

**Stimulating the Skies: Forecasting Market Response to Capacity Expansion in
International Air Travel Using Deep Learning**

Matias Jose Montero

A Thesis

In The Department

of

John Molson School of Business

Present in Partial Fulfillment of the Requirements

for the Degree of

Master of Business Analytics and Technology Management at

Concordia University

Montréal, Québec, Canada

December 2025

© Matias Jose Montero, 2025

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared,

By: Matias Jose Montero

Entitled:

“Stimulating the Skies: Forecasting Market Response to Capacity Expansion in International Air
Travel Using Deep Learning”

and submitted in partial fulfillment of the requirements for the degree of

Master of Business Analytics and Technology Management

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair

Pr. Dongliang Sheng

_____ Examiner

Pr. Dongliang Sheng

_____ Examiner

Pr. Danielle Morin

_____ Supervisor

Pr. Salim Lahmiri

Approved by

Dr. Rustam Vahidov Chair of Department

Dr. Anne-Marie Croteau Dean of John Molson School of Business

Abstract

Stimulating the Skies: Forecasting Market Response to Capacity Expansion in International Air Travel Using Deep Learning

Matias Jose Montero

Accurate air travel demand forecasting is critical for airline profitability, yet predicting the market stimulation resulting from new capacity additions remains a significant challenge. This research addresses this gap by developing and evaluating a series of deep learning models to forecast incremental passenger demand following capacity expansion. Utilizing a longitudinal dataset of Canadian international Origin-Destination (O-D) pairs from 2014 to 2024, this study integrates passenger flow data with airline schedule information, including engineered features to capture capacity growth and new route introductions. Four distinct neural network architectures were implemented and compared: a baseline Artificial Neural Network (ANN), a 1D Convolutional Neural Network (CNN), a Long Short-Term Memory (LSTM) network, and a Gated Recurrent Unit (GRU) network. The GRU model achieved the lowest Mean Absolute Percentage Error (MAPE) at 5.23%, outperforming all other models. In contrast, the ANN and CNN models produced higher error rates, indicating limited ability to capture temporal dependencies. The GRU maintained consistent performance even during the volatile 2020–2024 period, suggesting its suitability for forecasting under disrupted market conditions. These results support the use of recurrent architectures for modeling demand stimulation and provide a data-driven framework for route planning, revenue optimization, and resource allocation in the airline industry.

Key Words: Deep Learning, Market Stimulation, Air Travel, Predictive Analytics

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Salim Lahmiri, for his kindness, attentiveness, and invaluable guidance throughout this research. His mastery in the field of machine learning has been a tremendous source of insight and inspiration, and his support has greatly enriched the quality of this work.

I am profoundly thankful to my parents, Benjamina Castellanos and Rafael Montero, and to my brothers, Facundo and Nicolás Montero, for their unwavering love, encouragement, and belief in me. I am also deeply grateful to my aunts, Teresa and Ercilia and my whole family. Their support has been the foundation upon which I could pursue this academic journey.

I would also like to express my profound gratitude and heartfelt appreciation to my girlfriend and future wife, Marie-Sophie, for being my unwavering partner and companion through every challenge. Moving to a new country and building a life while studying and working is no easy feat, and I never could have done it without her love, support, and understanding.

Finally, I wish to acknowledge my friends in Montreal, whose companionship and encouragement made this experience more meaningful. I am equally grateful to my friends in Argentina, who have always cheered me on from afar. A special mention goes to my friend Julian Carrera, with whom I once dreamt of going abroad to study—thank you for sharing that vision and inspiring me to make it a reality. I am also deeply grateful to my friend and classmate Pierre, whose friendship and support as a study partner have been invaluable throughout this journey.

Table of Contents

List of Figures	vii
List of Tables.....	viii
Acronyms.....	ix
1. Introduction.....	1
1.1.1 Network Planning	4
1.1.2 Origin-Destination Flows and Demand Forecasts.	4
1.1.3 Target Variable	4
1.1.4 Structure of Thesis.	4
2. Literature Review.....	6
2.1.1 The need for accurate demand forecasting in the modern commercial aviation industry. 6	
2.1.2 Regression-based Econometric Models and Time Series Models for Predicting Air Travel Demand.	8
2.1.3 Machine Learning and Deep Learning Models for Predicting Air Travel Demand.	9
3. Research Methodology	23
3.1.1 Data Sources and Key Variables	23
3.1.2 Passenger Flow Data (Demand-Side)	24
3.1.3 Schedule and Capacity Data (Supply-Side).....	24
3.1.4 Final Dataset Characteristics and Variables	25
4. Research Objectives and Contributions	27
4.1 Data Collection	27
4.2 Data integration and Preprocessing	27
4.3 Data Transformation and Partitioning.....	28
4.4 Feature Engineering	29
4.5 Forecasting Models & Hyper-Parameter Tuning.....	29
4.6 Deep Learning Model Architectures	30
4.6.1 Artificial Neural Network (ANN).....	30
4.6.2 Long Short-Term Memory (LSTM) Network.....	31
4.6.3 Gated Recurrent Unit (GRU) Network	32
4.6.4 1D Convolutional Neural Network (CNN).....	33
4.7 Evaluation Metrics	34

5.	Results.....	35
5.1	Quantitative Model Performance.....	35
5.2	Detailed Performance Analysis of Individual Models.....	37
5.2.1	Artificial Neural Network (ANN).....	37
5.2.2	1D Convolutional Neural Network (CNN).....	37
5.2.3	Long Short-Term Memory (LSTM) Network.....	37
5.2.4	Gated Recurrent Unit (GRU) Network.....	37
5.3	Comparative Analysis and Benchmarking.....	38
6.	Findings and Discussion.....	43
7.	References.....	50

List of Figures

Figure 1. Research Outcomes	3
Figure 2. ANN Structure	30
Figure 3. LSTM Structure	31
Figure 4. GRU Structure	33
Figure 5. CNN Architecture.....	34

List of Tables

Table 1. Literature Review Summary	22
Table 2. Description of Key Variables in the Final Dataset.....	26
Table 3. Performance Metrics of Neural Network Models	36
Table 4. Similar Travel Demand Forecasting Papers Summary	39
Table 5. Performance Benchmark Against Other Research Papers	40

Acronyms

ANN	Artificial Neural Network
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
O-D	Origin-Destination
ASM	Available Seat Miles
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
R ²	Coefficient of Determination
DDS	Direct Data Solutions
SSIM	Standard Schedules Information Manual
GDP	Gross Domestic Product
ROI	Return on Investment

1. Introduction

International travel has long been a transformative force in global connectivity and economic development. The evolution of this industry, from the era of ocean liners in the early 19th century to the modern network of intercontinental flights, has not only redefined travel but has also prompted significant technological, operational, and strategic advancements (Belobaba, Odoni, & Barnhart, 2015). As technological innovations such as highly fuel-efficient jet engines and pressurized cabins emerged in the mid-20th century, airlines were able to dramatically lower operating costs and expand service offerings, thus making air travel accessible to a broader population. This transition laid the foundation for the competitive and dynamic environment that is today's global aviation market.

Despite the maturity of the industry and the proliferation of forecasting models for air travel demand, a critical and relatively underexplored research question remains: How does the market respond when new capacity is introduced between two points in the map? In practical terms, when an airline adds a new flight (say, on a route connecting Montreal and Naples), there is often a stimulation in demand that goes beyond merely accommodating previously unmet needs. This phenomenon, referred to as market stimulation, encapsulates the latent demand that is unlocked when additional capacity alleviates existing constraints, thereby encouraging travelers who were previously unable or unwilling to fly to now consider air travel.

Capacity refers to the total number of seats and flights offered by airlines on a given route (measured by Seats, Flights, and Available Seat Miles (ASMs)). The forecasting problem centers on predicting how changes in capacity (especially new route introductions) stimulate additional passenger demand.

Forecasting market stimulation due to capacity expansion is essential for airline network planning and revenue management. The introduction of a new flight can trigger a cascade of effects: it may lower the effective seat scarcity, induce competitive fare adjustments, and create a perception of increased availability that encourages bookings from both leisure and business segments. Furthermore, the market response is not solely determined by the increased capacity but is influenced by a myriad of factors including fare elasticity, consumer behavior, network connectivity, and external economic conditions. For example, if a new Montreal-Naples flight is introduced, its success in stimulating additional demand may depend on how sensitive passengers are to fare changes and on whether the new service effectively captures the latent, unmet demand from potential passengers who previously faced capacity limitations.

Recent advances in forecasting methodologies, particularly in the realms of deep learning and hybrid modeling, offer promising tools to capture this complex interplay. Deep learning architectures, such as Convolutional Long Short-Term Memory (ConvLSTM) networks, have demonstrated superior abilities in modeling non-linear, dynamic, and spatial-temporal relationships within data (Shi et al., 2015). When integrated with econometric approaches that provide interpretability through elasticity measures (Kölker, Lütjens, & Gollnick, 2024), a hybrid

model can potentially forecast not only the absolute level of demand but also the incremental effect of capacity additions. This dual approach enables the identification of both the baseline demand and the stimulated demand that emerges as a direct consequence of new capacity entering the market.

The significance of this research extends beyond academic curiosity. In an era characterized by rapid technological evolution and increasing competition, airlines are under constant pressure to optimize network capacity, pricing, and service strategies. Accurate forecasting of market stimulation can serve as a critical decision support tool, enabling network planning managers to design optimal schedules, adjust fare structures, and implement targeted marketing strategies that maximize revenue while capturing untapped demand. Moreover, by quantifying the demand response to capacity expansion, airlines can better evaluate the return on investment (ROI) for new routes and aircraft acquisitions, ensuring that capital-intensive decisions are backed by robust predictive insights.

Although substantial literature exists on air travel demand forecasting using traditional time series methods and econometric models (Wang & Gao, 2021; Zachariah, Sharma, & Kumar, 2023), the specific phenomenon of market stimulation triggered by capacity increases has not received equivalent attention. Most conventional models focus on capturing historical trends and cyclical variations in passenger flows without adequately addressing the dynamic response induced by additional capacity. This research aims to fill that gap by developing a forecasting framework that explicitly models the incremental demand generated when new flights are added.

The value of this research extends beyond the immediate interests of airlines and network planning teams. By enabling more accurate forecasting of market stimulation following capacity expansion, the deep learning framework developed here supports strategic decision-making in route planning, fleet allocation, and investment, while also enhancing revenue management through improved demand anticipation and dynamic pricing. Airports, too, stand to benefit from these insights, as they can better anticipate passenger flows, optimize resource allocation, and plan infrastructure investments in response to changing market conditions. The ability to understand and forecast demand stimulation allows airports to coordinate more effectively with airlines, manage congestion, and improve the overall passenger experience.

Policy makers and regulators are provided with robust demand forecasts that inform market access policies, support the evaluation of new routes, and contribute to monitoring the health of the industry. Local governments and tourism boards can leverage demand forecasting to guide economic development strategies, anticipate visitor influx, and make informed decisions about infrastructure and promotional efforts. Insights into market stimulation also enable these stakeholders to collaborate with airlines in launching new routes that maximize regional benefits.

Investors and financial analysts gain a transparent methodology for assessing the risks and opportunities associated with airline and airport expansion projects. The research provides a replicable framework for benchmarking performance and evaluating market potential across

different regions and time periods, thereby supporting more informed investment decisions. Passengers and consumer advocacy groups, meanwhile, benefit indirectly from the improved matching of capacity and demand, which translates into more flight options, competitive fares, and enhanced travel experiences. Reliable demand forecasts help prevent situations of overcapacity or undercapacity, reducing the likelihood of cancellations, delays, or price spikes and ultimately protecting consumer interests.

The current work establishes a baseline for future comparative studies and contributes to the ongoing evolution of predictive analytics in aviation.

The research employs a hybrid approach that synergizes the strengths of deep learning techniques and econometric analysis. On one hand, deep neural networks will be utilized to extract complex, non-linear patterns from large-scale datasets, while on the other hand, econometric models will provide insight into the causal relationships between fare levels, capacity utilization, and passenger demand. This combined methodology is expected to yield forecasts that are not only highly accurate but also interpretable and actionable. The insights derived from this model will help airline decision-makers to understand how market demand shifts in response to capacity changes, thereby enabling more informed decisions regarding network expansion and resource allocation. Figure 1 illustrates the hybrid methodological framework employed in this research. The process is structured into three sequential stages: pattern extraction, causal analysis, and forecast generation. In the first stage, deep learning techniques are applied to large-scale datasets to uncover complex, non-linear patterns. The second stage involves econometric analysis to identify and interpret causal relationships between fare levels, capacity utilization, and passenger demand. Finally, the third stage consists in the integration of these insights to generate forecasts that are both accurate and actionable.

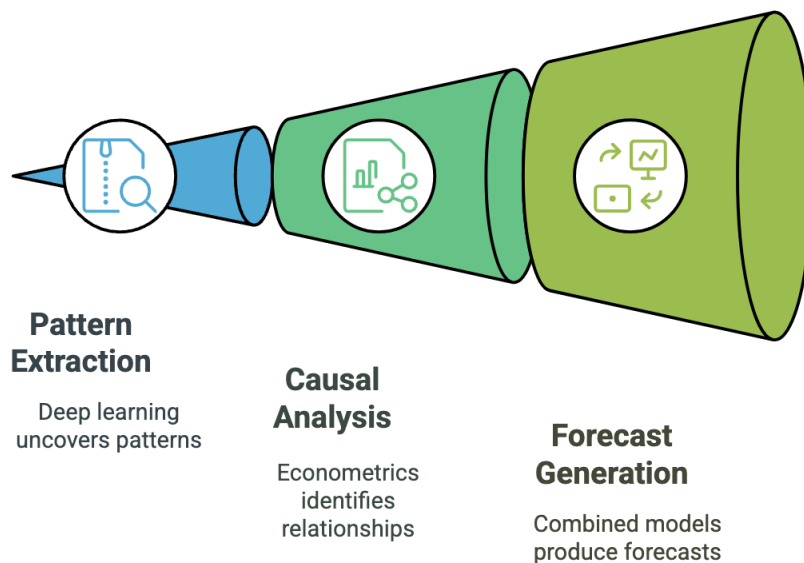


Figure 1. Research Outcomes

This study addresses a crucial and under-researched aspect of air travel demand forecasting: predicting the market stimulation resulting from new capacity additions. By leveraging state-of-the-art deep learning methods alongside established econometric models, the research seeks to offer a tool that can forecast incremental demand and provide actionable insights for airline network planning. This work aims to contribute to the literature by advancing forecasting techniques in the aviation industry and offering practical implications for managing capacity and stimulating market growth in an increasingly competitive and dynamic environment.

