al., 2021)

In this work, both univariate and multivariate LSTM models were used for forecasting market demand. The univariate LSTM model uses a single feature (market demand) at a time step to predict the same feature at the next time step. On the other hand, the multivariate LSTM model uses multiple features at a time step to predict a single feature or multiple features at the next time step. The performance of these models was evaluated using MSE, and MAE, internally, and using MAPE externally.

The LSTM models were compared with the ARIMA and ARIMAX models to determine which method provided the most accurate forecasts. The results of this comparison will provide valuable insights into the effectiveness of these advanced analytics techniques in optimizing airline fleet planning.

## Exogenous Variables

To address sub-question B of the research question, the study leverages the author's expertise in the aviation industry, combining expert opinion and intuition. Drawing on this background, the identified exogenous variables deemed most pertinent for inclusion in the forecast model are the number of airlines operating on a specific route and the anticipated competition level within a specific market. These variables, thoroughly detailed in Chapter 4, provide crucial insights for predicting market demand accurately.

A diagram of a graph

Description automatically generated

Figure A basic map of how the relationship between exogenous and endogenous variables are understood in the new method (Jacaruso, 2018)

The forecasting models employed encompass a Multivariate Long Short-Term Memory (LSTM), an advanced version extending from the Univariate LSTM. Additionally, the study utilizes an Autoregressive Integrated Moving Average (ARIMA) model, specifically SARIMA, which omits the seasonal S due to its limited prevalence in literature. This ARIMA model will be benchmarked against ARIMAX, essentially SARIMAX, enriched with an exogenous variable—Airlines and Competition—identified in the earlier chapters.

The final layer of analysis introduces a Multivariate LSTM, integrating the exogenous variables. A heuristic analysis will be systematically conducted to discern the most fitting model across various market segments, flight hauls, and load types. This rigorous testing and comparative approach aim to identify the most effective forecasting model, aligning with the intricacies of diverse market scenarios and operational conditions within the aviation industry.

## Fitness Metrics

In time series forecasting, various metrics are used to evaluate the fitness of a model. Commonly used metrics include Root Mean Square Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Akaike Information Criterion (AIC). These metrics are essential for assessing the accuracy and performance of forecasting models. In this chapter, we will discuss these metrics and establish the rationale for choosing Mean Absolute Percentage Error (MAPE) as the primary metric for our analysis.

### Common Fitness Metrics

#### Mean Absolute Error (MAE)

MAE is a measure of errors between paired observations, emphasizing the magnitude of the errors without considering their direction. It provides a more intuitive understanding of the average error.

A black and white math equation

Description automatically generated with medium confidence

#### Root Mean Square Error (RMSE)

RMSE is a widely used metric for assessing the accuracy of predictions obtained by a model. It measures the differences between actual and predicted values, penalizing large errors and scaling the scores in the same units as the forecast values.

A square root of a mathematical equation

Description automatically generated

#### Akaike Information Criterion (AIC)

AIC is a statistical measure used to evaluate the goodness of fit of a model. It takes into account the model's complexity and the amount of information it captures, allowing for model comparison. It quantifies the goodness of fit for the model as well as the simplicity of the model. It is only meaningful in relative comparison of fit and is extensively used for evaluating ARIMA. (Zhang, et al., 2021)

### Rationale for Choosing MAPE

While these metrics are valuable, we have chosen to primarily use MAPE for evaluating our forecasting models. It is short for Mean Absolute Percentage Error and is defined as follows:

A black and white math equation

Description automatically generated with medium confidence

Together with MSE, and RMSE, these three measures often, but not always, give the same rankings of fits. (Albright & Winston, 2020) However, MAPE is preferred because it is easier to interpret and provides a clear understanding of the average percentage error. Additionally, MAPE is a normalized metric, allowing for the evaluation of different loads in our data, such as passengers, freight, and mail, against each other. This normalization enables fair comparisons across diverse data categories, which is essential in the context of airline fleet planning where multiple types of data need to be evaluated consistently.

By selecting MAPE as our primary fitness metric, we aim to ensure a comprehensive and interpretable assessment of the forecasting models, considering the specific requirements of our analysis.

### R-Squared

In the context of evaluating the fitness of a model, the R-squared (coefficient of determination) has been incorporated to assess the goodness of fit of the model. This measure, in conjunction with the Mean Absolute Percentage Error (MAPE) of the forecast, serves the purpose of identifying overfitting in the model.

R-squared is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. It provides insights into how well the independent variables explain the variability of the dependent variable. When used alongside MAPE, it helps in understanding whether the model is overly tailored to the training data, which can lead to poor generalization to new data.

The inclusion of R-squared in the evaluation process enhances the assessment of model fitness by considering both the goodness of fit and the predictive accuracy. This comprehensive approach aids in identifying potential issues such as overfitting, thereby contributing to the robustness of the model evaluation.

## Results

### ARIMA

The research undertakes an analysis of 60 cross-category combinations, encompassing business model, haul, and load. The underlying assumption posits that each category holds unique characteristics resistant to aggregation. With a focus on LCCs and FSCs, the evaluation centres on passengers from very short, short, and medium-haul LCC markets, along with very short, short, medium, and long-haul FSC markets. Table 1 presents the MAPE scores from the ARIMA model and ARIMAX models with exogenous variables (Airlines and Competition):

Table MAPE Scores of Selected Categories with ARIMA

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Business Model | | Haul | ARIMA | ARIMAX Airlines | | ARIMAX Competition | |
| FSC | Long | | 122.31% | 116.86% | 121.45% | |
| FSC | Medium | | 96.32% | 94.67% | 95.16% | |
| FSC | Short | | 92.32% | 93.26% | 92.18% | |
| FSC | Very Short | | 201.41% | 196.22% | 190.15% | |
| LCC | Medium | | 90.07% | 86.96% | 87.92% | |
| LCC | Short | | 256.85% | 253.97% | 265.05% | |
| LCC | Very Short | | 81.07% | 83.58% | 92.37% | |

Table 2 displays the corresponding R-Squared scores:

Table R-Squared Scores of Selected Categories with ARIMA

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Business Model | Haul | ARIMA | ARIMAX Airlines | ARIMAX Competition |
| FSC | Long | 58.20% | 59.51% | 60.37% |
| FSC | Medium | 63.66% | 63.64% | 64.02% |
| FSC | Short | 72.35% | 72.56% | 71.04% |
| FSC | Very Short | 83.73% | 83.81% | 83.85% |
| LCC | Medium | 91.53% | 91.55% | 91.59% |
| LCC | Short | 90.82% | 91.00% | 90.98% |
| LCC | Very Short | 83.98% | 83.55% | 83.66% |

While none of the results strictly adhere to stringent acceptance criteria, acknowledging the lockdown era in the test set prompts a consideration of looser MAPE criteria. Notably, LCCs' very short-haul flights exhibit fewer errors, albeit diminishing with the inclusion of any exogenous variables. For FSC market long and medium-haul flights, and LCC market medium and short-haul flights, knowledge of the number of operating airlines can enhance ARIMA's forecasts by an average of 16%. Simultaneously, estimating competition levels for FSC market short or very short-haul routes can yield an average improvement of 6% in predictions.

An intriguing aspect of these results lies in the selection of models based on R-squared values, where the outcomes differ significantly. This emphasizes that overfitting constitutes a major source of error. It is crucial to emphasize that these averages have been aggregated for analytical convenience, and each route in the model possesses unique metrics and results. Figures 20-25 provide specific examples of route-specific outcomes:

A graph with lines and numbers

Description automatically generated

Figure An Example of FSC Forecast

A graph of a graph showing the same number of passengers

Description automatically generated with medium confidence

Figure An Example of FSC Fit

A graph with lines and numbers

Description automatically generated

Figure An Example of LCC Forecast

A graph of cargo freight

Description automatically generated

Figure An Example of LCC Fit

A graph of a graph with numbers and lines

Description automatically generated with medium confidence

Figure An Example of Cargo Forecast

A graph of a number of passengers

Description automatically generated

Figure An Example of Cargo Fit

### LSTM

While ARIMA results were obtained for the top 100 individual routes of each cross-category and then aggregated for each cross-category, LSTM, due to its computational expenses, necessitated a different approach. Routes were initially aggregated and then forecasted. While this solution may not be optimal, it was chosen solely for the purpose of an overall comparison between the two algorithms. It's crucial to note that our model is fully capable of training the LSTM algorithm for individual routes and forecasting based on individual route parameters.

Table 3 presents the MAPE results for the seven categories central to this work, and the R-squared values of each fit are displayed in Table 4:

Table MAPE Scores of Selected Categories with LSTM

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Business Model | | Haul | LSTM Airlines | | LSTM Competition | |
| FSC | Long | | 105.36% | 110.65% | |
| FSC | Medium | | 56.37% | 58.63% | |
| FSC | Short | | 52.62% | 55.48% | |
| FSC | Very Short | | 73.54% | 71.70% | |
| LCC | Medium | | 61.54% | 63.26% | |
| LCC | Short | | 58.97% | 59.51% | |
| LCC | Very Short | | 76.30% | 73.47% | |

Table R-Squared Scores of Selected Categories with LSTM

|  |  |  |  |
| --- | --- | --- | --- |
| Business Model | Haul | LSTM Airlines | LSTM Competition |
| FSC | Long | -44.22% | -12.73% |
| FSC | Medium | -93.68% | -146.81% |
| FSC | Short | -488.51% | -563.10% |
| FSC | Very Short | -132.48% | -45.08% |
| LCC | Medium | 1.11% | -1.89% |
| LCC | Short | -2.27% | -1.99% |
| LCC | Very Short | 0.30% | -0.34% |

While the MAPE values fall within an acceptable range, justified by the lockdown era, the R-squared values are considerably off, resembling those of a baseline model. Figures 26-31 display plots for the forecasts vs. actual and fit vs. actual of three sample categories.

A graph of a line graph

Description automatically generated with medium confidence

Figure LSTM Forecasts for LCC Medium-haul Passengers

A graph of a line graph

Description automatically generated with medium confidence

Figure LSTM Fit for LCC Medium-haul Passengers

A graph of a graph of a cargo

Description automatically generated with medium confidence

Figure LSTM Forecasts for Cargo Long-haul Freight

A graph of a plane

Description automatically generated with medium confidence

Figure LSTM Fit for Cargo Long-haul Freight

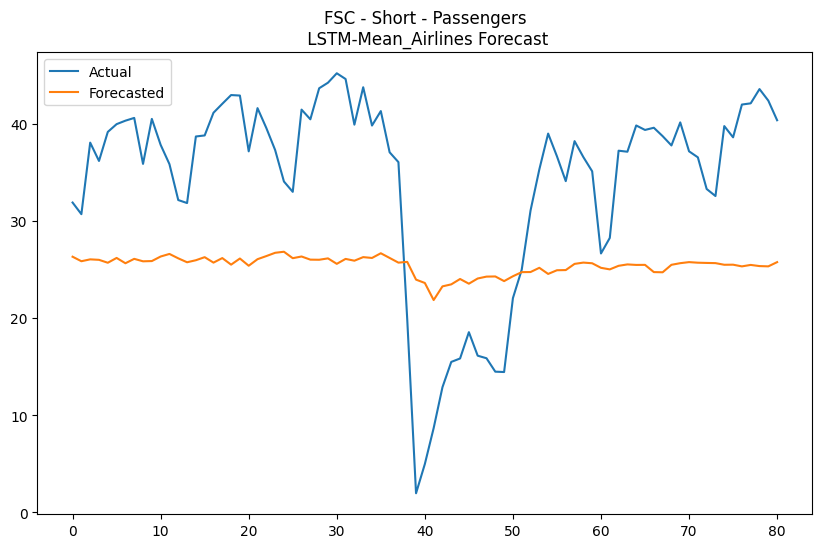


Figure LSTM Forecasts for FSC Short-haul Passengers

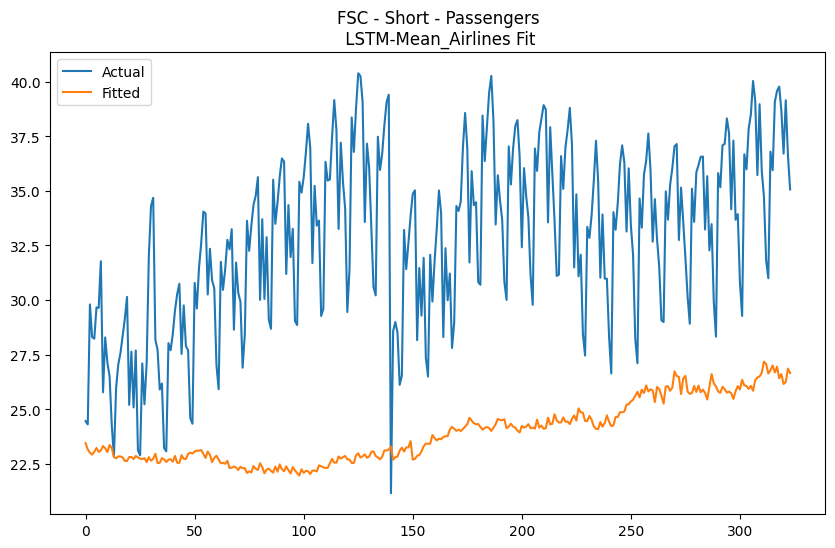


Figure LSTM Fit for FSC Short-haul Passengers

#### Hyperparameter Tuning

To enhance the performance of the LSTM model, a grid search hyperparameter tuning was conducted specifically for one of our focal categories. Subsequently, these optimized hyperparameters were applied uniformly across all categories. While this may not be the most favourable approach, it is a practical necessity given the computational expenses associated with LSTM.

Tables 5 and 6 present the MAPE and R-squared results, respectively, for LSTM after hyperparameter tuning:

Table MAPE Scores of Selected Categories with LSTM (Hyperparameter Tuning)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Business Model | | Haul | LSTM Airlines | | LSTM Competition | |
| FSC | Long | | 113% | 112% | |
| FSC | Medium | | 57% | 58% | |
| FSC | Short | | 52% | 52% | |
| FSC | Very Short | | 73% | 72% | |
| LCC | Medium | | 61% | 64% | |
| LCC | Short | | 59% | 59% | |
| LCC | Very Short | | 74% | 74% | |

Table R-Squared Scores of Selected Categories with LSTM (Hyperparameter Tuning)

|  |  |  |  |
| --- | --- | --- | --- |
| Business Model | Haul | LSTM Airlines | LSTM Competition |
| FSC | Long | 1% | -3% |
| FSC | Medium | -4% | -3% |
| FSC | Short | 10% | -9% |
| FSC | Very Short | 19% | -4% |
| LCC | Medium | 1% | 0% |
| LCC | Short | 0% | 0% |
| LCC | Very Short | 0% | 0% |

The presented tables provide valuable insights into the performance of the LSTM model across various categories, along with the outcomes following the fine-tuning of hyperparameters for one focus category, applied universally across all categories. This underscores the pragmatic compromise necessitated by computational challenges while striving to maintain a level of optimization. Additionally, it acknowledges the computational constraints that led to the aggregation of routes. It is crucial to interpret these results within the context of the chosen approach, recognizing that the option for individual route training exists within the model.