1.1.1 Network Planning.

Network planning is the strategic backbone of an airline, dictating where it flies, how frequently, and with what aircraft. This complex process involves a hierarchical set of decisions, including route network design, flight scheduling, and fleet assignment, all aimed at maximizing profitability and market share. The central challenge lies in matching an airline's finite capacity (its aircraft and crew) with fluctuating passenger demand across thousands of potential origin-destination markets. Decisions made in network planning have long-term structural implications, influencing operational costs, revenue potential, and competitive positioning. As this research addresses the market's response to capacity additions, it directly informs the most critical output of network planning: the flight schedule. By accurately forecasting the stimulation effect of adding a new flight, network planners can more effectively evaluate the viability of new routes, optimize frequencies on existing routes, and ultimately design a network that is both resilient and profitable.

1.1.2 Origin-Destination Flows and Demand Forecasts.

Air travel demand refers to the demand to travel between a point of origin and a final destination (O-D). Passengers may be served by a single non-stop flight or by a complex itinerary involving one or more connections. The total flow of passengers between an O-D pair represents the true market demand. Consequently, airline forecasting cannot focus merely on individual flight legs but must consider the entire network of O-D flows. Demand forecasts are the primary input for nearly all commercial planning functions, from the strategic decisions in network planning to the tactical execution in revenue management and pricing. The accuracy of these forecasts is paramount. An overestimation of demand leads to empty seats and lost revenue, while an underestimation results in missed opportunities and potential market share loss to competitors. This thesis focuses on a specific, yet critical, component of O-D demand forecasting: quantifying the incremental demand stimulated by a capacity increase, thereby providing a more accurate and dynamic picture of the true market potential.

1.1.3 Target Variable

The variable to be predicted is Sum of Total Market Pax: the total number of passengers traveling on a specific non-directional Origin-Destination (O-D) pair within a given month.

1.1.4 Structure of Thesis.

This thesis is structured into six chapters to systematically address the research question.

- Chapter 1 has introduced the research problem, highlighting the significance of forecasting market stimulation in the airline industry and outlining the core objectives and contributions of this work.
- Chapter 2 provides a review of the relevant literature, covering the evolution of air travel demand forecasting from traditional econometric and time series models to modern machine learning and hybrid approaches. It identifies the existing research gap regarding the specific phenomenon of demand stimulation from capacity expansion.
- Chapter 3 describes the methodology employed in this study. It details the data sources, including historical passenger booking data and external indicators, and explains the architecture of the proposed hybrid forecasting framework, which integrates deep learning (ConvLSTM) with econometric principles.
- Chapter 4 presents the empirical results of the forecasting model. It evaluates the model's performance against several benchmark models using key accuracy metrics and analyzes its ability to predict both baseline and stimulated demand on the selected route.
- Chapter 5 discusses the key findings and their implications. It interprets the model's outputs in the context of airline network planning and revenue management, offering actionable insights for strategic decision-making. The limitations of the study are also candidly discussed.
- Chapter 6 concludes the thesis by summarizing the research, reiterating the main contributions, and suggesting promising directions for future research in the domain of air travel demand forecasting.

2. Literature Review.

2.1.1 The need for accurate demand forecasting in the modern commercial aviation industry.

The forecasting of air travel demand has long been recognized as a critical element in the commercial aviation industry, influencing strategic planning, fleet allocation, network design, revenue management, and overall operational efficiency. Accurate demand forecasts enable airlines to optimize pricing, manage capacity, reduce underutilization or overcapacity risks, and ultimately improve profitability in an industry historically characterized by high capital intensity and cyclical performance. This review synthesizes the evolution of forecasting techniques—from early econometric models and time series methods to contemporary machine learning and hybrid approaches—highlighting their theoretical underpinnings, empirical applications, strengths, and limitations in the context of modern commercial aviation.

Historically, the rapid growth of the airline industry was driven primarily by technological advancements, such as the introduction of jet aircraft in the 1950s and the emergence of wide-body jets in the 1970s. During these periods, government policies and regulatory frameworks favored technological progress, often at the expense of profitability and competitive market dynamics (Belobaba, Odoni, & Barnhart, 2015). However, the landscape of air travel demand changed significantly following the deregulation of the airline industry in the United States in 1978. Deregulation ushered in an era of cost efficiency, enhanced competition, and market-driven decision making. These shifts have subsequently spread to many industrialized nations, resulting in a highly competitive international airline market where profit margins are thin and external shocks can have drastic effects on demand patterns (Belobaba et al., 2015).

In the decades following deregulation, further technological progress and globalization of the industry prompted rapid increases in both passenger numbers and flight frequencies, even as airlines continued to grapple with cyclical profitability and heightened vulnerability to disruptive events. The impacts of geopolitical tensions, global warming, natural disasters, and pandemics have underscored the need for more adaptive and accurate forecasting models. For instance, the severe downturns following the September 11, 2001 attacks, the global recession of 2008–2010, and the unprecedented challenges posed by the COVID-19 pandemic have all highlighted the fragility of the aviation sector in the face of sudden demand shocks (Biolini et al., 2021; Truong, 2021). Consequently, accurate demand forecasting has emerged not only as a tool for operational optimization but also as a strategic imperative to mitigate risks in an increasingly volatile global environment (Alam et al., 2024).

At the heart of demand forecasting is the need to understand and predict the willingness of passengers to travel between origin and destination pairs, a process that involves evaluating multiple itineraries differentiated by factors such as the number of connecting flights, total travel time, departure and arrival times, and fare levels (Kölker, Lütjens, & Gollnick, 2024). This

conceptualization of air travel demand recognizes the interaction between two networks: the structural network, defined by the flight options offered by airlines based on profit maximization strategies, and the global functional passenger network, which reflects the actual flow of passengers driven by their preferences and constraints (Kölker et al., 2024). The interplay between these networks underscores the complexity of forecasting demand, as changes in service offerings, such as the introduction of new routes, alterations in flight frequency, or variations in in-flight service quality, can have significant repercussions on passenger flows.

Over the past decade, a significant body of research has focused on systematically reviewing the factors influencing air travel demand and evaluating the performance of various forecasting methodologies. For instance, Wang and Gao (2021) conducted a comprehensive review of 87 studies published between 2010 and 2020, finding that most research concentrated on national-level demand using time series data linked to socio-economic and operational factors. Their review highlighted that time series methods were predominant, and that the most frequently mentioned causal factors influencing demand were GDP, income levels, market size (often proxied by population), and price indicators. Similarly, Zachariah, Sharma, and Kumar (2023) synthesized findings from a range of studies, identifying additional key drivers such as geographical location, service quality, customer preferences, tourism trends, seasonality, and the influence of low-cost carriers. Their work also emphasized that significant disruptive events, ranging from system failures and natural disasters to major political and economic shocks, play a crucial role in shaping demand patterns (Zachariah et al., 2023).

A wide range of studies have identified several key attributes that significantly influence passenger choices in air travel. These attributes include ticket price, the number of connecting flights, and total travel time, including connecting time (Adler et al., 2005; Coldren & Koppelman, 2005; Hess, 2007; Hsiao & Hansen, 2011; Theis et al., 2006). Other important factors are the time of day, weekdays, service type, size and type of aircraft, and various carrier-specific aspects (Barnhart et al., 2014; Coldren et al., 2003; Coldren & Koppelman, 2005; Grammig et al., 2005; Koppelman et al., 2008; Wei & Hansen, 2005). These attributes collectively shape passenger preferences and, consequently, the overall passenger flows within the air transportation network (Kölker, Lütjens, & Gollnick, 2024).

Kölker et al. (2024) calculated elasticities of market share for these attributes and introduced the concept of the value of travel time (VOTT) and the quality of travel index (QTI). The VOTT measures passengers' willingness to pay for reduced travel time, while the QTI assesses the overall utility of itineraries from the passenger perspective.

2.1.2 Regression-based Econometric Models and Time Series Models for Predicting Air Travel Demand.

Given the dynamic nature of the aviation industry, various forecasting techniques have been developed and refined over time. Traditional methods include econometric regression models and time series approaches. Econometric models attempt to capture the causal relationships between air travel demand and economic variables such as GDP, income per capita, and population size. One widely used class of econometric models is the gravity model, which estimates passenger volumes based on the “attraction” between city-pairs. Gravity models express the passenger volume as a function of socio-economic, demographic and geographic attraction parameters associated with origin and destination cities (Tillmann, Joormann, & Ammann, 2023).

Grosche, Rothlauf, and Heinzl (2007) demonstrated that gravity models using geo-economic variables, such as the income and population of metropolitan areas, could effectively estimate air passenger volumes, even in markets where detailed service-related data were unavailable. Hazledine (2017) extended this approach by incorporating endogenous factors such as flight frequency and service quality, achieving high explanatory power (with R^2 values around 0.76–0.77) in forecasting cross-border and domestic air travel. Additional innovations in gravity modeling include the incorporation of ethnic links, employment in knowledge-intensive sectors, and the presence of low-cost carriers, all of which have been shown to significantly stimulate demand (Boonekamp, Zuidberg, & Burghouwt, 2018).

Zhou et al. (2018) used gravity models with different airport catchment areas to study their predicting power for air passenger demand, incorporating Poisson pseudo-maximum likelihood estimation (PPML) into the model. They included factors like distance, airfare, population, tourism, and the mining sector in their model. This study contributed to predicting air travel demand by highlighting the importance of accurately defining catchment areas and considering various socio-economic factors.

Carmona-Benítez, Nieto, and Miranda (2017) introduced an Econometric Dynamic Model (EDM) to estimate passenger demand in the air transport industry. Their model used socio-economic indicators to predict demand in Mexican states and employed the Arellano-Bover method to address issues of endogeneity and autocorrelation. The results showed that passenger demand is highly elastic with respect to the number of flights, followed by economic activity and consumer prices. This research contributed to predicting air travel demand by providing a robust econometric model that incorporates various economic indicators and addresses potential statistical issues (Carmona-Benítez et al., 2017).

Gelhausen, Berster, and Wilken (2018) introduced a direct demand model for long-term forecasting of air passenger and flight volumes at German airports, addressing limitations of the traditional four-step approach. Unlike data-intensive models that rely on detailed sub-models for trip generation, distribution, and mode choice, their framework employs co-integration theory to capture long-run equilibrium relationships between macroeconomic indicators and air traffic

variables. The model integrates EU GDP growth, passengers-per-flight trends, and calibrated demand shocks (e.g., German reunification, 9/11, and the 2004 EU enlargement) within a parsimonious econometric specification validated with out-of-sample testing. Results indicate a GDP elasticity of approximately 1.31 and low out-of-sample error, offering a practical alternative for strategic planning that reduces data requirements while maintaining robustness—particularly useful for long-term scenario analysis and policy evaluation (Gelhausen et al. 2018)

Despite their strengths, econometric models face several challenges. They are particularly effective for long-term forecasting but can suffer from issues like endogeneity, reverse causality, and an inability to capture complex, nonlinear interactions. These limitations have paved the way for the use of time series forecasting methods, which rely on historical data to extrapolate trends into the future. Techniques such as Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), and Seasonal ARIMA (SARIMA) have been widely applied in the field. Wang and Gao (2021) note that time series methods are especially useful for short-term forecasting as they can effectively model seasonality and cyclic variations inherent in aviation data. However, classical time series models often struggle to account for sudden structural breaks or irregularities, and those limitations have become increasingly apparent during disruptive events such as the COVID-19 pandemic (Zachariah et al., 2023). Dantas et al. (2017) addressed some of these shortcomings by applying a hybrid approach that combines bootstrap aggregating (Bagging) with the Holt-Winters method. By decomposing the time series into seasonal, trend, and remainder components using STL (Seasonal-Trend decomposition with Loess), and then aggregating forecasts from resampled versions of the remainder component, their method achieved substantial improvements in forecast accuracy, reducing the symmetric mean absolute percentage error (sMAPE) by 31% to 43% relative to traditional methods.

2.1.3 Machine Learning and Deep Learning Models for Predicting Air Travel Demand.

The limitations of both econometric and classical time series models have derived in the development of more advanced methodologies based on machine learning and deep learning. Machine learning models are particularly appealing for air travel demand forecasting because of their ability to detect complex nonlinear patterns in large datasets. These models are typically categorized into supervised and unsupervised learning methods, with supervised learning being the most prevalent in forecasting applications. In supervised learning, models are trained on historical data where input features (e.g., past demand, economic indicators, flight frequencies) are mapped to output labels (e.g., passenger numbers). The training process involves data cleaning, feature selection, model selection, iterative training using algorithms such as gradient descent, and rigorous evaluation on test datasets (Lingrui & Xin, 2024). Deep learning models, which are an extension of machine learning, use multi-layer architectures to capture hierarchical representations of data, often leading to improved accuracy in prediction tasks.

Recent studies have demonstrated the superiority of machine learning approaches over traditional methods in predicting air travel demand. Koç and Arslan (2018), for example, applied artificial neural networks (ANNs) to forecast domestic air travel demand in Turkey, achieving high levels of accuracy with Pearson's correlation coefficients ranging from 0.79 to 0.93. Similarly, Ramadhani et al. (2020) compared the performance of a backpropagation neural network (BPNN) against an ARIMA model for forecasting monthly passenger data on key routes in Indonesia, finding that the neural network significantly outperformed the ARIMA model, with markedly lower mean absolute percentage errors (MAPE). Furthermore, Li Long et al. (2021) employed Neural Granger Causality models to integrate Google Trends data into forecasting models for Singapore Changi Airport, achieving an impressive R^2 value of 0.919. These studies underscore the potential of machine learning and deep learning models to not only improve prediction accuracy but also to accommodate the nonlinearities and high-dimensional interactions inherent in air travel demand data.