Notably, some results from hyperparameter tuning exhibit a deterioration compared to pre-tuning conditions, emphasizing the need for individualized tuning. As mentioned earlier, the conceptual framework, the codebase, and the model itself are versatile, enabling both hyperparameter tuning for individual categories (and even individual routes) and individual route forecasting. This flexibility allows for optimization based on specific requirements and preferences.

Figures 32-37 depict the same categories presented in Figures 26-31 but with the new tuned model. It is noteworthy that the fit for FSC Short-haul Passengers has significantly improved. This improvement was evident in Table 6, where the R-squared value improved by 9-fold. The reason for this dramatic enhancement is that this category was chosen as the focus category for hyperparameter tuning.

A graph of a graph

Description automatically generated

Figure Tuned LSTM Forecasts for LCC Medium-haul Passengers

A graph with blue lines and orange lines

Description automatically generated

Figure Tuned LSTM Fit for LCC Medium-haul Passengers

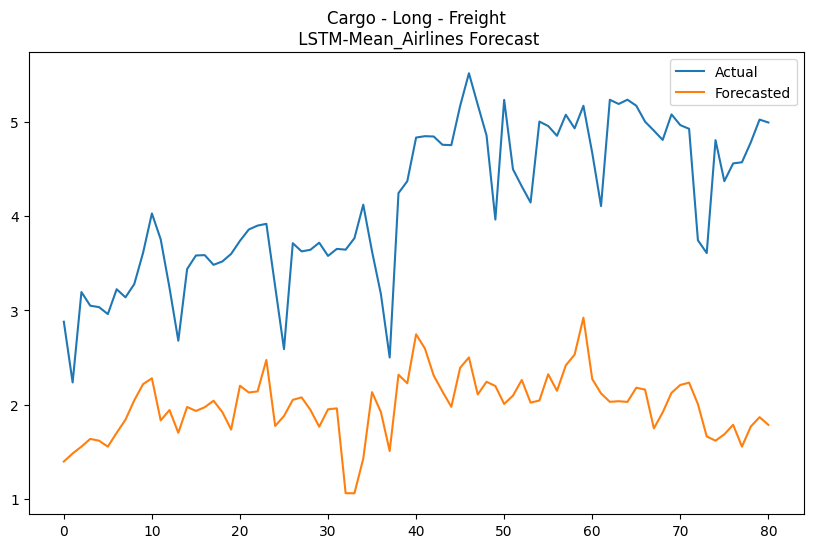


Figure Tuned LSTM Forecasts for Cargo Long-haul Freight

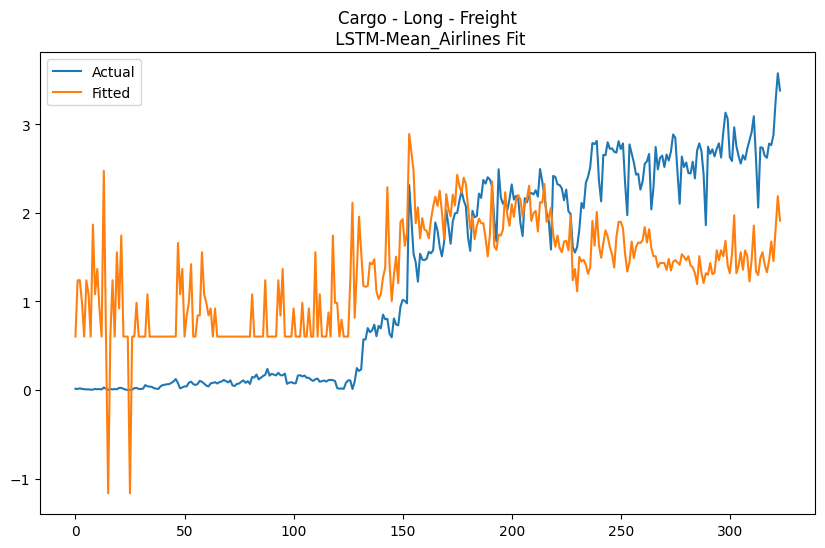


Figure Tuned LSTM Fit for Cargo Long-haul Freight

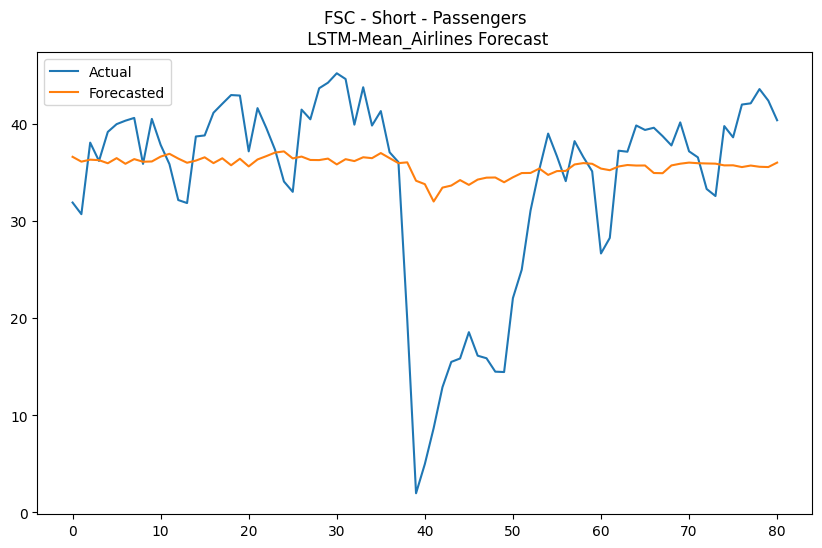


Figure Tuned LSTM Forecasts for FSC Short-haul Freight

A graph of a flight

Description automatically generated with medium confidence

Figure Tuned LSTM Fit for FSC Short-haul Freight

## Comparison and Conclusion

The comparative analysis between ARIMA and LSTM models requires careful consideration of nuances introduced by different aggregation processes. ARIMA, trained and tested on the top 100 individual routes and then averaged, contrasts with LSTM, which, due to computational limitations, aggregated all routes before training and testing.

### Results on Aggregated Routes Data

Tables 7 and 8 present the results of ARIMA modelling on aggregated routes data, confirming our hypothesis of the need for tailored forecasting approaches:

Table MAPE Results for ARIMA on Aggregated Routes Data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Business Model | | Haul | ARIMA | ARIMAX Airlines | | ARIMAX Competition | |
| FSC | Long | | 128.93% | 130.07% | 129.13% | |
| FSC | Medium | | 58.60% | 66.07% | 58.38% | |
| FSC | Short | | 59.01% | 64.65% | 61.78% | |
| FSC | Very Short | | 60.98% | 73.02% | 72.11% | |
| LCC | Medium | | 58.02% | 57.64% | 57.85% | |
| LCC | Short | | 64.74% | 66.03% | 65.92% | |
| LCC | Very Short | | 73.19% | 73.10% | 73.03% | |

Table R-Squared Results for ARIMA on Aggregated Routes Data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Business Model | | Haul | ARIMA | ARIMAX Airlines | | ARIMAX Competition | |
| FSC | Long | | 74.69% | 75.01% | 74.77% | |
| FSC | Medium | | 66.13% | 62.84% | 66.14% | |
| FSC | Short | | 53.81% | 54.28% | 54.21% | |
| FSC | Very Short | | 59.12% | 63.98% | 63.60% | |
| LCC | Medium | | 97.05% | 97.06% | 97.08% | |
| LCC | Short | | 94.93% | 94.47% | 94.97% | |
| LCC | Very Short | | 85.07% | 85.07% | 85.10% | |

Comparing these results with LSTM’s results in Tables 5 and 6 highlights the diverse behaviors exhibited by distinct markets (business models) and varying flight lengths (hauls), substantiating our hypothesis that distinct and varying flight demands require tailored forecasting approaches.

### Best Predictors for Each Focus Category

Table 9 outlines the best-performing predictors for each category based on a heuristic analysis, emphasizing the need for a nuanced approach based on the specific characteristics of each market and flight length.

Table Best Predictor for each Focus Category

|  |  |  |
| --- | --- | --- |
| Business Model | Haul | Best Predictor |
| FSC | Long | LSTM Airlines |
| FSC | Medium | LSTM Airlines |
| FSC | Short | LSTM Competition |
| FSC | Very Short | ARIMA |
| LCC | Medium | ARIMAX Airlines |
| LCC | Short | LSTM Airlines |
| LCC | Very Short | ARIMAX Competition |

### Key Insights and Future Directions

Our analysis underscores the effectiveness of LSTM with the average number of airlines operating on a route category as exogenous variable, in FSC long and medium-haul and in LCC short-haul markets. It also highlights the imperative for individual hyperparameter tuning across the 60 categories and the need for additional computational power for individual route tuning. Looking forward, individual route forecasting remains crucial for capturing the intricacies of specific routes and contributing to the optimization of airline fleet planning through advanced analytics. This conclusion sums up this work’s answer to sub-question C, whereby it shows the best approach to address the inherent uncertainty of the demand pattern is to categorize the market, individualize the routes, and perform forecasting on individual routes while experimenting different exogenous variable per individual route.

# Business optimisation model description and analysis

## Underlying Concepts

### Hub and Spoke Network

In the airline industry, the Hub and Spoke (HS) network is a system where traffic is routed through central hub airports, with spokes connecting the hubs to the final destinations. This model contrasts with the point-to-point system, where flights travel directly between destinations without any intermediate stops. The HS network has become a prevalent model for airlines due to its efficiency and ability to offer more frequent flights between a larger number of destinations.

A diagram of a pair of pairs of shoes

Description automatically generated with medium confidence

Figure Point to Point and Hub and Spoke Networks (Skalovskaia, 2019)

Our model addresses an integrated Hub Location Problem (HLP), which is a critical aspect of the HS network design. The HLP involves determining the optimal locations for hubs that will serve as central points for traffic consolidation and distribution. In our model, the hub location decision is solved simultaneously with the fleet planning model, which is essential for ensuring that the chosen hubs can be effectively served by the airline's fleet. (Mohri, Nasrollahi, Pirayesh, & Mohammadi, 2022)

The integrated approach to hub location and fleet planning is a sophisticated approach that simultaneously solves for hub locations and fleet planning. This dual consideration is vital for designing an optimal hub network that aligns with the airline's operational capabilities and strategic objectives, ultimately leading to a more efficient and cost-effective airline operation.

### Variable Routing Cost and DOC

In our study, we employ a model to estimate the operational costs of an aircraft for a given route within a Hub and Spoke Network. This model is grounded in the Direct Operating Cost (DOC) framework, which delineates the fixed and variable components of flight operating costs. The DOC model, as formulated by (Yu, Hong, & Peiwen, 2012), is expressed mathematically as:

Here, signifies the DOC per cycle in dollars, is the flight length in nautical miles, stands for seat capacity, and is the design range. The coefficients through are regression coefficients that remain constant across different aircraft types.

The model uses the NASA 97 method to generate aircraft DOC samples, which are then used to determine the functional form between DOC and flight lengths. It rests on three foundational assumptions: (i) the fuel density is 1.5 Lb/Gal, (ii) the fuel price is set at 1.5 Dollars/Gal, and (iii) the maintenance labour rate is fixed at 25 Dollars. These assumptions are reflective of the fuel prices and maintenance labour costs as of the year 2011.

The original formula was restructured into a new form:

In this revised equation, and are variables that depend on the aircraft type, while represents the distance of the route and varies based on the origin-destination (OD) pair. Both the slope (variable component) and the intercept (constant component) are influenced by the prices of oil and labour, which necessitates accurate forecasting to ensure effective optimization of the model.

### Specific Range

In the realm of aircraft operations, particularly within the Hub and Spoke Network, the concept of specific range is a critical factor in determining the efficiency and cost-effectiveness of flight routes. The specific range of an aircraft refers to the optimal distance it can travel under certain conditions before requiring additional fuel, which has significant implications for the Direct Operating Cost (DOC).

Antonio Filippone, in his work (Filippone, 2012), elucidates the complexities associated with the specific range of aircraft. He notes that for long-haul flights, carrying additional fuel for the entire journey increases fuel consumption. There exists a breakeven point where it becomes more efficient to stop and refuel rather than continue flying non-stop. This is due to the extra weight of carrying additional fuel, which increases fuel burn. For example, a Boeing 777-300 is more fuel-efficient on non-stop flights that are less than 3,000 nautical miles. Beyond this distance, the efficiency gains from refuelling at a midpoint outweigh the losses from additional take-off and landing cycles.

The specific range is not only a function of aircraft design but also of operational strategy. Very long non-stop flights may require limiting the number of seats to accommodate the fuel weight, which affects the revenue potential of the flight. The critical fiscal factor for these flights is the fuel burned per seat-nautical mile. This was a contributing factor to the cancellation of some of the world's longest commercial flights around 2013, such as the Singapore Airlines' New York to Singapore route, which was limited to 100 business class passengers due to the weight of the fuel required for the 10,300-mile journey. (Park, 2013)

In our work, we consider the specific range of aircraft as a constraint to ensure that the operational costs are optimized. By doing so, we can model more realistic scenarios that reflect the economic realities of long-haul flights, considering the trade-offs between carrying additional fuel and the potential revenue from passenger seats. This approach allows us to develop a more accurate and financially viable fleet planning and route optimization strategy within the Hub and Spoke Network.

## Our Contributions

While fleet planning models have been in existence since the 1990s, the integration of these models with the Hub and Spoke (HS) network has only recently been explored. In this relatively short period, significant research has been conducted on this integration, and our work contributes to two previously unexplored domains. Addressing our research sub-question B, our model incorporates two pragmatic and underappreciated managerial decision-making options:

### Retrofit

Retrofitting involves integrating new technologies or modifications into existing aircraft to enhance their performance, efficiency, and compliance with evolving regulations. Particularly relevant for older aircraft facing challenges like non-compliance with new regulations or uneconomical operation in changing conditions (such as increasing fuel prices), retrofitting allows for the enhancement of existing aircraft without the need for complete replacement. This, in turn, extends their operational lifespan and improves their environmental and economic performance.