Building on the success of standalone machine learning and deep learning models, researchers have increasingly explored comparative approaches to further enhance forecasting accuracy. Kanavos et al. (2021) conducted a comparative analysis of classical time series models (ARIMA, SARIMA) and deep learning neural networks (DLNN) for forecasting aviation demand. Using a dataset of 125,263 records from 48 U.S. airports (2005–2018), their study indicated that while SARIMA improves upon ARIMA for seasonal data, the deep learning network, optimized with four hidden layers of 512 nodes each, ReLU activation, and Adam optimizer, achieves the best performance. This study demonstrated the superior predictive capability of deep architectures for capturing nonlinear and seasonal patterns in aviation demand, outperforming traditional time series approaches.

Beyond individual ML and DL models, ensemble approaches have gained traction for their ability to synthesize multiple forecasting perspectives and reduce model-selection risk.

In this context, Hu (2023) proposed an innovative combination forecasting method using a nonadditive Choquet fuzzy integral. This technique aggregates predictions from four univariate grey models (GM(1,1), NGBM(1,1), FGM(1,1), and FNGBM(1,1)) into a composite forecast known as NCMG. Applied to monthly passenger flow data from six Taiwanese airports (2013-2019), Hu's model consistently outperformed both linear (e.g., simple average, variance-covariance) and nonlinear (e.g., MLP, RBFN, SVR) combination methods across MAPE and RMSE metrics. Statistical validation using Friedman/Nemenyi and Diebold-Mariano tests further confirmed NCMG's robustness and its ability to capture interactions among constituent forecasts.

This study exemplifies the growing emphasis on ensemble learning in air travel demand forecasting, where combining diverse models can enhance accuracy and mitigate the risks associated with relying on a single forecasting technique. It also aligns with broader trends in

hybrid modeling, which seek to balance interpretability and predictive power by integrating traditional statistical methods with advanced AI algorithms.

Hybrid models represent the next evolutionary step in forecasting methodologies, as they seek to combine the interpretability and causal insights of traditional models with the pattern recognition and high predictive accuracy of machine learning techniques. Hybrid approaches can take various forms. One common strategy involves decomposing the original time series into components using techniques such as variational mode decomposition (VMD), and then modeling each component separately using different methods, for example, employing ARMA for stationary components and kernel extreme learning machines (KELM) for non-stationary components (Jin et al., 2020). The individual forecasts are then integrated to produce a final demand forecast. This approach has been shown to reduce forecast errors and improve robustness, particularly in the context of short-term forecasting where data volatility is high.

An example of a hybrid approach is Saâdaoui et al. (2020) forecasting model that coupled a feedforward neural network (FNN) with nonlinear least squares (NLS)-based regression curve fitting to predict U.S. air traffic. Using monthly data from 2000 to 2017 (available seat miles, passenger enplanements, and load factors), the approach decomposed time series into trend and seasonal components, fitted them using polynomial and sinusoidal functions, and modeled residuals with an FNN trained via backpropagation. Their results showed that seasonality extraction significantly improves accuracy, while detrending offers little benefit. The SS-FNN model achieved remarkably low MAPE values. This model outperformed traditional SARIMA and ARIMA–SVM hybrids, demonstrating the effectiveness of combining parametric decomposition with neural networks.

Expanding the scope to long-term forecasting, Emami Javanmard et al. (2024) introduced a hybrid framework that integrates pairwise ML algorithms with optimization modeling. Using Canadian air transportation data from 1970 to 2019, their approach combined models such as ARIMA, SARIMA, MIDAS, and ANN, optimized via Particle Swarm Optimization. The hybrid models significantly improved accuracy for both passenger and freight demand, highlighting the potential of optimization-enhanced ML ensembles in strategic planning.

Recent advancements in hybrid forecasting architectures have explored the integration of unconventional data sources and spatial dependencies to enhance predictive accuracy. Liang et al. (2024) proposed a novel framework that leverages spatial-temporal search engine data (STSED) combined with advanced decomposition and optimization techniques. Their approach introduces a two-stage data preprocessing strategy to mitigate noise and collinearity in STSED, followed by decomposition using the Time-Varying Filtering-based Empirical Mode Decomposition (TVF-EMD) method. The forecasting process employs a multi-model fusion strategy integrating convolutional neural networks (CNN), CNN-LSTM, and ConvLSTM models, with ensemble weights optimized via an Improved Hunger Games Search (IHGS)

algorithm applied to a Kernel-based Extreme Learning Machine (KELM) integrator. Empirical results demonstrate that this hybrid architecture significantly outperforms traditional statistical models and standalone AI models across multiple evaluation metrics, underscoring the value of incorporating spatial effects and real-time behavioral indicators into air passenger demand forecasting. This study highlights the growing importance of big data sources, such as search engine queries, in capturing latent demand signals and improving forecast robustness under dynamic market conditions.

Another promising hybrid methodology involves the integration of ANNs with meta-heuristic algorithms to optimize the training process. Mostafaeipour, Goli, and Qolipour (2018) developed a hybrid ANN model enhanced with Bat and Firefly algorithms to predict air travel demand in Iran. The Bat Algorithm, inspired by the echolocation behavior of bats, and the Firefly Algorithm, inspired by the flashing behavior of fireflies, serve as optimization tools that help fine-tune the weight parameters of the neural network. By incorporating these meta-heuristic algorithms, the hybrid model achieved a substantial improvement in prediction accuracy, with performance metrics such as the coefficient of determination (R^2) increasing from 0.2 to nearly 0.9. This improvement illustrates the effectiveness of hybrid models in capturing the complex relationships between input variables and air travel demand.

Another important direction in hybrid modeling involves enhancing input selection and capturing interdependencies across multiple time series. These strategies help improve both the relevance of model inputs and the robustness of forecasts.

Sun et al. (2019) exemplified this approach by combining the Mean Impact Value (MIV) method with a Nonlinear Vector Auto-Regression Neural Network (NVARNN). In this model, the MIV method is used to identify and select the most significant input variables based on their impact on the output, while the NVARNN is employed to capture both linear and nonlinear relationships among multiple time series. This hybrid model outperformed traditional forecasting techniques such as SARIMA and standalone neural network models, demonstrating better level accuracy, directional accuracy, and overall robustness. The success of these hybrid methods suggests that future research should further explore the integration of alternative data sources, such as real-time social media signals or search query data, to enhance forecasting accuracy even further (Li Long et al., 2021).

An emerging area of research involves the use of hybrid choice models that integrate traditional discrete choice frameworks with latent variable analysis. These models capture both observable factors (such as fare and flight frequency) and unobservable factors, including passenger attitudes, perceptions, and risk aversion. By incorporating latent variables, hybrid choice models provide a richer understanding of passenger decision-making processes, which in turn can inform more effective forecasting and revenue management strategies. Although the application of hybrid choice models in air travel demand forecasting is still in its early stages, preliminary studies

suggest that they can improve predictive performance by better accounting for the nuanced influences of consumer behavior on demand (Ben-Akiva et al., 2002; Abou-Zeid & Ben-Akiva, 2014).

Despite the advancements in forecasting methodologies, several challenges remain. One of the key difficulties is the inherent uncertainty and volatility in air travel demand, which can be exacerbated by disruptive events such as pandemics or geopolitical crises. Traditional models often rely on historical data and may fail to predict sudden shifts in demand patterns. Although machine learning and hybrid models offer improved adaptability, they are not immune to the “black box” problem, where the interpretability of the model’s internal workings is limited. This lack of transparency can be a barrier to adoption by decision-makers who require clear causal explanations to justify strategic investments and operational adjustments. Additionally, the integration of diverse data sources, ranging from historical demand and economic indicators to real-time social media data, poses both methodological and computational challenges. Data heterogeneity, missing data, and noise can all adversely affect forecasting accuracy, necessitating robust preprocessing and data integration techniques.

The literature also highlights the importance of addressing the granularity of demand forecasts. In yield management applications, for example, it is critical to generate forecasts at very granular levels, such as daily or even hourly forecasts, to optimize seat inventory and pricing decisions. However, obtaining accurate forecasts at such a granular level can be challenging, particularly for new routes or markets with limited historical data. One common approach to overcoming this limitation is to aggregate demand at a higher level and then distribute it across lower-level components using methods such as “aggregate and distribute” forecasting. This approach decomposes the demand forecast into an aggregate level and a series of disaggregation factors that allocate the aggregate forecast across different time periods or market segments (Crystal et al., 2007). While effective in certain contexts, this method requires careful calibration to avoid propagating errors from the aggregate forecast into the disaggregated components.

Furthermore, the cyclical nature of the aviation industry adds another layer of complexity to demand forecasting. Airline profits and passenger demand exhibit significant cyclicity, influenced by macroeconomic conditions, seasonal variations, and even long-term trends such as globalization and technological change. As a result, forecasting models must be capable of capturing both short-term fluctuations and long-term trends. Hybrid models that combine the strengths of econometric models for long-term forecasting with machine learning techniques for short-term prediction have shown promise in this regard. For instance, models that integrate time series decomposition methods (such as STL decomposition) with neural networks or support vector regression can effectively capture seasonal patterns and cyclic variations while also adapting to sudden market changes (Dantas et al., 2017; Tillmann, Joormann, & Ammann, 2023).

In addition to methodological advancements, the evolution of air travel demand forecasting has important practical implications for airline network planning and revenue management. Accurate demand forecasts enable airlines to make informed decisions regarding fleet allocation, route network expansion, and capacity planning. They also play a critical role in optimizing pricing strategies by balancing demand and capacity, thereby maximizing revenue per available seat kilometer (RASK). Fiig et al. (2019) demonstrated that even modest improvements in forecast accuracy can lead to significant enhancements in airline revenue management systems, particularly through more precise fare class availability and dynamic pricing strategies.

Recent industry developments further underscore the importance of advanced forecasting models. As airlines continue to invest in new technologies and expand their route networks, the ability to accurately predict passenger demand becomes even more critical. For example, Airbus's upgraded 20-year demand forecast, which projects a substantial increase in the global fleet and improved fuel efficiency through modernized aircraft, relies on sophisticated forecasting models that integrate both macroeconomic trends and technological innovations (Airbus, 2024; Reuters, 2024). Similarly, the shift of budget airlines towards hybrid models that combine low fares with enhanced amenities reflects changing consumer preferences and underscores the need for models that can capture evolving market dynamics (The Wall Street Journal, 2024).

Looking forward, the field of air travel demand forecasting is likely to benefit from further integration of advanced data analytics, machine learning, and hybrid modeling approaches. Emerging data sources such as social media, mobile phone data, and online search queries offer new opportunities to capture real-time shifts in consumer behavior. When combined with traditional economic and operational data, these sources can provide a more holistic view of market dynamics and improve the responsiveness of forecasting models to sudden shocks. Moreover, advancements in computational power and the development of new algorithmic techniques, such as deep reinforcement learning and ensemble learning, promise to further enhance the predictive accuracy and robustness of forecasting systems.

In conclusion, the literature on air travel demand forecasting has evolved considerably over the past few decades. Early models based on econometric regressions and time series methods laid the groundwork for understanding the fundamental relationships between economic indicators and passenger demand. However, the increasing complexity of the aviation industry, driven by globalization, technological advancements, and frequent disruptive events, has necessitated the development of more sophisticated forecasting methodologies. Machine learning and deep learning models have emerged as powerful tools capable of capturing complex nonlinearities and high-dimensional interactions, while hybrid approaches that integrate multiple modeling

techniques offer promising avenues for further improvement. Despite the advances, challenges remain in terms of model interpretability, data integration, and the need for granular forecasts. Addressing these challenges will be critical for enabling airlines to navigate an increasingly volatile market and optimize their operations in the face of rapid technological and economic change.

As the industry continues to evolve, future research should focus on the integration of real-time data, the development of more transparent hybrid models, and the continuous refinement of forecasting methodologies to better capture the multifaceted nature of air travel demand. Such efforts will not only enhance operational efficiency but also contribute to more sustainable and profitable airline operations in an era characterized by rapid change and heightened uncertainty.

Study	Purpose	Data/Market/Period	Approach	Findings
Belobaba, Odoni, & Barnhart (2015)	To describe the historical evolution and strategic advancements of the airline industry.	Global aviation industry, from the early 19th century to modern times.	Historical and strategic analysis.	Technological progress (e.g., jet engines) and deregulation were key drivers of industry growth, competition, and a shift towards cost efficiency.
Wang & Gao (2021)	To review air travel demand forecasting literature.	87 studies published between 2010 and 2020.	Comprehensive literature review.	Most research focuses on national-level demand using time series data. The most common causal factors studied are GDP, income, population, and price.
Zachariah, Sharma, & Kumar (2023)	To synthesize findings on air travel demand drivers.	A range of existing studies.	Literature synthesis.	Key demand drivers include geography, service quality, tourism, seasonality, and low-cost carriers. Disruptive events play a crucial role in shaping demand.
Kölker, Lütjens, & Gollnick (2024)	To analyze passenger itinerary choice and the value of travel attributes.	European aviation market, using 2019 booking data from a major GDS.	Econometric modeling, calculating elasticities, Value of Travel Time (VOTT), and Quality of Travel Index (QTI).	Passenger demand is an interaction between the airline's structural network and the passenger's functional network. VOTT and QTI can quantify passenger preferences.

Multiple Passenger Choice Studies (Adler et al., 2005; Coldren et al., 2003, 2005; Hess, 2007; etc.)	To identify key attributes influencing passenger choice in air travel.	Primarily US and European air travel markets; data from various years (early 2000s).	Discrete choice modeling.	Ticket price, number of connections, travel time, time of day, aircraft type, and carrier are significant factors shaping passenger preferences and flows.
Grosche, Rothlauf, & Heinzl (2007)	To estimate air passenger volumes using geo-economic variables.	Markets where service-related data were unavailable.	Gravity models.	Gravity models using income and population of metropolitan areas can effectively estimate passenger volumes.
Hazledine (2017)	To improve air travel forecasting by adding endogenous factors to gravity models.	Cross-border and domestic air travel.	Augmented Gravity Model.	Adding factors like flight frequency and service quality achieved high explanatory power, with R ² values of 0.76–0.77.
Boonekamp, Zuidberg, & Burghouwt (2018)	To identify additional factors that stimulate demand in gravity models.	North Atlantic market (Europe-North America); data for 2015.	Gravity models.	The presence of ethnic links, knowledge-intensive employment, and low-cost carriers significantly stimulates demand.
Zhou et al. (2018)	To study the impact of airport catchment area definitions on demand prediction.	Australian domestic air passenger market; data for the year 2015.	Gravity models with Poisson pseudo-maximum likelihood (PPML) estimation.	Accurately defining catchment areas and including socio-economic factors (tourism, mining) is important for prediction accuracy.

Carmona-Benítez, Nieto, & Miranda (2017)	To estimate passenger demand using socio-economic indicators.	Mexican air transport market.	Econometric Dynamic Model (EDM) with the Arellano-Bover method.	Passenger demand is highly elastic to the number of flights, followed by economic activity and consumer prices.
Dantas et al. (2017)	To improve the accuracy of time series forecasts for air travel demand.	Brazilian domestic air passenger demand; monthly data from 2004-2014.	Hybrid approach combining bootstrap aggregating (Bagging) with the Holt-Winters method and STL decomposition.	The hybrid method significantly improved forecast accuracy, reducing sMAPE by 31% to 43% compared to traditional methods.
Koç and Arslan (2018)	To forecast domestic air travel demand using neural networks.	Domestic air travel in Turkey.	Artificial Neural Networks (ANNs).	ANNs achieved high accuracy, with Pearson's correlation coefficients between 0.79 and 0.93.
Ramadhani et al. (2020)	To compare the performance of neural networks and ARIMA for forecasting.	Key air travel routes in Indonesia (monthly passenger data).	Backpropagation Neural Network (BPNN) vs. ARIMA.	The BPNN significantly outperformed the ARIMA model, achieving much lower MAPE.
Li Long et al. (2021)	To improve air travel forecasting by integrating search engine data.	Singapore Changi Airport.	Neural Granger Causality models with Google Trends data.	Integrating Google Trends data into forecasting models achieved a very high R^2 value of 0.919.

Jin et al. (2020)	To develop a hybrid model for short-term forecasting.	Singapore Changi Airport passenger arrivals; monthly data from 2004-2019.	Hybrid model using Variational Mode Decomposition (VMD), ARMA, and Kernel Extreme Learning Machines (KELM).	Decomposing the time series and modeling components separately reduced forecast errors and improved robustness.
Mostafaeipour, Goli, & Qolipour (2018)	To predict air travel demand using a meta-heuristic optimized neural network.	Air travel demand in Iran.	Hybrid ANN enhanced with Bat and Firefly meta-heuristic algorithms.	The hybrid model substantially improved prediction accuracy, increasing the R^2 from 0.2 to nearly 0.9.
Sun et al. (2019)	To develop a hybrid model for selecting variables and capturing nonlinear relationships.	Air passenger traffic for Beijing Capital International Airport (PEK); monthly data from 2008-2017.	Hybrid approach combining Mean Impact Value (MIV) for variable selection with a Nonlinear Vector Auto-Regression Neural Network (NVARNN).	The MIV-NVARNN model outperformed SARIMA and standalone neural network models in level accuracy, directional accuracy, and robustness.
Hybrid Choice Model Literature (Ben-Akiva et al., 2002; Abou-Zeid & Ben-Akiva, 2014)	To capture both observable and unobservable factors in passenger	Various transportation contexts, primarily in Europe and North America.	Hybrid choice models integrating discrete choice with latent variable analysis.	These models provide a richer understanding of passenger decision-making by including factors like attitudes and risk aversion,

	choice.			improving predictive performance.
Crystal et al. (2007)	To address forecasting at granular levels when data is limited.	Yield management applications.	"Aggregate and distribute" forecasting methodology.	Demand can be forecasted at a higher level and then distributed to lower-level components, though it requires careful calibration.
Fiig et al. (2019)	To demonstrate the revenue impact of improved forecast accuracy.	Airline revenue management systems.	O&D forecasting system simulation.	Modest improvements in forecast accuracy can lead to significant enhancements in airline revenue, especially via dynamic pricing and fare class management.
Hu (2023)	To develop a nonadditive forecast combination method based on grey prediction for air passenger flow.	Six major airports in Taiwan; monthly passenger flows; Jan 2013–Dec 2019; one- and three-step-ahead evaluation with rolling windows.	Combine GM(1,1), NGBM(1,1), FGM(1,1), FNGBM(1,1) via the Choquet fuzzy integral (λ -fuzzy measure); benchmark against linear (SA, VACO, DMSFE, QP) and nonlinear (MLP, RBFN, SVR)	NCMG outperforms alternatives on average across airports and horizons; improves robustness and mitigates incorrect model choice risk.

			combinations; accuracy via MAPE and RMSE; significance via Friedman/Nemenyi and Diebold–Mariano tests.	
Kanavos et al. (2021)	To compare ARIMA, SARIMA, and deep learning models for aviation demand forecasting.	U.S. aviation data from 48 airports; 125,263 records; 2005–2018.	ARIMA(3,1,2), SARIMA(0,1,1)(0,1,2) ₁₂ vs. Deep Learning Neural Network (4 hidden layers × 512 nodes, ReLU, Adam, learning rate 0.002).	Deep learning model outperforms ARIMA and SARIMA (MAE = 0.02995, RMSE = 0.04257, MAPE = 7.98%, R ² = 0.99121), proving its superiority for nonlinear and seasonal patterns.
Emami Javanmard et al. (2024)	To forecast long-term air passenger and freight demand and assess energy and emission impacts.	Canadian air transport; 1970–2019; forecasts to 2050.	Hybrid framework combining ARIMA, SARIMA, MIDAS, ANN with pairwise optimization (Particle Swarm Optimization); sensitivity analysis via SVM.	Pairwise models improve accuracy by up to 25%; predict 94% passenger growth and 28–34% freight growth by 2050; demand surges significantly increase energy use and CO ₂ emissions.

Saâdaoui et al. (2020)	To improve U.S. air traffic forecasting using a hybrid ANN approach.	U.S. air traffic data (ASM, PE, PLF); monthly; 2000–2017.	Hybrid Feedforward Neural Network (FNN) + NLS-based curve fitting for trend and seasonality; residuals modeled via FNN.	SS-FNN achieves MAPE of 0.57% (12 months), outperforming SARIMA and ARIMA–SVM; seasonality extraction improves accuracy, detrending less effective.
Industry Reports (Airbus, 2024; Reuters, 2024; The Wall Street Journal, 2024)	To provide context on current industry trends and future outlook.	Global airline industry.	Market analysis and forecasting reports.	Industry projects substantial growth in the global fleet, improved fuel efficiency, and a shift by budget airlines towards hybrid service models.

Table 1. Literature Review Summary

3. Research Methodology

3.1.1 Data Sources and Key Variables

To test the performance of different machine learning techniques at forecasting the incremental demand stimulated by a capacity increase, a dataset was built from two distinct but complementary sources: the first one consists of commercial airlines' capacity data between Canada and the world, and the second one consists of air travel ticketed passengers' data from tickets flown between Canada and the world. Demand data was sourced from IATA's Direct Data Solutions (DDS), which aggregates multiple inputs including airline-contributed ticketing data, transactions from IATA's Billing and Settlement Plan (BSP), ARC's Area Settlement Plan (ASP), and both direct and indirect sales channels (e.g., airline websites and travel agencies). Capacity data was obtained from Cirium's DIIO aviation analytics platform, which compiles airline schedules submitted by over 900 carriers through Standard Schedules Information Manual (SSIM) files.

This longitudinal dataset captures the dynamics of the Canadian international air travel market from 2014 to 2024, providing a granular view of both passenger demand and airline capacity. The primary objective in constructing this dataset was to create a unified view that enables the direct analysis of how changes in airline supply (specifically the introduction of new flights or routes) influence passenger demand flows. By merging passenger flow information with detailed airline schedule data, it is possible to systematically investigate the phenomenon of market stimulation. The resulting panel dataset, which tracks multiple Origin-Destination (O-D) pairs over time, is uniquely suited to the time-series and deep learning methodologies employed in this research.

Beyond the variables already included (fare, seats, flights, ASMs, new route indicator, lagged features), other factors that could influence forecasting accuracy include:

- Macroeconomic indicators: GDP, unemployment rates, consumer confidence.
- Competitor activity: Changes in competitor schedules, pricing, and marketing efforts.
- External shocks: Events like pandemics, geopolitical tensions, or natural disasters.
- Marketing expenditure and tourism trends.
- Customer preferences and weather conditions.

Nevertheless, limiting the model to core variables directly related to airline capacity and realized demand allows for a direct analysis of the relationship between capacity expansion and market stimulation, which is the central research question. Also, the chosen variables are consistently available for all O-D pairs and time periods, ensuring robust longitudinal analysis. External variables may vary in availability, definition, and reliability, introducing noise and potential bias. Finally, many benchmark studies in air travel demand forecasting use similar core variables, facilitating direct comparison of results. Including a broader set of external variables would make it difficult to benchmark performance and validate findings against established models.

For the purpose of this work, narrowing the focus to the current set of variables ensures clarity, interpretability, data reliability, and actionable insights. The final dataset is an amalgamation of a demand-side component, detailing passenger flows and fares, and a supply-side component, which describes the capacity offered by airlines.

3.1.2 Passenger Flow Data (Demand-Side)

The demand dataset quantifies the realized passenger demand between city pairs, serving as the primary measure of market activity. The key attributes leveraged from this source include:

- **Sum of Total Market Pax:** This variable represents the total number of passengers traveling on a specific non-directional O-D pair within a given month. It is the central target variable (y) for the forecasting models developed in this thesis.
- **Blended Fare:** This represents the average fare paid by passengers on the O-D pair for the month. It serves as a critical economic input feature, as demand for air travel is inherently sensitive to price.
- **Total Stage Length:** This feature measures the total distance flown for the O-D pair. It acts as a proxy for route characteristics and is used as a static input feature in the analysis.
- **True OD ND:** This identifier for the non-directional Origin-Destination pair is the primary key used to track the performance of a specific route over time.

3.1.3 Schedule and Capacity Data (Supply-Side)

The supply-side data consists of monthly scheduled airline operations and capacity deployment between Canada and international destinations from 2014 to 2024. This dataset provides the crucial explanatory variables related to airline capacity that are hypothesized to stimulate demand. The key attributes utilized are:

- **Flights:** The total number of flights operated on an O-D pair in a given month. This is a direct measure of service frequency.
- **Seats:** The total number of seats offered by airlines on the route in a month. This variable, along with flights, is a primary indicator of available capacity.
- **ASMs (Available Seat Miles):** A capacity measure that accounts for both the number of seats offered and the distance they are flown.
- **New/old route:** A categorical indicator identifying whether a service on a given O-D pair is 'New' or 'Existing'. This feature is fundamental to the thesis, as it allows for the explicit modeling of market stimulation following a new route launch. It was transformed into a binary indicator for use in the models.

3.1.4 Final Dataset Characteristics and Variables

Following a process of cleaning, integration, and feature engineering (detailed in Chapter 4), the two data sources were merged into a single, analysis-ready data frame. This final dataset contains a monthly record for each O-D pair, spanning from January 2014 through December 2024, and encompasses a total of 82,357 individual observations with a set of features describing both the market's state and its month-over-month evolution. It is worth noticing that in the current longitudinal dataset the period going from 2020 to 2022 includes the COVID-19 pandemic, which was characterized by travel disruptions and high travel demand volatility. The key variables used for modeling are described in Table 2.

Table 2: Description of Key Variables in the Final Dataset

Variable Category	Variable Name	Description	Role in Model
Target Variable	Sum of Total Market Pax	Total monthly passengers on an O-D pair.	Target (y)
Demand-Side Features	Blended Fare	The average fare paid by passengers on the O-D pair.	Input Feature
	Total Stage Length	The total distance flown for the O-D pair.	Input Feature
Supply-Side Features	Flights	The number of flights operated on the O-D pair.	Input Feature
	Seats	The total number of seats offered on the O-D pair.	Input Feature
	ASMs	Available Seat Miles; a measure of total capacity.	Input Feature
	New_Route_Indicator	A binary flag (1 for new, 0 for existing) derived from the New/old route column.	Input Feature

Engineered Features	Lag_1_Pax	The Sum of Total Market Pax from the previous month.	Input Feature
	Lag_1_Seats	The number of Seats from the previous month.	Input Feature
	Lag_1_Flights	The number of Flights from the previous month.	Input Feature
	Lag_1_Fare	The Blended Fare from the previous month.	Input Feature
Engineered Features	Seats_Growth_Pct	The percentage change in Seats from the previous month.	Input Feature
	Flights_Growth_Pct	The percentage change in Flights from the previous month.	Input Feature
	Month_Cos, Month_Sin	Trigonometric transformations of the month to capture seasonality.	Input Feature
Identifier	ND OD City	Unique integer ID for each O-D pair after encoding.	Input to Embedding Layer

Table 2. Description of Key Variables in the Final Dataset

In conclusion, the data set compiled for this study provides a multi-faceted view of the air travel market. Its longitudinal structure and the combination of both demand-side and supply-side variables (especially the explicit identification of new route introductions) make it an adequate foundation for developing and testing advanced forecasting models. This dataset enables a direct and quantitative exploration of the relationship between capacity expansion and market stimulation, addressing the core research question of this thesis.

4. Research Objectives and Contributions

This chapter outlines the methodology employed to forecast air travel demand and quantify the impact of capacity changes. The research framework encompasses a multi-stage process, beginning with the collection and integration of raw data, followed by extensive preprocessing and feature engineering to prepare the dataset for modelling. Subsequently, a suite of deep learning models is developed, trained, and evaluated. The entire workflow, from data ingestion to results analysis, is designed to ensure reproducibility and robustness, providing a clear and logical path to answering the core research questions of this thesis.

4.1 Data Collection

The foundation of this study rests on two primary datasets, as detailed in Chapter 3. The data collection phase involved sourcing:

1. **Passenger Flow Data:** A series of yearly CSV files from 2014 to 2024 pulled from IATA DDS, containing monthly passenger counts, blended fares, and route characteristics for non-directional Origin-Destination (O-D) pairs.
2. **Schedule and Capacity Data:** An Excel file detailing airline operations from 2010 to 2024 pulled from Cirium DIIO platform, including the number of flights, seats, and Available Seat Miles (ASMs) for each O-D pair, along with a critical indicator for new route introductions.