Various retrofit options are available, including:

1. Blended Winglets: Boeing and Airbus offer winglets as a retrofit option for various aircraft types, with the potential to achieve significant fuel savings.
2. Cabin Weight Reduction: Using lightweight components contributes to lower fuel consumption and emissions, thereby improving the economic performance of the aircraft as well as its range.
3. Electric Taxiing: This technology significantly reduces fuel burn during taxiing, leading to lower emissions and operational costs.
4. Re-engining: This retrofit option offers substantial fuel savings and emission reductions, contributing to the environmental and economic performance of the aircraft.

In our model, we introduced retrofit as a component, where aircraft can undergo one or multiple available retrofit options, resulting in decreased values for and in the routing cost formula. This, however, comes at the expense of the initial cost of . In our formulation, is multiplied by and is multiplied by , where is the decision variable indicating whether a retrofit was selected for aircraft or not. Furthermore, and are cost reducing coefficients.

### Reconfiguration

Each aircraft type can have multiple seating configurations. For example, while Ryanair accommodates 189 passengers on its Boeing 737-800s, others opt for a two-class seating with only 162 passengers in total. (Adriana, 2023) Interestingly, there has been minimal research attempting to find the optimal configuration. In our work, we endeavoured to integrate seat configuration into the broader context of fleet planning, Hub Location Problem (HLP), and retrofit options.

Our model begins with an aircraft's initial seating configuration, assuming all seats are economy class, and then determines the optimal number of business class seats to be added to the airplane.

Considering the Direct Operating Cost (DOC) that was presented earlier formula for operational costs per route, the total number of seats influences variable costs primarily through its impact on fuel consumption and secondary effects on items like meals. Given that the ratio of space occupied by a business seat to the space occupied by an economy seat, the DOC formula can be expressed as:

Simplified, this formula becomes:

In this revised equation, and are variables that depend on the aircraft type, M and N are constant numbers (equivalent to and ), while represents the distance of the route varying based on the origin-destination (OD) pair. All four components of slope (variable component) and the intercept (constant component) are influenced by oil and labour prices, necessitating accurate forecasting for effective model optimization.

## Model

Our model is part of a broader approach to airline operations, which includes fleet planning. There are two main types of models in fleet planning: maximisers and minimizers. Maximiser models aim to maximize an airline's profit by increasing the flow of revenue passengers carried by the airline, considering the limitations of the airline’s fleet. In this approach, market demand is a constraint, and the airline’s objective is to increase its share of the market.

On the other hand, minimizer models, like the one proposed here, seek the most cost-effective way to transport all passengers from origin to destination. In this approach, market demand (or the company’s share thereof) is a given, and the airline’s fleet capacity is the constraint. The objective function is to minimise the operating costs.

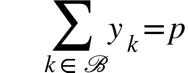
Our model integrates fleet optimization, retrofit decisions, hub location selection, and seating configuration in a unified approach. This comprehensive model allows airlines to make more informed decisions about their operations, leading to increased efficiency and profitability.

### Nomenclature

### Mathematical Formulation

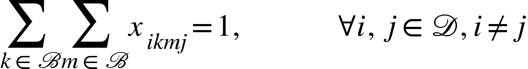
{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><mo>&#xA0;</mo><mtext>Min</mtext><munder><mo>&#x2211;</mo><mrow><mi>i</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi></mrow></munder><munder><mo>&#x2211;</mo><mrow><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>j</mi><mo>&#x2260;</mo><mi>i</mi></mrow></munder><munder><mo>&#x2211;</mo><mrow><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mrow></munder><munder><mo>&#x2211;</mo><mrow><mi>s</mi><mo>&#x2208;</mo><mi mathvariant=\"script\">R</mi></mrow></munder><mfenced><mrow><msub><mi>d</mi><mrow><mi>i</mi><mi>j</mi></mrow></msub><mfenced><mrow><mfenced><mrow><mn>1</mn><mo>-</mo><msup><mi>g</mi><mrow><mi>a</mi><mi>s</mi></mrow></msup><msup><mi>&#x3C3;</mi><mrow><mi>a</mi><mi>s</mi></mrow></msup></mrow></mfenced><msup><mi>&#x391;</mi><mi>a</mi></msup><mo>+</mo><mfenced open=\"&#x230A;\" close=\"&#x230B;\"><mrow><msup><mi>&#x3B2;</mi><mi>a</mi></msup><mfenced><mrow><mn>1</mn><mo>-</mo><mi>&#x3B3;</mi></mrow></mfenced></mrow></mfenced><mi>&#x39C;</mi></mrow></mfenced><mo>+</mo><mfenced><mrow><mn>1</mn><mo>-</mo><msup><mi>g</mi><mrow><mi>a</mi><mi>s</mi></mrow></msup><msup><mi>&#x3C4;</mi><mrow><mi>a</mi><mi>s</mi></mrow></msup></mrow></mfenced><msup><mi>&#x392;</mi><mi>a</mi></msup><mo>+</mo><mfenced open=\"&#x230A;\" close=\"&#x230B;\"><mrow><msup><mi>&#x3B2;</mi><mi>a</mi></msup><mfenced><mrow><mn>1</mn><mo>-</mo><mi>&#x3B3;</mi></mrow></mfenced></mrow></mfenced><mi>&#x39D;</mi></mrow></mfenced><msubsup><mi>v</mi><mrow><mi>i</mi><mi>j</mi></mrow><mi>a</mi></msubsup><mspace linebreak=\"newline\"/><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>&#xA0;</mo><mo>+</mo><munder><mo>&#x2211;</mo><mrow><mi>k</mi><mo>&#x2208;</mo><mi>&#x212C;</mi></mrow></munder><msub><mi>H</mi><mi>k</mi></msub><msub><mi>y</mi><mi>k</mi></msub><mo>+</mo><munder><mo>&#x2211;</mo><mrow><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mrow></munder><msup><mi>F</mi><mi>a</mi></msup><msup><mi>z</mi><mi>a</mi></msup><mo>+</mo><munder><mo>&#x2211;</mo><mrow><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mrow></munder><msup><mi>&#x3B2;</mi><mi>a</mi></msup><msup><mi>b</mi><mi>a</mi></msup><msup><mi>z</mi><mi>a</mi></msup><mo>+</mo><munder><mo>&#x2211;</mo><mrow><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mrow></munder><munder><mo>&#x2211;</mo><mrow><mi>s</mi><mo>&#x2208;</mo><mi mathvariant=\"script\">R</mi></mrow></munder><msup><mi>g</mi><mrow><mi>a</mi><mi>s</mi></mrow></msup><msup><mi>f</mi><mrow><mi>a</mi><mi>s</mi></mrow></msup><msup><mi>z</mi><mi>a</mi></msup><mo>&#xA0;</mo><mo>-</mo><mo>&#xA0;</mo><munder><mo>&#x2211;</mo><mrow><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mrow></munder><msup><mi>E</mi><mi>a</mi></msup><msup><mi>&#x3B7;</mi><mi>a</mi></msup></mstyle></math>","origin":"MathType for Microsoft Add-in"}(1)

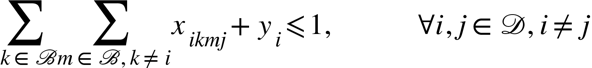
s.t.:

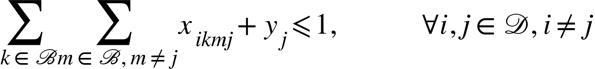
 (2)

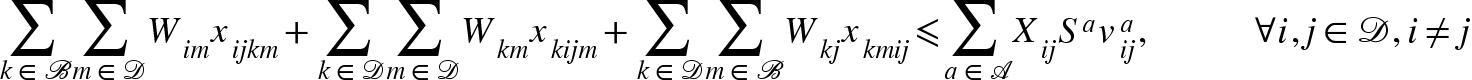
{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>x</mi><mrow><mi>i</mi><mi>k</mi><mi>m</mi><mi>j</mi></mrow></msub><mo>&#x2A7D;</mo><msub><mi>y</mi><mi>k</mi></msub><mo>,</mo><mspace width=\"2em\"/><mo>&#xA0;</mo><mo>&#x2200;</mo><mi>m</mi><mo>,</mo><mi>k</mi><mo>&#x2208;</mo><mi mathvariant=\"fraktur\">B</mi><mo>,</mo><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mo>&#xA0;</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (3)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>x</mi><mrow><mi>i</mi><mi>k</mi><mi>m</mi><mi>j</mi></mrow></msub><mo>&#x2A7D;</mo><msub><mi>y</mi><mi>m</mi></msub><mo>,</mo><mspace width=\"2em\"/><mo>&#xA0;</mo><mo>&#x2200;</mo><mi>m</mi><mo>,</mo><mi>k</mi><mo>&#x2208;</mo><mi mathvariant=\"fraktur\">B</mi><mo>,</mo><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mo>&#xA0;</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (4)

 (5)

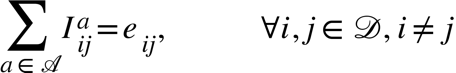
 (6)

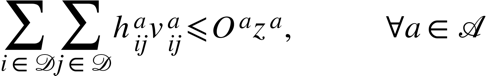
 (7)

 (8)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>x</mi><mrow><mi>k</mi><mi>k</mi><mi>m</mi><mi>m</mi></mrow></msub><mo>&#x2A7E;</mo><msub><mi>y</mi><mi>k</mi></msub><mo>+</mo><msub><mi>y</mi><mi>m</mi></msub><mo>&#x2212;</mo><mn>1</mn><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>k</mi><mo>,</mo><mi>m</mi><mo>&#x2208;</mo><mi>&#x212C;</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (9)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>e</mi><mrow><mi>i</mi><mi>j</mi></mrow></msub><mo>&#x2A7D;</mo><msub><mi>y</mi><mi>i</mi></msub><mo>+</mo><msub><mi>y</mi><mi>j</mi></msub><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (10)

 (11)

 (12)

{"mathml":"<mml:math style=\"font-family:stix;font-size:16px;\" xmlns:m=\"http://schemas.openxmlformats.org/officeDocument/2006/math\" xmlns:mml=\"http://www.w3.org/1998/Math/MathML\"><mml:mstyle mathsize=\"16px\"><mml:mi>&#x3B3;</mml:mi><mml:msup><mml:mi>&#x3B2;</mml:mi><mml:mi>&#x3B1;</mml:mi></mml:msup><mml:mo>&#x2264;</mml:mo><mml:msup><mml:mi>S</mml:mi><mml:mi>&#x3B1;</mml:mi></mml:msup><mml:mo>,</mml:mo><mml:mspace width=\"2em\"/><mml:mo>&#x2200;</mml:mo><mml:mi>a</mml:mi><mml:mo>&#x2208;</mml:mo><mml:mi class=\"mathscr\" mathvariant=\"script\">A</mml:mi></mml:mstyle></mml:math>","origin":"MathType for Microsoft Add-in"} (13)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msubsup><mi>g</mi><mrow/><mrow><mi>a</mi><mi>s</mi></mrow></msubsup><mo>&#x2208;</mo><mfenced open=\"{\" close=\"}\"><mrow><mn>0</mn><mo>,</mo><mn>1</mn></mrow></mfenced><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi><mo>,</mo><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (14)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>y</mi><mi>k</mi></msub><mo>&#x2208;</mo><mfenced open=\"{\" close=\"}\"><mrow><mn>0</mn><mo>,</mo><mn>1</mn></mrow></mfenced><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>k</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (15)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msubsup><mi>I</mi><mrow><mi>i</mi><mi>j</mi></mrow><mi>a</mi></msubsup><mo>&#x2208;</mo><mfenced open=\"{\" close=\"}\"><mrow><mn>0</mn><mo>,</mo><mn>1</mn></mrow></mfenced><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi><mo>,</mo><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (16)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>e</mi><mrow><mi>i</mi><mi>j</mi></mrow></msub><mo>&#x2208;</mo><mfenced open=\"{\" close=\"}\"><mrow><mn>0</mn><mo>,</mo><mn>1</mn></mrow></mfenced><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (17)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msub><mi>x</mi><mrow><mi>i</mi><mi>k</mi><mi>m</mi><mi>j</mi></mrow></msub><mo>&#x2A7E;</mo><mn>0</mn><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>k</mi><mo>,</mo><mi>m</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (18)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msup><mi>&#x3B2;</mi><mi>a</mi></msup><mo>&#x2208;</mo><mi mathvariant=\"normal\">&#x2115;</mi><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>&#x3B1;</mi><mo>&#x2208;</mo><mi mathvariant=\"script\">A</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (19)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msubsup><mi>v</mi><mrow><mi>i</mi><mi>j</mi></mrow><mi>a</mi></msubsup><mo>&#x2208;</mo><mi mathvariant=\"normal\">&#x2115;</mi><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>i</mi><mo>,</mo><mi>j</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">D</mi><mo>,</mo><mi>i</mi><mo>&#x2260;</mo><mi>j</mi><mo>,</mo><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (20)

{"mathml":"<math style=\"font-family:stix;font-size:16px;\" xmlns=\"http://www.w3.org/1998/Math/MathML\"><mstyle mathsize=\"16px\"><msup><mi>z</mi><mi>a</mi></msup><mo>&#x2208;</mo><mi mathvariant=\"normal\">&#x2115;</mi><mo>,</mo><mspace width=\"2em\"/><mo>&#x2200;</mo><mi>a</mi><mo>&#x2208;</mo><mi class=\"mathscr\" mathvariant=\"script\">A</mi></mstyle></math>","origin":"MathType for Microsoft Add-in"} (21)

### Model Description

#### Objective Function

Eq. (1) is the objective function which aims to minimize the airline’s operational and fleet configuration costs. The first term calculates the total transportation cost of aircraft type serving the passenger flows associated with each OD pair, while the aircraft has undergone retrofit and seat reconfiguration. The second term covers the hub setting cost for the HS network, the third term sums the fixed cost of fleet that is assigned to serve the passenger flows of all OD pairs. This cost could be the purchase cost if the aircraft is not part of the airline’s fleet, or the fixed cost of the aircraft if it’s already in the fleet. The fourth term is the cost of adding new business class seats to the aircraft. The fifth term is the cost of purchasing and installing retrofit options for the airline’s fleet. And the sixth last term, calculates the selling gains of unincorporated fleet. It should be noted here that this term should be disregarded if the planning period is short, as ceasing operations and selling the airline’s fleet would be the baseline solution for a cost minimizer objective function.