These raw data sources were gathered to create a holistic view of the market, capturing both the realized demand (passengers) and the offered supply (capacity).

4.2 Data integration and Preprocessing

To create a cohesive dataset suitable for analysis, the raw data underwent a thorough integration and preprocessing pipeline.

First, all yearly passenger flow CSVs were systematically consolidated. This involved cleaning each file by stripping extraneous header rows and standardizing column names. Dates were parsed into a consistent format, and the True OD ND column was renamed to ND OD City to facilitate merging.

Concurrently, the schedule data from the Excel file was prepared. Metadata and summary rows were programmatically removed, and the date column was parsed to extract year and month components.

The crucial step was the integration of these two sources. The processed passenger data was merged with the schedule data using a left-merge operation, joining on the composite key of ND OD City, Year, and Month. This produced a single, unified data frame containing both supply and

demand variables for each O-D pair over time. Finally, the merged dataset was filtered to the 2014-2024 period, and any rows with missing values for the target variable (Sum of Total Market Pax) were dropped to ensure the integrity of the training data.

4.3 Data Transformation and Partitioning.

With a clean dataset, the next stage involved transforming the data for model consumption and partitioning it for training and evaluation.

- Before modeling, the dataset undergoes several preprocessing steps to stabilize and normalize the data. First, the original passenger demand (Sum of Total Market Pax) is preserved in a separate column (OriginalPax) for reference. A natural log transformation is then applied to the passenger demand and related continuous features (flights, seats, ASMs, block minutes, and their lagged values) to reduce skewness and stabilize variance. Extreme outliers in the capacity growth features (absolute and percentage changes in seats and flights) are capped at the 1st and 99th percentiles to limit their impact. For each route with at least two years of data, an STL (Seasonal-Trend Loess) decomposition is performed on the log-transformed passenger demand. This process extracts the smooth trend and seasonal components, which are stored in new features (Pax_trend and Pax_seasonal), and the passenger demand target is replaced with the STL residual (log demand minus trend minus seasonal), effectively removing seasonal and trend effects from the series.
- Numerical Feature Scaling: All numerical input features were scaled using StandardScaler. This process transforms each feature to have a mean of zero and a unit variance ($z=(x-\mu)/\sigma$). Scaling is essential for the optimal performance of neural networks, as it prevents features with larger magnitudes from disproportionately influencing the model's weight updates during training.
- Categorical Feature Encoding: The ND OD City identifier, being a categorical string, was converted into unique integer IDs using LabelEncoder. This numerical representation allows the model to process the route identifiers, which are subsequently fed into an embedding layer to learn dense vector representations for each route.
- Data Partitioning and Sequencing: Given the time-series nature of the data, a strict chronological split was employed to prevent data leakage and to realistically simulate a forecasting scenario. The dataset was partitioned into a training set (the first 90%) and a test set (the final 10%). For the recurrent (LSTM, GRU) and convolutional (CNN) models, the data was further transformed into sequences. A window of 3 historical time steps (e.g., 3 months) was used to create input sequences for each O-D pair, with the target being the passenger demand in the subsequent month.

4.4 Feature Engineering

To enhance the predictive power of the models and explicitly capture the dynamics of market stimulation, several new features were engineered from the base data:

- **New Route Indicator:** A binary feature (`New_Route_Indicator`) was created from the categorical New/old route column, with a value of 1 signifying the first months of a new route's operation and 0 otherwise. This allows the model to specifically learn the distinct demand patterns associated with new service introductions.
- **Lagged Features:** To provide the models with historical context and momentum, lagged values from the previous month were created for key variables, including Sum of Total Market Pax (`Lag_1_Pax`), Seats (`Lag_1_Seats`), Flights (`Lag_1_Flights`), and Blended Fare (`Lag_1_Fare`).
- **Capacity Growth Metrics:** To quantify changes in supply, month-over-month growth metrics were calculated for Seats and Flights, both in absolute (`Seats_Growth_Abs`) and percentage (`Seats_Growth_Pct`) terms.
- **Cyclical Temporal Features:** To represent the cyclical nature of monthly seasonality without imposing a false ordinal relationship, the month number was transformed using sine and cosine functions:
 - $\text{MonthCos} = \cos(2\pi \cdot 12\text{month})$
 - $\text{MonthSin} = \sin(2\pi \cdot 12\text{month})$

The final dataset was sorted chronologically by O-D pair to ensure the integrity of all time-dependent feature calculations.

4.5 Forecasting Models & Hyper-Parameter Tuning

This study develops and evaluates a suite of deep learning models to forecast passenger demand. The model development process was standardized to ensure a fair comparison. All models were built using the TensorFlow/Keras framework and compiled with the **Adam optimizer**, a widely used and effective gradient descent algorithm. The **Mean Squared Error (MSE)** was chosen as the loss function to be minimized during training.

To prevent overfitting and find an optimal number of training epochs, an **EarlyStopping** callback was employed. This technique monitors the model's performance on a validation set (a portion of the training data) and halts the training process if the validation loss does not improve for a specified number of consecutive epochs (a "patience" of 10). The model then restores the weights from the epoch with the best validation performance, ensuring a robust and well-generalized model.

4.6 Deep Learning Model Architectures

Four distinct neural network architectures were implemented, each designed to capture different types of patterns within the data.

4.6.1 Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) (Grossi and Buscema, 2007) are a class of machine learning models inspired by the biological neural networks that constitute animal brains. ANNs consist of layers of interconnected nodes, or "neurons," where each connection has an associated weight. These networks are designed to recognize patterns and make predictions by learning from data.

In contrast to recurrent neural networks (RNNs), ANNs are typically feedforward networks where information flows in one direction—from input to output—without cycles or loops. This structure makes ANNs suitable for a variety of tasks, including classification, regression, and clustering, but it limits their ability to handle sequential data effectively.

ANNs learn to map inputs to outputs through a process known as backpropagation, which involves adjusting the weights of connections based on the error between the predicted and actual outputs. The activation function σ applied to each neuron's output introduces non-linearity into the model, allowing ANNs to capture complex relationships within the data. Figure 2 (Buscema et al., 2014) depicts the ANN structure and how it is defined.

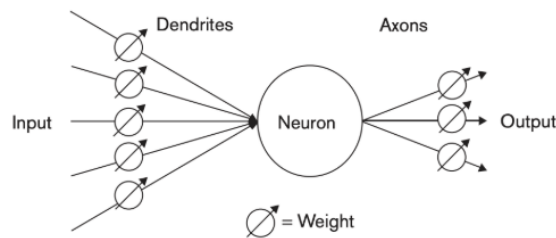


Figure 2. ANN Structure

The basic mathematical formulation for the output γ of a neuron in an ANN layer can be expressed as:

$$\gamma = \sigma(\sum_i w_i x_i + b) \quad (1)$$

where w_i are the weights, x_i are the input features, and b is the bias term. The activation function σ , such as the sigmoid, tanh, or ReLU, determines the output of the neuron based on the weighted sum of inputs.

ANNs are particularly effective for tasks where the input data is not sequential, such as image classification, where each input image can be processed independently. Despite their simplicity, ANNs can achieve high performance on a wide range of problems when appropriately designed and trained.

One of the main challenges in training ANNs is avoiding overfitting, where the model performs well on training data but poorly on unseen data. Techniques such as regularization, dropout, and early stopping are commonly used to mitigate this issue and improve the generalization capability of ANNs.

While ANNs are powerful tools for many machine learning tasks, their feedforward architecture is best suited for static, non-sequential data. Their simplicity and versatility make them a fundamental building block in the field of neural networks, laying the groundwork for more advanced architectures like RNNs, Convolutional Neural Networks (CNNs), and others.

4.6.2 Long Short-Term Memory (LSTM) Network

Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) networks are a specialized variant of Recurrent Neural Networks (RNNs) crafted to overcome the challenge of capturing long-range dependencies in sequential data. Introduced by Hochreiter and Schmidhuber in 1997, LSTMs have gained widespread recognition for their efficacy in handling sequential information. Its architecture has been depicted in Figure 3 (Géron, n.d.).

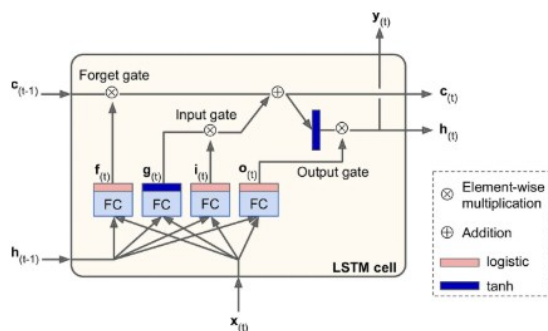


Figure 3. LSTM Structure

LSTMs feature a unique memory cell structure with three gates (input, forget, and output) that regulate the flow of information. These gates enable LSTMs to selectively retain or discard information over time steps, allowing them to effectively capture and utilize long-term dependencies in sequential data without facing the vanishing gradient problem encountered by traditional RNNs. Each memory cell is built around a central linear unit with a fixed self-connection (the CEC). In addition to net_{c_j} , c_j gets input from a multiplicative unit out_j , and from

another multiplicative unit in_j . in_j 's activation, at time t is denoted by $y^{in_j}(t)$, out_j 's by $y^{out_j}(t)$. We have:

$$y^{out_j}(t) = f_{out_j}(net_{out_j}(t));$$

(3)

$$y^{in_j}(t) = f_{in_j}(net_{in_j}(t));$$

(4)

The input gate decides the relevance of new information, the forget gate manages the retention or deletion of outdated data, and the output gate controls the output based on the current cell state and input. These mechanisms, facilitated by sigmoid and tanh activation functions, empower LSTMs to retain essential information, enabling them to capture intricate patterns and dependencies across sequences.

LSTMs find extensive application in tasks requiring sequential modeling, such as machine translation, speech recognition, sentiment analysis, and time series forecasting. Their ability to comprehend contextual information and long-range dependencies makes them invaluable in diverse fields.

However, LSTMs might demand substantial computational resources, particularly in training larger networks or with extensive datasets. Despite their ability to capture long-term dependencies effectively, modeling very extended dependencies could still present challenges.

4.6.3 Gated Recurrent Unit (GRU) Network

GRUs (Cho et al., 2014) consist of a simplified architecture with two gates (an update gate and a reset gate) designed to manage the flow of information throughout the network. Unlike LSTMs, GRUs integrate the reset and update gates, allowing for a more straightforward and computationally efficient structure.

The update gate in GRUs controls the blending of new input information with the previous memory state, determining the extent to which the new information is incorporated. Meanwhile, the reset gate regulates the amount of past information to forget or retain based on the current input. Figure 4 (Harshed, 2024) depicts the GRU structure and how it is defined.

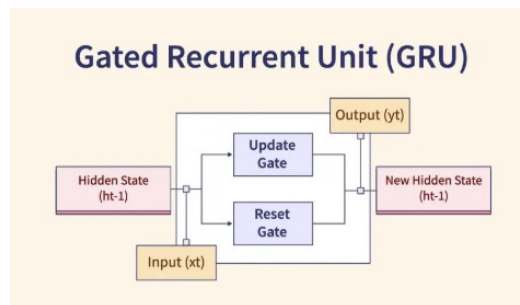


Figure 4. GRU Structure

The simplicity of the GRU architecture, with its reduced number of gates, results in a more straightforward training process and computational efficiency compared to LSTMs. Despite their simpler design, GRUs have demonstrated competitive performance in various sequential modeling tasks.

The gated recursive convolutional neural network (grConv) consists of four weight matrices W^l , W^r , G^l and G^r . At each recursion level $t \in [1, T - 1]$, the activation of the j^{th} hidden unit $h_j^{(t)}$ is computed by:

$$h_j^{(t)} = \omega_c h_j^{(t)} + \omega_l h_{j-1}^{(t-1)} + \omega_r h_j^{(t-1)} \quad (5)$$

where ω_c , ω_l and ω_r are the values of a garter that sum to 1.

GRUs are widely applied in natural language processing, speech recognition, and time series analysis, among other areas requiring sequential data processing. Their capability to capture dependencies and context across sequences makes them valuable in scenarios where computational efficiency and streamlined architecture are essential.

4.6.4 1D Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) stand as pivotal deep learning models tailored for processing grid-like data such as images and time series. Known for their hierarchical feature extraction capabilities, CNNs employ convolutional layers that apply learnable filters to detect local patterns within the input data. These filters progressively capture features, starting from basic edges and textures, and evolving to more complex representations in deeper layers.

1D Convolutional Neural Networks (CNNs) are specialized deep learning architectures adapted for processing sequential data, particularly time series, signals, and textual data. Unlike traditional CNNs primarily designed for spatial data like images, 1D CNNs are specifically structured to capture dependencies and patterns along a single axis, such as a temporal sequence.

The fundamental concept of 1D CNNs involves employing one-dimensional convolutional filters that move across the input sequence, extracting local patterns and capturing relationships between adjacent data points. These filters slide through the sequential input, performing convolutions to detect temporal patterns, transitions, and dependencies, akin to how 2D CNNs capture spatial features in images. Figure 5 (Kiranyaz et. al, 2021) depicts a sample on how the CNN architecture is built.

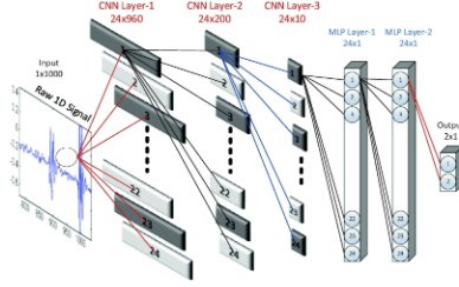


Figure 5. CNN Architecture

In each CNN layer, 1D forward propagation (1D-FP) (Kiranyaz et al., 2021) is expressed as follows:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} \text{conv1d}(w_{ik}^{l-1}, s_i^{l-1}) \quad (1)$$

Where x_k^l is defined as the input, b_k^l is defined as the bias of the k^{th} neuron at layer l , s_i^{l-1} is the output of the i^{th} neuron at layer $l - 1$, w_{ik}^{l-1} is the kernel of the i^{th} neuron at layer $l - 1$ to the k^{th} neuron at layer l . $\text{conv1d}(\dots)$ is used to perform ‘in-valid’ 1D convolution without zero-padding.