#### Constraints

Eq. (2) to (13) are the model’s constraints. Eq. (2) ensures that a total number of *p* hubs are located from the candidate set of hubs. Eq. (3) and (4) restrict passenger flows through intermediate airports that are not one of the selected hub locations. Eq. (5) makes sure that the whole demand from origin node to the destination node routed via some hub cities. Eq. (6) and (7) ensure that in the case when the origin or destination are not hub locations, the flow cannot be more than 1. Eq. (8) calculates the total amount of passenger-hours performed between any two desired nodes in the network and ensures that enough fleet capacity is allocated for that route. Eq. (9) guarantees that flow between hub locations is done only directly. Eq. (10) ensures that routes that are opened either connect a hub city to a non-hub city, or hub to hub cities, and Eq. (11) restrict allocating aircraft types for each route to only one. Eq. (12) states that the total time each aircraft type spends on its routes cannot exceed the block time that the aircraft type can provide. Eq. (13) sets the upper bound of business class seats for an aircraft type, to the available space within that type. Eq. (14) to (21) are domain constraints.

# Conclusion

This work commenced with the formulation of research questions and aims in Chapter 3, outlining three sub-questions as the focal point of the study. The subsequent exploratory data analysis (EDA) in Chapter 4 delved into various facets of market demand data, revealing patterns crucial for addressing sub-question C, which revolves around forecasting market demand. Additionally, this phase shed light on variables influencing optimal fleet composition.

In Chapter 5, our focus shifted to identifying the best approach for market demand forecasting. Through segregated algorithm training, fine-tuning, and heuristic solutions, we concluded that this multi-pronged strategy offered the most effective forecasting methodology.

Chapter 6 witnessed the culmination of our efforts, manifesting in the construction of a holistic mathematical model. This model seamlessly integrates fleet optimization, retrofit decisions, hub location selection, and seating configuration. Beyond empowering airlines to make informed decisions for heightened efficiency and profitability, the model also provided insights into sub-question A—unravelling the variables impacting an airline’s fleet composition. Notably, our findings underscored the influence of planning horizon and decisional variables on the optimum fleet, offering pragmatic insights for strategic decision-making in the aviation industry.

In addition to the key findings, our work proposes avenues for future research. We strongly advocate for a more segmented and personalized forecasting approach, emphasizing the need to categorize demand by class. Despite challenges in obtaining reliable data on passenger seat classes, we propose incorporating different class flow demand into our model, decoupling seating configuration from existing data constraints.

In conclusion, this research contributes valuable insights into forecasting methodologies and decision variables shaping optimal fleet composition. Future endeavours should explore the recommended avenues for a more nuanced understanding of market demand and seating configuration dynamics. The transparent acknowledgment of challenges and limitations enriches the credibility of our findings, paving the way for continued advancements in aviation research.

# Bibliography

Özener, O. Ö., Matoglu, M. Ö., Günes, E., Haouari, M., & Sözer, H. (2017). Solving a large-scale integrated fleet assignment and crew pairing problem. *Annals of Operations Research*, 1-24.

Adriana. (2023, 7 26). *A Deep Dive into the Passenger Capacity of a Boeing 737*. Retrieved from We In The Sky: https://flynewleaf.ca/passenger-capacity-of-a-boeing-737/

Albright, S. C., & Winston, W. L. (2020). *Business analytics - data analysis and decison making.* Boston, MA, USA: Cengage.

Ardil, C. (2020). Trainer Aircraft Selection Using Preference Analysis for Reference Ideal Solution (PARIS). *International Journal of Aerospace and Mechanical Engineering,*.

Bazargan, M., & Hartman, J. (2012). Aircraft replacement strategy: model and analysis. *Journal of Air Transport Management*, 26-29.

Berhart, C., Farahat, A., & Lohatepanont, M. (2009). Airline Fleet Assignment with Enhanced Revenue Modeling. *Operations Research*, 231-244.

Bitzan, J., & Peoples, J. (2016). A Comparative Analysis of Cost Change for Low-Cost, Full-Service, and Other Carriers in the US Airline Industry. *Research in Transportation Economics*, 25-41.

Box, G., & Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Chadwick, W. J. (2023, 2 9). *Department of Transportation.* Retrieved from US Federal Register Office: https://public-inspection.federalregister.gov/2023-03224.pdf

Chatfield, C. (2003). *The Analysis of Time Series An Introduction, Sixth Edition.* New York: Chapman and Hall/CRC. Retrieved from GitHub: https://phosgene89.github.io/sarima.html

Chowdhury, E. (2007). Low Cost Carriers: How Are They Changing the Market Dynamics of the U.S. Airline Industry? by Erfan Chowdhury. *(Unpublished master’s thesis)*.

Clarke, L. W., Hane, C. A., Johnson, E. L., & Nemhauser, G. L. (1996). Maintenance and Crew Considerations in Fleet Assignment. *Transportation Science*, 249-260.

Cornell Law School. (2019). *14 CFR § 291.45 - BTS Schedule T–100, U.S. Air Carrier Traffic and Capacity Data by Nonstop Segment and On-Flight Market.* Retrieved from Legal Information Institute: https://www.law.cornell.edu/cfr/text/14/291.45

Daft, J., & Albers, S. (2015). An empirical analysis of airline business model convergence. *Journal of Air Transport Management*, 3-11.

Daraban, B. (2012). The Low Cost Carrier Revolution Continues: Evidence From The US Airline Industry. *Journal of Business & Economics Research*.

Do, Q. H., Lo, S.-K., Chen, J.-F., Le, C.-L., & Anh, L. H. (2020). Forecasting Air Passenger Demand: A Comparison of LSTM and SARIMA. *Journal of Computer Science*.

E, D. (2020, 1 23). *Types of Airlines and Airline Business Models*. Retrieved from One Education: https://www.oneeducation.org.uk/types-of-airlines-and-business-models/

Eurocontrol. (2011, 1). *Study into the impact of the global economic crisis on airframe utilisation.* Retrieved from Eurocontrol: https://web.archive.org/web/20150606044528/https://www.eurocontrol.int/sites/default/files/content/documents/official-documents/facts-and-figures/coda-reports/study-impact-global-economic-crisis-2011.pdf

Filippone, A. (2012). *Advanced Aircraft Flight Performance.* Cambridge: Cambridge University Press.

Gomes, L. F., Fernandes, J. E., & Soares de Mello, J. C. (2012). A fuzzy stochastic approach to the multicriteria selection of an aircraft for regional chartering. *Journal of Advanced Transportation*.

Haire, A. R., & Machemehl, R. B. (2009). A Methodology for Incorporating Fuel Price Impacts into Short-term Transit Ridership Forecasts. *Engineering, Economics, Business*.

Holloway, S. (2008). *Straight and Level, Practical Airline Economis.* Hampshire, England; Burlington, USA: Ashgate Publishing.

Hsu, C. I., L. H., Liu, S. M., & Chao, C. C. (2011). Aircraft replacement scheduling: a dynamic programming approach. *Transportation Research Part E: Logistics and Transportation Review*, 41-60.

Huang, C. C. (2021). Assessing the financial performance of airlines in the Asia-Pacific region. *Investment Management and Financial Innovations*, 234-244.

IATA Sustainability and Economics, S. G. (n.d.).

International Air Transport Association (IATA). (2023). *Airlines Set to Earn 2.7% Net Profit Margin on Record Revenues in 2024.* Geneva: The International Air Transport Association (IATA).

Jacaruso, L. (2018). A method of trend forecasting for financial and geopolitical data: inferring the effects of unknown exogenous variables. *Journal of Big Data*.