Similar to their 2D counterparts, 1D CNN architectures consist of convolutional layers, often followed by pooling layers that reduce dimensionality, thereby extracting and emphasizing crucial features from the input sequence. These layers enable hierarchical feature extraction, progressively learning higher-level representations as information flows through the network.

The advantages of 1D CNNs lie in their ability to automatically learn relevant features from raw sequential data without relying on manual feature engineering. They excel in capturing local patterns and dependencies, making them valuable for various tasks such as time series forecasting, speech recognition, sentiment analysis in text, and bioinformatics.

4.7 Evaluation Metrics

To quantitatively assess and compare the performance of the forecasting models, four standard regression metrics were employed. These metrics were calculated on the unseen test set.

- Mean Absolute Error (MAE) computes the average of absolute differences between predicted and actual values, focusing on the average magnitude of errors, regardless of their direction.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- Root Mean Squared Error (RMSE) calculates the square root of the average of squared differences between predicted and actual values over 'n' observations.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- Mean Absolute Percentage Error (MAPE): MAPE measures the average magnitude of errors in predictions as a percentage of actual values. It is scale-independent and useful for comparing forecast accuracy across datasets. A lower MAPE indicates better predictive performance.

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

The calculation of MAE, RMSE and MAPE offers insights into the model's adaptability to changing patterns over time. This evaluation process enhances the reliability and applicability of the findings, offering valuable insights into the model's performance in forecasting time-dependent numerical values.

5. Results

This section presents an analysis of the empirical results obtained from the evaluation of the four neural network architectures: Artificial Neural Network (ANN), Long Short-Term Memory (LSTM) network, Gated Recurrent Unit (GRU) network, and 1D Convolutional Neural Network (CNN). The primary objective is to assess their efficacy in forecasting air travel demand, represented by the Sum of Total Market Pax, and to understand their ability to model the complex dynamics introduced by factors such as seasonality, fare changes, and capacity adjustments, including market stimulation from new route introductions. The models were evaluated on a held-out test set, representing the most recent 10% of the chronologically sorted data from 2014 to 2024. Performance was quantified using four standard regression metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE).

5.1 Quantitative Model Performance

The performance of each developed model on the unseen test data is summarized in Table 3. This table provides a quantitative basis for comparing the predictive accuracy and explanatory power of the different architectures. The metrics collectively offer an insight on model performance: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) quantify the average prediction error in absolute terms (passengers), with RMSE penalizing larger errors more strongly.

Mean Absolute Percentage Error (MAPE) measures the average error as a percentage of the actual values, making it scale-independent, where, for example, a MAPE of 5% indicates the forecast is off by 5% on average. Lower values on all these metrics indicate better performance.

Table 3: Performance Metrics of Neural Network Models

Model	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)	Mean Absolute Percentage Error (MAPE)
1. Artificial Neural Network (ANN)	322.89	772.50	17.57
2. Long Short-Term Memory (LSTM)	132.28	364.74	5.78
3. Gated Recurrent Unit (GRU)	118.59	316.66	5.23
4. 1D Convolutional Neural Network (CNN)	172.74	388.07	5.76

Table 3. Performance Metrics of Neural Network Models

Overall, the results reveal a clear pattern: recurrent sequence-learning models outperform the non-sequential models. The Gated Recurrent Unit (GRU) network achieved the highest accuracy, with an MAE of ~118.6, RMSE ~316.7, and MAPE 5.23%. This means the GRU’s predictions deviated from actual passenger counts by only about 5.2% on average – indicating a high degree of reliability for monthly route-level forecasts. The Long Short-Term Memory (LSTM) network was the next best, with a MAPE of 5.78% (MAE ~132.3). In contrast, the feed-forward Artificial Neural Network (ANN) had by far the worst performance (MAE ~322.9, MAPE 17.57%), indicating its predictions were nearly 18% off on average – an error level so high that the model is barely better than a naive prediction using the historical mean. The 1D Convolutional Neural Network (CNN) performed substantially better than the ANN (MAE ~172.7, MAPE 5.76%) but still underperformed compared to the recurrent models. Notably, the CNN’s MAPE (5.76%) was in the same range as the LSTM’s, but its higher RMSE suggests it made some larger errors that the LSTM avoided. In summary, the two recurrent architectures (LSTM and GRU) significantly outperformed the non-recurrent architectures (ANN and CNN) across all metrics – underscoring the importance of sequence memory for capturing the temporal dynamics of air travel demand.

5.2 Detailed Performance Analysis of Individual Models

To understand the reasons for the observed performance differences, the next step is to analyze each model's results in the context of its architecture and the nature of the dataset.

5.2.1 Artificial Neural Network (ANN)

The ANN, implemented as a multi-layer perceptron, served as the baseline model. It produced the highest error rates, with an MAE of approximately 322.9 and a MAPE of 17.57%. This level of error indicates that the model's predictions were not sufficiently accurate for practical use. The ANN processes each input independently and lacks the ability to model temporal relationships. Although lagged features were included, the model treated them as static inputs and did not learn sequential patterns. These results suggest that the ANN is not well suited for forecasting tasks involving time-dependent data such as monthly air travel demand.

5.2.2 1D Convolutional Neural Network (CNN)

The CNN was designed to identify short-term patterns within a three-month input window. It achieved an MAE of approximately 172.7 and a MAPE of 5.76%, which is a substantial improvement over the ANN. The model was able to detect local trends, such as sudden increases or decreases in demand, but lacked the capacity to retain information beyond the input window. This limitation affected its ability to model longer-term seasonal or trend-based patterns. The CNN performed moderately well but did not match the accuracy of the recurrent models.

5.2.3 Long Short-Term Memory (LSTM) Network

The LSTM model achieved an MAE of approximately 132.3 and a MAPE of 5.78%. Its architecture includes mechanisms for retaining and discarding information over time, which allowed it to learn temporal dependencies in the data. The model was able to capture seasonal patterns, trends, and the influence of lagged variables more effectively than the non-recurrent models. The LSTM's performance confirms that models with memory components are better suited for forecasting tasks involving sequential data.

5.2.4 Gated Recurrent Unit (GRU) Network

The GRU, a simplified variant of the LSTM, achieved the best performance among the models tested, with an MAE of approximately 118.6, RMSE of 316.7, and MAPE of 5.23%. This level of error suggests that for a route with 10,000 monthly passengers, the GRU's forecast would typically be within about 523 passengers. Its architecture uses a single update gate and merges the cell and hidden states, which makes it easier to train and may contribute to better generalization. In this study, the GRU captured long-range dependencies, seasonal patterns, and the effects of engineered

features such as new route indicators and capacity growth metrics as effectively as the LSTM, and showed less tendency to overfit. The GRU's smaller parameter set may have been better suited to the dataset, which includes monthly data for over 100 routes across a ten-year span. The results indicate that the GRU was able to learn meaningful representations of route states and produce more accurate predictions than the other models. Based on these findings, the GRU is the preferred model for forecasting air passenger demand in this context.

5.3 Comparative Analysis and Benchmarking

This chapter presents a comparison between the results of this thesis and those reported in recent benchmark studies.

To contextualize the results obtained, recent papers related to air travel demand forecasting were used as benchmarks. In Table 4 the main characteristics of each study are briefly described.

Study (Year)	Approach & Model	Data & Features
Present Thesis (2025)	Deep learning (ANN, CNN, LSTM, GRU) no hybridization; best model GRU.	Historical monthly flight capacity and passenger data (no external predictors)
Liang et al. (2022)	Hybrid decomposition-ensemble model (MIC_MEMD_BiLSTM) combining EMD decomposition with BiLSTM neural network.	Chinese airports data: monthly passenger volumes + Baidu Index internet search data as exogenous variable.
Jafari & Lewison (2024)	Two-stage deep neural network with multiple inputs/outputs (GRU model identified as best performer)	U.S. aviation market data: monthly passenger counts with airline-level disaggregation (competition between carriers incorporated).
Saâdaoui et al. (2020)	Hybrid Feedforward ANN + NLS regression (trend & seasonality explicitly modeled then combined with ANN output)	U.S. air traffic (national) monthly series, decomposed into trend and seasonal components (for regression).

Gunter & Zekan (2021)	Econometric Global Vector Autoregression (GVAR) – models interdependence among multiple airports and global factors	Global panel of the world’s busiest airports; included domestic & international economic indicators and oil prices as predictors.
Emami Javanmard et al. (2024)	Hybrid machine learning + optimization framework (included a MIDAS-ARIMA model among best performers).	Canadian air passenger & freight demand with diverse features: GDP, population, energy consumption, CO ₂ emissions, etc. (economic, social, environmental data).
Kanavos et al. (2021)	Classical time-series vs. deep learning: compared ARIMA and SARIMA models with a deep learning neural network (DLNN) approach.	Global (multi-route) demand data: monthly aviation passenger numbers (e.g. U.S. domestic routes, 2005–2018); included route-wise data and socio-economic features (e.g. population, GDP, income for origin/destination) for the DL model.

Table 4. Similar Travel Demand Forecasting Papers Summary

In table 5 the described papers are compared using three metrics: MAE, RMSE and MAPE.

Study	Model	MAE	RMSE	MAPE	Data scope
Present Thesis (2025)	GRU	118	316	5.23%	Route level, monthly
Present Thesis (2025)	LSTM	132	364	5.78%	Route level, monthly
Present Thesis (2025)	CNN	319	617	5.76%	Route level, monthly
Present Thesis (2025)	ANN	322	772	17.57%	Route level, monthly
Liang et al. (2022)	MIC_MEM D_BiLSTM	103.1k	117.3k	1.62%	Airport level, monthly
Liang et al. (2022)	BiLSTM	139.2k	155.2k	2.21%	Airport level, monthly
Liang et al. (2022)	SARIMA	163.8k	205.7k	2.62%	Airport level, monthly
Gunter & Zekan (2021)	GVAR h=1	0.054M		0.35%	Airport level, quarterly
Gunter & Zekan (2021)	VAR h=1	0.032M		0.21%	Airport level, quarterly
Gunter & Zekan (2021)	ARIMA h=1	0.026M		0.17%	Airport level, quarterly
Jafari & Lewison	LSTM		1.56M	2.20%	Route group level,

(2024)					monthly
Jafari & Lewison (2024)	Conv1D		4.04M	5.69%	Route group level, monthly
Jafari & Lewison (2024)	ConvLSTM 2D		3.54M	4.90%	Route group level, monthly
Saâdaoui et al. (2020)	SS-FNN h=1			0.57%	Airport level, yearly
Emami Javanmard et al. (2024)	ANN	41.9k	43.6k	51.54%	Route group level, yearly
Emami Javanmard et al. (2024)	ARIMA	140.1k	1973	1.59%	Route group level, yearly
Emami Javanmard et al. (2024)	MIDAS	698.6k	785.0k	7.92%	Route group level, yearly
Emami Javanmard et al. (2024)	MIDAS-ARIMA	111.1k	159.5k	1.27%	Route group level, yearly
Kanavos et al. (2021)	ANN	299.5	425.7	7.98%	Route level, monthly

Table 5. Performance Benchmark Against Other Research Papers

To put the forecasting results into context, the performance of the GRU and LSTM models is compared with findings from recent air travel demand forecasting studies (see Table 5). Direct comparisons are made cautiously due to differences in data scope (route-level vs. airport or national level), feature sets (historical internal data vs. inclusion of external predictors), and time periods (particularly whether the volatile 2020–2022 pandemic years are included in model training).

Overall Performance vs. Literature: The GRU and LSTM models in this thesis achieved MAPE values of approximately 5.2% and 5.8%, respectively, on route-level monthly data, with corresponding MAE and RMSE values summarized in Table 5. These errors are moderate compared to the very low errors reported in some studies using aggregated data or exogenous inputs. For example, Liang et al. (2022), using a hybrid BiLSTM model with a decomposition-ensemble strategy and Baidu search index as an external predictor, reported an extremely low MAPE of about 1.6% when forecasting monthly passenger throughput at major Chinese airports. Similarly, Günter and Zekan (2021), employing a Global Vector Autoregression on quarterly airport traffic, achieved MAPE below 0.5%. By contrast, studies focusing on route-level or market-level predictions tend to report higher error rates. Kanavos et al. (2021), for instance, obtained roughly 7.98% MAPE for a deep learning model on multi-route monthly data (even with some socio-economic features included), which is higher than the GRU’s error in this thesis. The results here (GRU ~5.2% MAPE) indicate better accuracy than that older deep learning benchmark, likely due to the use of a more suitable recurrent architecture and extensive feature engineering (lags,

seasonality adjustments, etc.). They are, however, not as low as the error rates seen in studies that leverage more aggregated series or richer external information.

Differences in Data Scope: one reason for the disparity in error rates is the level of data aggregation. Forecasting route-level monthly demand (as done in this thesis) is inherently more difficult than forecasting airport-level or national totals. Route-level data are noisier and more volatile, because they capture localized effects (competition, specific regional events, etc.) and have lower traffic volumes where percentage changes can be large. In contrast, airport-level or country-level passenger numbers aggregate many routes, smoothing out idiosyncratic fluctuations. The literature reflects this: models predicting total airport passengers can achieve very low MAPE (often 1–3%) since the data are more stable. For example, Liang et al.'s aforementioned study dealt with airport totals and benefited from smoother series, whereas the GRU in the present thesis deals with dozens of individual routes, including low-density and high-variance markets. Thus, the 5% MAPE at route-level can be considered a reasonable outcome when compared to a ~2% MAPE at airport-level in other work – these figures are not directly equivalent in difficulty.

Effect of External Features: The use (or omission) of exogenous input features is another critical factor. This thesis deliberately uses only internal historical data (past demand, capacity, fares, etc.) for forecasting. In contrast, several high-performing models in the literature integrate external predictors that provide additional context. For instance, Liang et al. (2022) included Internet search query indices (a leading indicator of travel interest) and demonstrated a 11–20% improvement in accuracy over models using only historical data. Emami Javanmard et al. (2024) incorporated a wide range of economic and demographic variables (GDP, population, energy prices, etc.) to forecast Canadian air travel demand; their best hybrid model (a MIDAS-ARIMA combination) achieved MAPE ~1.27%, far lower than an uninformed ANN which had MAPE >50% when it lacked those external inputs. These examples underscore that neural network models require the right inputs to be effective. The models in this thesis mitigated the lack of exogenous data through feature engineering (lagged features, growth rates, seasonal decompositions), but they still rely solely on patterns in the historical route data. As a result, certain kinds of forecast errors—especially those due to factors like economic shocks or policy changes—could not be preemptively corrected by the model. In comparison, studies that fed such information directly into the model were able to achieve lower errors by capturing influences that pure time-series patterns alone might miss.

Impact of the COVID-19 Pandemic Period: A distinguishing aspect of this work is the inclusion of the 2020–2022 period in model training and evaluation. This was an extraordinarily volatile time for air travel demand: global passenger volumes plummeted in 2020 and rebounded irregularly through 2021–2022. Including this period means the model had to contend with non-stationary behavior and structural breaks vastly beyond the range of the pre-2020 patterns. Many contemporary studies avoid this issue by excluding or separately modeling the pandemic era. For example, Jafari (2022) observed that including 2020–2021 data caused traditional models to break down, and Jafari & Lewison (2024) explicitly restricted their training data to end at 2019 to

maintain stability in their models. In this thesis, no such exclusion was made; the models were trained through 2024, forcing them to learn across the pandemic shock. This undoubtedly increased the overall error metrics: even though the GRU and LSTM can, to some extent, adapt to trend changes (especially with the help of STL detrending in preprocessing), accurately forecasting the sudden collapse and surge in demand is extremely difficult for any model. The GRU's performance (~5% MAPE) still indicates that it captured general post-pandemic recovery trends reasonably well (achieving about 95% accuracy overall), but its errors during 2020–2022 would have been larger than in normal years, raising its average MAPE. This factor likely accounts for much of the gap between the ~5% MAPE here and the <3% MAPE reported in some pre-pandemic studies. In summary, including the COVID-19 disruption makes the forecasting task significantly more challenging, and the models' moderate error rates should be viewed in light of this more rigorous test. By reflecting real-world volatility, the evaluation in this thesis is more conservative than those in studies limited to stable periods.

Alignment with Literature Trends: Despite the above differences, the relative performance pattern observed here is consistent with broader research findings. Recurrent neural networks (sequence-based models) outperform simpler architectures for air travel demand forecasting. The fact that both the LSTM and GRU surpassed the ANN and CNN in this thesis mirrors results in other studies where LSTM/GRU models outshine traditional models or feed-forward nets when handling sequential data. For example, Jafari and Lewison (2024) found a GRU-based sequence model to be the top performer on a multi-route airline dataset, and Kanavos et al. (2021) noted that a deep learning approach (even a basic one) beat out ARIMA on route-level data. Thus, the value of sequence learning emphasized by this thesis (with GRU as the best model) is well supported by the literature. Furthermore, the GRU's edge over the LSTM here (slightly lower error and better generalization) might be explained by its simpler gating mechanism requiring fewer parameters; this advantage in generalization has been reported in other contexts as well, especially on relatively limited datasets.

Future Directions for Improvement: While the GRU model performed well relative to the constraints, there is evident scope for reducing errors further by adopting some of the strategies from these benchmark studies. Incorporating exogenous variables is a primary avenue: relevant features such as economic indicators (GDP, employment rates), oil prices, or online search frequencies could enhance the model's ability to anticipate demand changes, especially turning points that are driven by external events. The literature strongly suggests benefits from this approach (e.g., adding Baidu Index search data in Liang et al. improved BiLSTM forecast accuracy, and macroeconomic features in Javanmard et al. enabled their hybrid model to outperform simpler models by a wide margin). Another direction is to explore decomposition-ensemble methods more deeply. This thesis used a basic STL decomposition to detrend and deseasonalize the passenger series; future work could employ more advanced techniques like Empirical Mode Decomposition (EMD) or wavelet transforms to break down the time series, followed by specialized models for each component (as done by Liang et al. and others).

Combining the strengths of statistical models and machine learning in a hybrid framework (for example, an ARIMA model to handle long-term trend and seasonality, and a neural network for the residuals) could yield better accuracy and robustness, as evidenced by studies like Saâdaoui et al. (2020) and Emami Javanmard et al. (2024). Additionally, examining alternative neural architectures may prove fruitful. Transformer-based sequence models, which have shown promise in other time-series domains, could potentially capture very long-range dependencies across years of data more effectively than LSTM/GRU units. Hybrid architectures (such as CNN-LSTM combinations or attention-augmented RNNs) might also uncover complex patterns that a standalone GRU might miss. Finally, methods to address shock events explicitly could be integrated – for example, incorporating regime-switching models or anomaly detection mechanisms to handle periods like pandemics. Such approaches could allow the model to adjust its behavior when the underlying data-generating process undergoes a sudden change, thereby maintaining accuracy during extreme events.

6. Findings and Discussion.

The results show that sequence-based models (GRU and LSTM) achieved significantly better forecasting accuracy than the non-sequential models (ANN and 1D CNN). This confirms that capturing the temporal dynamics of air travel demand—its seasonality, trends, and momentum—is crucial for accurate prediction. The GRU model delivered the strongest performance, with a coefficient of determination $R^2 \approx 0.58$ on the test set (indicating it explained about 58% of the variance in passenger demand). This level of accuracy illustrates the value of architectures capable of learning long-range dependencies. In contrast, the feed-forward ANN and the CNN, which cannot retain sequential context beyond the given input window, struggled to forecast the volatile demand patterns. These findings validate that recurrent neural networks are particularly well-suited to the problem of route-level passenger demand forecasting, especially when modeling the “market stimulation” effect of new capacity.

An important question is whether these models can inform decisions about the optimal time of day to schedule a flight. In their current form, the models cannot directly answer this question, because time of day was not included as an input feature. The dataset is aggregated at the monthly level, so the predictions reflect total monthly passengers on a route rather than variations by hour of day. Consequently, factors like morning versus evening flight preferences are beyond the model’s scope. Past research works indicate that time-of-day can influence passenger choices (e.g., business travelers prefer morning departures), as demonstrated by Wei, Vaze, & Jacquillat (2020), who showed that aligning flight schedules with passenger time-of-day preferences improves both customer satisfaction and airline profitability, and by Muñoz, Laniado, & Córdoba (2020), whose stated-preference modeling revealed strong traveler preferences for specific departure and return times in international round-trip itineraries. However, the current model does not capture such granularity. To use deep learning for optimal flight timing decisions, one would need a more fine-grained approach. For example, incorporating departure time as a feature or developing a separate model on daily/hourly data that accounts for time-of-day demand patterns. In this thesis, since

time-of-day was outside the modeling scope, the forecast results are applicable to overall route capacity decisions (whether or not to add flights on a route and total monthly seats) rather than tactical scheduling of flights at specific hours. In practice, airlines could complement the model's route-level demand forecasts with internal scheduling analysis or customer preference data to determine the best departure times, an area for potential future work.

Another point to take into account is that the GRU model had an acceptable result even when trained and tested on the highly volatile 2020–2024 period, which included the COVID-19 pandemic. The inclusion of the pandemic years meant the model faced extreme disruptions in the training data; global passenger volumes plummeted in 2020 and rebounded irregularly in 2021–2022. Despite this, the GRU was able to extract enough signal to maintain reasonable accuracy, suggesting that sequence-based deep learning models can be valuable tools even during periods of significant demand upheaval. However, including the COVID-19 shock did increase the overall error compared to a scenario without this period. Many contemporary studies improved accuracy by excluding or separately modeling the pandemic era; for example, Jafari & Lewison (2024) report that omitting 2020–2021 data helped maintain model stability. In the current case, if the COVID-19 years were removed from the dataset, the forecasting task would become easier (limited to more typical demand patterns) and the models' performance would likely improve further. Lower error metrics (e.g. MAPE) would be expected because the models would not need to accommodate such an extreme period, unprecedented fluctuation. In essence, the pandemic introduced structural breaks that no model could capture perfectly; excluding that period would let the GRU and LSTM focus on regular seasonal and trend relationships, likely reducing their forecast errors. This trade-off was consciously accepted in this thesis to test the model's resilience. The fact that the GRU achieved approximately 5% MAPE with the pandemic included is encouraging. Therefore, had the COVID-19 period been excluded, the GRU's average accuracy might have been higher (perhaps in the 3–4% MAPE range, judging by how much other studies' errors increased when including pandemic data). This discussion underscores how shock events affect model performance: including real-world disruptions makes evaluation more rigorous but can inflate error metrics, whereas restricting the training set to stable periods yields more optimistic results.

From an industry perspective, the improved accuracy of the GRU (and LSTM) model has several practical implications for airlines and airport authorities:

- **Route Planning:** More reliable forecasts of passenger uptake on new or expanded routes can inform network expansion decisions. A model that anticipates demand stimulation from added capacity helps evaluate the viability of a new route or a frequency increase, estimating expected load factors and revenue. This reduces the financial risk of launching routes: if the model predicts that adding a flight will increase monthly passengers by, say, 15%, planners can judge whether that demand justifies the added costs. In this study, the GRU's ability to quantify stimulated demand means airlines could use it to prioritize which

route additions are most promising (e.g., choosing between several candidate new destinations by comparing their forecasted demand growth).

- **Revenue Management:** With better demand predictions, airline revenue management teams can optimize pricing and seat inventory control. By anticipating high-demand periods or market growth due to added capacity, airlines can adjust fares and allocate seat classes in advance to maximize revenue per available seat-mile (RASM) while remaining competitive. For example, if the model foresees a significant uplift in demand on a route after a new flight is introduced, the airline might initially set more aggressive promotional fares to stimulate early bookings, and then yield-manage by raising prices as the flights start filling up closer to departure. On the contrary, if demand stimulation is predicted to be modest, the airline may maintain higher fares or limit capacity deployment. The forecasting of stimulation allows proactive pricing strategies, aligning with the expected demand curve influenced by capacity changes.
- **Operational Planning:** Accurate demand forecasts enable more efficient resource allocation both for airlines and airports. Airlines can align fleet assignment (aircraft type and size) and crew scheduling to match expected passenger volumes, avoiding under-utilization (too large an aircraft for the demand) or overcrowding (too small an aircraft or too few flights). For instance, if the model predicts strong growth on a route, an airline might upgauge from a narrow-body to a wide-body aircraft or add another daily flight to accommodate the forecasted passengers. Airport operators, on their side, can use these forecasts to plan ground handling, check-in counters, security staffing, and gate assignments, ensuring that airport services are scaled appropriately to the demand. In essence, planning based on a robust forecast helps prevent both excess capacity (empty seats, idle staff) and bottlenecks (insufficient check-in counters, long lines) in operations. The models' demonstration of reliability even during volatile periods means they could be trusted for operational planning in uncertain times – for example, airports saw unpredictable surges during the pandemic recovery, and a model like this paper's GRU could aid in anticipating those surges.

Another point of discussion is the rationale behind the input features chosen for the models. The feature set in this study was designed to capture key demand and supply factors in a way that is broadly applicable across the airline industry (rather than proprietary airline-specific data). Table 2 (see Section 3.1.4) lists the variables: notably, demand-side metrics were included like total passengers (the target) and average fare, and supply-side metrics like number of flights, seats, and ASMs (Available Seat Miles) on each route. These core variables were selected because they directly relate to the phenomenon of interest – capacity expansion and resulting market stimulation – and they are standard measures in air transport research and practice. In essence, the inputs represent the essential forces of demand and supply: fare is a price indicator affecting demand, while flights/seats/ASMs indicate capacity offered. Engineered features derived from these basics were also included, such as a `New_Route_Indicator` (binary flag for the first months of a new route) to explicitly mark when a capacity expansion occurs, lagged variables (e.g., previous month's passengers, fare, capacity) to provide the models with recent historical context, and

seasonal indicators (sine/cosine of month) to capture recurring patterns. This set of features was inspired by both domain knowledge and data availability:

- **Domain Knowledge:** The literature and industry studies point to certain drivers of air travel demand – price, service frequency, connectivity, and seasonality are repeatedly shown to influence passenger volumes. Therefore, fare and direct capacity metrics were included to reflect these drivers. For instance, prior passenger choice studies highlight *ticket price*, *number of connections*, *travel time*, *time of day*, etc., as significant factors shaping demand. In the context of market-level forecasting, “time of day” and connection options were not explicitly modeled, but flight frequency (flights per month) and route distance (stage length) serve as proxies for service availability and travel convenience, respectively. The focus was set on variables that airlines themselves can tweak (adding flights, adjusting capacity) or observe (fares, demand levels), aligning the model with decision factors that a network planner or revenue manager understands.
- **Data Availability and Consistency:** The chosen features were all consistently available for the entire dataset and for all routes, sourced from industry data aggregators (IATA DDS for demand and fares, Cirium for schedule capacity). By using common industry datasets, it was ensured that the inputs are standardized and comparable across different markets and time periods. Importantly, features that would be company-specific were avoided, such as an individual airline’s marketing spend, frequent-flyer membership statistics, or internal operational costs, since those are not publicly available for all routes and would limit the model’s generality. Instead, the inputs like total seats or average fare are *industry-wide metrics* at the route level – they encompass all carriers in the market, not just one airline’s contribution. This makes the model’s findings relevant to any airline or stakeholder examining the route, not just a single company’s perspective. The downside is that the effect of one carrier vs. another is not explicitly distinguished (e.g., a low-cost carrier versus a full-service carrier both add seats, and the model treats a seat as a seat regardless of who offers it). But this abstraction was deliberate to keep the analysis at the market level.

Overall, the feature selection was guided by the study’s focus on capacity-driven demand changes and by pragmatic choices to ensure robust modeling. As noted in the methodology, limiting the model to core variables related to supply and demand has advantages: it allows a clearer analysis of the relationship between capacity and stimulated demand, avoids the noise that many extraneous variables might introduce, and aligns with variables used in similar forecasting studies (facilitating benchmarking). In other words, the inputs are largely industry-standard factors that any airline would consider when evaluating a market. For example, “Flights” and “Seats” are fundamental to capacity planning, and “Blended Fare” is a commonly tracked metric indicating market pricing. By using these, the model’s outputs can be interpreted in industry terms – e.g., “if 10% more seats are added, passenger volume is forecasted to rise by X%.” This industry orientation improves the applicability of the results: an airline could apply the model to its routes using those same inputs (readily obtainable from schedules and ticket data) without needing special proprietary data.