Kaplan, T. (2017). *INSIGHT FROM FLIGHTGLOBAL: Mid-life aircraft trading patterns and the impact of lessors .* Retrieved from FlightGlobal: https://www.flightglobal.com/insight-from-flightglobal-mid-life-aircraft-trading-patterns-and-the-impact-of-lessors/123219.article

Khoo, H. L., & Teoh, L. E. (2014). A bi-objective dynamic programming approach for airline green fleet planning. *Transportation Research Part D: Transport and Environment*, 166-185.

*List of Low-Cost-Carriers (LCCs) based on ICAO definition.* (2017, 06 14). Retrieved from ICAO: http://www.icao.int/sustainability/Pages/GATO2030.aspx

Listes, O., & Dekker, R. (2005). A scenario aggregation-based approach for determining a robust airline fleet composition for dynamic capacity allocation. *Transportation Science*, 367-82.

Müller, C., Kieckhäfer, K., & Spengler, T. S. (2018). The influence of emission thresholds and retrofit options on airline fleet planning: An optimization approach. *Energy Policy*, 242-257.

Mohri, S. S., Nasrollahi, M., Pirayesh, A., & Mohammadi, M. (2022). An integrated global airline hub network design with fleet planning. *Computers & Industrial Engineering*.

New, C. C. (1975). TRANSPORT FLEET PLANNING FOR MULTI-PERIOD OPERATIONS. *Operational Research Quarterly*, 151-166.

Oliveira, A. V., Caliari, T., & Narcizo, R. R. (2022). An empirical model of fleet modernization: On the relationship between market concentration and innovation adoption by airlines. *Research in Transportation Business & Management*.

Oliveira, A. V., Narcizo, R. R., Caliari, T., Morales, M. A., & Prado, R. (2021). Estimating fuel-efficiency while accounting for dynamic fleet management: Testing the effects of fuel price signals and fleet rollover. *Transportation Research Part D: Transport and Environment*.

Oum, T., Zhang, A., & Zhang, Y. (2000). Optimal demand for operating lease of aircraft. *Transportation Research Part B*, 17-29.

Ozdemir, M. S., Basligil, H., & Karaca, M. (2011). Aircraft Selection Using Analytic Network Process: A Case for Turkish Airlines. *Proceedings of the World Congress on Engineering*.

Park, K. (2013, 11 1). Why the Longest Nonstop Flights Are Ending. *Bloomberg*.

Powell, W. B., & Carvalho, T. A. (1997). Dynamic control of multicommodity fleet management problems. *European Journal of Operational Research*, 522-541.

Procurement Resource. (2022). *Kerosene Price Trends and Forecast*. Retrieved from Procurement Resource, Insights That Matter: https://www.procurementresource.com/resource-center/kerosene-price-trends

R, K. (2023, 08 19). *Pulse*. Retrieved from LinkedIn: https://www.linkedin.com/pulse/low-cost-carriers-business-model-kannan-r

Richter, F. (2023). *Liftoff: Airline Industry Returns to Profits in 2023.* Statista.

Rosskopf, M., Lehner, S., & Gollnick, V. (2014). Economic-environmental trade-offs in long-term airline fleet planning. *Journal of Air Transport Management*, 109-115.

Sa, C. A. (2016, 05 27). *Robust fleet planning under stochastic demand (Master's Thesis).* Retrieved from TU Delft Repoistory: http://resolver.tudelft.nl/uuid:13fd72f6-946b-4fc1-bcfd-ee1d737abe85

Sa, C. A., Santos, F. B., & Clarke, J.-P. B. (2019). Portfolio-based airline fleet planning under stochastic demand. *Omega*.

Schick, G., & Stroup, J. (1981). Experience with a multi-year fleet planning model. *Omega*, 389-396.

Shube, D., & Stroup, J. (1975). Fleet Planning Model. *Winter Computer Simulation Conference Proceedings* (pp. 45-50). Society for Computer Simulation.

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. *IEEE International Conference on Machine Learning and Applications.* Orlando, FL, USA: IEEE.

Skalovskaia, E. (2019). *Potential of air transport in Cabo Verde focusing on hub-and-spoke model.* Worms: BSc Thesis.

Spiewanowski, P. (2015). Digital Natives: LCCs still rule in online engagement. *RPubs by RStudio*.

Subramanian, R., Scheff, R. P., Quillinan, J. D., Wiper, D., & Marsten, R. E. (1994). Cold-start: Fleet Assignment at Delta Air Lines. *Interfaces*, 104-120.

Suissa, A. (2010, 7 28). *Number 295 - Passengers Transported on All-Cargo Aircraft And Nonscheduled Passengers Transported on Scheduled Service.* Retrieved from US Department of Transportation: https://www.transportation.gov/regulations/guidance/number-295-passengers-transported-all-cargo-aircraft-and-nonscheduled

Teoh, L. E., & Khoo, H. L. (2015). Airline Strategic Fleet Planning Framework. *Journal of the Eastern Asia Society for Transportation Studies*.

United States Government Accountability Office . (2014, 6). *The Average Number of Competitors in Markets Serving the Majority of Passengers Has Changed Little in Recent Years, but Stakeholders Voice Concerns about Competition.* Retrieved from AIRLINE COMPETITION: https://www.gao.gov/assets/gao-14-515.pdf

US Department of Transportation. (2023, 12). *Air Carriers : T-100 Domestic Market (All Carriers)* . Retrieved from Bureau of Transportation Statistics: https://www.transtats.bts.gov/DL\_SelectFields.aspx?gnoyr\_VQ=GED&QO\_fu146\_anzr=Nv4+Pn44vr45

US Department of Transportation. (2023, 12 14). *Data Bank 28DM-T-100*. Retrieved from Bureau of Transportation Statistics: https://www.bts.gov/browse-statistical-products-and-data/bts-publications/data-bank-28dm-t-100-domestic-market-data

US Department of Transportation. (2024). *Air Carrier Statistics*. Retrieved from Bureau of Transportation Statistics: https://www.transtats.bts.gov/DatabaseInfo.asp?QO\_VQ=EEE&Yv0x=D

US Energy Information and Administration. (2023, 12). *US Gulf Coast Kerosene-Type Jet Fuel Spot Price FOB*. Retrieved from EIA: https://www.eia.gov/dnav/pet/hist/eer\_epjk\_pf4\_rgc\_dpgD.htm

Wikipedia. (2023, 12 24). *List of low-cost airlines* . Retrieved from Wikipedia: https://en.wikipedia.org/wiki/List\_of\_low-cost\_airlines

Wu, J., Zhang, P.-w., Wang, Y., & Shi, J. (. (2022). Integrated aviation model and metaheuristic algorithm for hub-and-spoke network design and airline fleet planning. *Transportation Research Part E*.

Yu, W., Hong, S., & Peiwen, Z. (2012). Aircraft trip DOC parameters: A function of stage length, seat capacity and design range. *2012 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 2322-2325). Hong Kong, China: IEEE.

Zhang, R., Gue, Z., Meng, Y., Wang, S., Li, S., Niu, R., . . . Guo, Q. (2021). Comparison of ARIMA and LSTM in Forecasting the Incidence of HFMD Combined and Uncombined with Exogenous Meteorological Variables in Ningbo, China. *International Journal of Environmental Research and Public Health*.

# Appendices

A graph of a number of cities

Description automatically generated with medium confidence

Figure Pandemic Effects on Top 10 Routes

A screenshot of a graph

Description automatically generated

Figure Top 5 Airlines of Each Model

A screenshot of a graph

Description automatically generated

Figure Flight Haul Distribution in 1990 (Up) and in 2023(Bottom)