Limitations of the Study

Several limitations should be acknowledged. First, the modeling and results are based on a dataset specific to the Canadian international air travel market. The patterns learned by the models (and their performance levels) may not directly generalize to other regions or domestic markets that have different competitive dynamics, travel regulations, or passenger behavior. The underlying methodology is transferable, but retraining and validation would be needed for other contexts. Second, the best-performing models (GRU and LSTM) operate essentially as black-box predictors. They do not provide easily interpretable parameters or coefficients that explain how much each factor (e.g., fare or capacity) influences demand. This lack of interpretability contrasts with classical models (like regression-based approaches) and can pose challenges for adoption, as decision-makers often prefer models that offer transparent insights into cause and effect. Third, the feature set used in this study, while tailored to capture capacity changes and lagged demand effects, is limited in scope. Important external factors such as macroeconomic indicators, competitive actions (e.g., rival airlines' capacity on the same routes), or marketing efforts were not included. The `New_Route_Indicator` variable, implemented as a simple binary flag for the first few months of a route, provides only a coarse representation of market stimulation; it does not account for how the effect might build up or decay over time. These omissions mean the model might miss some drivers of demand or oversimplify certain effects. Expanding the feature space to include such factors could further improve model accuracy and applicability.

Conclusion

This thesis investigated the challenge of forecasting air passenger demand at the route level, with a focus on capturing the market stimulation effect that occurs when airlines introduce new routes or increase capacity. By developing a suite of deep learning models and evaluating their performance, the study demonstrated that Gated Recurrent Unit (GRU) networks offer exceptional accuracy for this task. The key insight is that models which process data sequentially and remember long-term information (like GRUs and LSTMs) markedly outperform those that do not (ANNs and CNNs). In practical terms, the GRU model was able to learn the complex relationship between added capacity (supply-side changes) and the resulting passenger volumes (demand-side response), even under volatile conditions. This represents a shift from traditional demand forecasting approaches that rely heavily on exogenous demand drivers such as GDP, economic growth, or population. Instead, the present work emphasizes capacity as an active factor in shaping demand. The success of the GRU model in detecting and quantifying stimulated demand suggests

that airlines can use capacity deployment not just as a reaction to demand, but as a lever to influence demand.

The contributions of this research are both methodological and practical. Methodologically, it introduces a deep learning framework that integrates flight schedule data with passenger trends, using engineered features (lags, growth rates, and new-route indicators) to bridge the supply-demand nexus. This approach attained high predictive performance (with errors around 5% MAPE for the best model) on a challenging dataset that includes unprecedented demand disruptions. Practically, the findings provide airline planners with a validated tool for forecasting route performance. A GRU-based model can help quantify expected passenger gains from adding a new route or extra flights, supporting data-driven decision-making in network strategy and capacity management. The fact that the model maintained solid performance through the pandemic period adds confidence that it could be deployed in real-world scenarios that include sudden shocks or recoveries.

Looking forward, there are several opportunities to extend this work. Future research could incorporate external data sources to enrich the model's understanding of demand drivers. For instance, macroeconomic indicators (GDP, exchange rates, tourism indices), traveler search and booking trends, or competitor capacity and pricing information could be integrated as additional features. These exogenous factors often influence travel demand and could improve forecast accuracy, especially for anticipating turning points in demand that are not evident in the historical passenger series alone. Additionally, more nuanced feature engineering could be explored. Instead of a binary new-route flag, one could model the duration and diminishing influence of a new route's novelty (for example, using a decaying function or multi-period indicators to represent the stimulation effect over time). Another avenue is the application of advanced model architectures. Transformer-based time series models or hybrid models (combining neural networks with classical time-series components) might capture very long-term dependencies or heterogeneous effects across routes even better. These architectures have shown promise in other sequence prediction domains and could be beneficial when sufficient data is available. Finally, expanding the scope of the study to other markets or finer time granularities (such as weekly or daily data, if accessible) would test the generality of the approach and potentially reveal new insights about demand dynamics at different scales.

To summarize everything that was discussed in this work, this thesis highlights the importance of sequential modeling techniques in forecasting air travel demand at the route level, particularly when assessing the impact of capacity changes. The GRU model emerged as a powerful tool, providing accurate predictions and demonstrating resilience in the face of volatile market conditions. By treating capacity not just as supply, but as a catalyst for demand, the study offers a new perspective on demand forecasting that could be highly valuable for airlines looking to optimize their networks and operations in an ever-changing industry.

7. References

- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2623–2631). ACM. <https://doi.org/10.1145/3292500.3330701>
- Alam, M. S., Deb, J. B., Al Amin, A., & Chowdhury, S. (2024). An artificial neural network for predicting air traffic demand based on socio-economic parameters. *Decision Analytics Journal*, 10, 100382.
- Belobaba, P., Odoni, A., & Barnhart, C. (Eds.). (2015). *The global airline industry*. John Wiley & Sons.
- Birolini, S., Jacquillat, A., Cattaneo, M., & Antunes, A. P. (2021). Airline Network Planning: Mixed-integer non-convex optimization with demand–supply interactions. *Transportation Research Part B: Methodological*, 154, 100-124.
- Boonekamp, Thijs, Joost Zuidberg, and Guillaume Burghouwt. "Determinants of air travel demand: The role of low-cost carriers, ethnic links and aviation-dependent employment." *Transportation Research Part A: Policy and Practice* 112 (2018): 18-28.
- Buscema, P. M., Massini, G., & Maurelli, G. (2014). Artificial neural networks: An overview and their use in the analysis of the AMPHORA-3 dataset. *Substance Use & Misuse*, 49(12), 1555–1568. <https://doi.org/10.3109/10826084.2014.933009>
- Carmona-Benítez, R. B., Nieto, M. R., & Miranda, D. (2017). An Econometric Dynamic Model to estimate passenger demand for air transport industry. *Transportation Research Procedia*, 25, 17-29.
- Carmona-Benítez, R. B., & Nieto, M. R. (2023). A methodology for calculating the unmet passenger demand in the air transportation industry. *Research in Transportation Business & Management*, 50, 101039.
- Coldren, G. M., & Koppelman, F. S. (2005). Modeling the competition among air-travel itinerary shares: GEV model development. *Transportation Research Part A: Policy and Practice*, 39(4), 345-365.
- Coldren, G. M., Koppelman, F. S., Kasturirangan, K., & Mukherjee, A. (2003). Modeling aggregate air-travel itinerary shares: Logit model development at a major US airline. *Journal of Air Transport Management*, 9, 361-369.

- Dantas, T. M., Oliveira, F. L. C., & Repolho, H. M. V. (2017). Air transportation demand forecast through Bagging Holt Winters methods. *Journal of Air Transport Management*, 59, 116-123.
- Emami Javanmard, M., Tang, Y., & Martínez-Hernández, J. A. (2024). Forecasting air transportation demand and its impacts on energy consumption and emission. *Applied Energy*, 364, 123031. <https://doi.org/10.1016/j.apenergy.2024.123031>
- Garrow, L. A., Jones, P. J., & Parker, R. A. (2007). How much airline customers are willing to pay: An analysis of price sensitivity in online distribution channels. *Journal of Revenue and Pricing Management*, 5(4), 271-290.
- Gelhausen, M. C., Berster, P., & Wilken, D. (2018). A new direct demand model of long-term forecasting air passengers and air transport movements at German airports. *Journal of Air Transport Management*, 71, 140–152. <https://doi.org/10.1016/j.jairtraman.2018.04.001>
- Géron, A. (2018). *Neural networks and deep learning (Chapter 4: Recurrent Neural Networks)*. O'Reilly Media. <https://www.oreilly.com/library/view/neural-networks-and/9781492037354/ch04.html>
- Grammig, J., Hujer, R., & Scheidler, M. (2005). Discrete choice modelling in airline network management. *Journal of Applied Econometrics*, 20(4), 467-486.
- Grosche, T., Rothlauf, F., & Heinzl, A. (2007). Gravity models for airline passenger volume estimation. *Journal of Air Transport Management*, 13(4), 175-183.
- Gunter, U., & Zekan, B. (2021). Forecasting air passenger numbers with a GVAR model. *Annals of Tourism Research*, 89, 103252. <https://doi.org/10.1016/j.annals.2021.103252>
- Harshed. (2024, November 29). *Understanding gated recurrent units in deep learning*. Medium. <https://medium.com/@harshedabdulla/understanding-gated-recurrent-units-grus-in-deep-learning-4404599dcefb>
- Hazledine, T. (2017). An augmented gravity model for forecasting passenger air traffic on city-pair routes. *Journal of Transport Economics and Policy (JTEP)*, 51(3), 208-224.
- Hess, S. (2007). Posterior analysis of random taste coefficients in air travel behaviour modelling. *Journal of Air Transport Management*, 13(4), 203-212.
- Hu, Y.-C. (2023). Air passenger flow forecasting using nonadditive forecast combination with grey prediction. *Journal of Air Transport Management*, 112, 102439. <https://doi.org/10.1016/j.jairtraman.2023.102439>

- Hsiao, C.-Y., & Hansen, M. (2011). A passenger demand model for air transportation in a hub-and-spoke network. *Transportation Research Part E: Logistics and Transportation Review*, 47(6), 1112-1125.
- Jin, F., Li, Y., Sun, S., & Li, H. (2020). Forecasting air passenger demand with a new hybrid ensemble approach. *Journal of Air Transport Management*, 83, 101744.
- Jafari, N., & Lewison, M. (2024). Forecasting air passenger traffic and market share using deep neural networks with multiple inputs and outputs. *Frontiers in Artificial Intelligence*, 7, 1429341. <https://doi.org/10.3389/frai.2024.1429341>
- Kanavos, A., Kounelis, F., Iliadis, L., & Makris, C. (2021). Deep learning models for forecasting aviation demand time series. *Neural Computing and Applications*, 33(16), 16329–16343. <https://doi.org/10.1007/s00521-021-06232-y>
- Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2020). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151, 107398. <https://doi.org/10.1016/j.ymssp.2020.107398>
- Koç, İ., & Arslan, E. (2018, October). Demand forecasting for domestic air transportation in Turkey using artificial neural networks. In 2018 6th International Conference on Control Engineering & Information Technology (CEIT) (pp. 1-6). IEEE.
- Kölker, K., Lütjens, K., & Gollnick, V. (2024). Analyzing global passenger flows based on choice modeling in the air transportation system. *Journal of Air Transport Management*, 115, 102530.
- Koppelman, F. S., Coldren, G. M., & Parker, R. A. (2008). Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B: Methodological*, 42(3), 263-273.
- Liang, X., Hong, C., Chen, J., Wang, Y., & Yang, M. (2024). A hybrid forecasting architecture for air passenger demand considering search engine data and spatial effect. *Journal of Air Transport Management*, 118, 102611. <https://doi.org/10.1016/j.jairtraman.2024.102611>
- Long, C. L., Guleria, Y., & Alam, S. (2021). Air passenger forecasting using Neural Granger causal Google trend queries. *Journal of Air Transport Management*, 95, 102083.
- Lingrui, L., & Xin, W. (2024). Towards smart aviation with sustainable development: artificial intelligence insights into the airline and advanced air mobility industries. *Decision Support Systems for Sustainable Computing*, 187-204.
- Mostafaiepour, A., Goli, A., & Qolipour, M. (2018). Prediction of air travel demand using a hybrid artificial neural network (ANN) with Bat and Firefly algorithms: a case study. *The journal of supercomputing*, 74(10), 5461-5484.

- Munoz, C., Laniado, H., & Córdoba, J. (2020). Development of a robust customer satisfaction index for domestic air journeys. *Research in Transportation Business & Management*, 37. <https://doi.org/10.1016/j.rtbm.2020.100519>
- Ramadhani, S., Dhini, A., & Laoh, E. (2020, November). Airline Passenger Forecasting using ARIMA and Artificial Neural Networks Approaches. In 2020 International Conference on ICT for Smart Society (ICISS) (pp. 1-5). IEEE.
- Saâdaoui, F., Saadaoui, H., & Rabbouch, H. (2020). Hybrid feedforward ANN with NLS-based regression curve fitting for US air traffic forecasting. *Neural Computing and Applications*, 32(15), 10073–10085. <https://doi.org/10.1007/s00521-019-04539-5>
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems* (pp. 802–810).
- Sun, S., Lu, H., Tsui, K. L., & Wang, S. (2019). Nonlinear vector auto-regression neural network for forecasting air passenger flow. *Journal of Air Transport Management*, 78, 54-62.
- Tillmann, A. M., Joormann, I., & Ammann, S. C. (2023). Reproducible air passenger demand estimation. *Journal of Air Transport Management*, 112, 102462.
- Truong, D. (2021). Estimating the impact of COVID-19 on air travel in the medium and long term using neural network and Monte Carlo simulation. *Journal of Air Transport Management*, 96, 102126.
- Wang, S., & Gao, Y. (2021). A literature review and citation analyses of air travel demand studies published between 2010 and 2020. *Journal of Air Transport Management*, 97, 102135.
- Wei, K., Vaze, V., & Jacquillat, A. (2020). Airline Timetable Development and Fleet Assignment Incorporating Passenger Choice. *Transportation Science*, 54(1), 139–163. <https://doi.org/10.1287/trsc.2019.0924>
- Xiao, Y., Liu, J. J., Hu, Y., Wang, Y., Lai, K. K., & Wang, S. (2014). A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting. *Journal of Air Transport Management*, 39, 1-11.
- Zachariah, R. A., Sharma, S., & Kumar, V. (2023). Systematic review of passenger demand forecasting in aviation industry. *Multimedia tools and applications*, 82(30), 46483-46519.
- Zhou, H., Xia, J. C., Luo, Q., Nikolova, G., Sun, J., Hughes, B., ... & Falkmer, T. (2018). Investigating the impact of catchment areas of airports on estimating air travel demand: A case study of regional Western Australia. *Journal of Air Transport Management*, 70, 91-103.