Airline Scheduling and Optimization: A Systematic Literature Review and Bibliometric Analysis

[1]Dr Richa Kalpesh Saxena, [2]Anvi Paliwal, [3]Aritra Chakraborty, [4]Arnav Aggarwal,

[5]Arush Garg, [6]Aryan Chauhan

[1]Assistant Professor, Marketing & Operations, NMIMS, Anil Surendra Modi School of Commerce, Mumbai Campus, India

[2,3,4,5,6]BBA, NMIMS, Anil Surendra Modi School of Commerce, Mumbai Campus, India

Abstract

This study presents a systematic literature review-cum-bibliometric analysis of research on airline scheduling and optimisation. Following the PRISMA framework, 2,828 records retrieved from Elsevier's Scopus database were refined by time span, subject area, document type, language and source type to a final dataset of 1,045 journal articles and review papers published between 2011 and 2026. The data were analysed using RStudio and Biblioshiny, and a VOSviewer cluster analysis of a separate RIS database supported a thematic literature review. Annual publications rose from 26 in 2011 to 113 in 2026, with China and the United States contributing the largest share of country records. Revenue management, fleet assignment, aircraft routing and integer programming emerged as motor themes, whereas machine learning and deep learning appeared less connected to the core of the field. The thematic review identified three interconnected clusters and a gap in coordination among airlines, airports and air traffic management under uncertainty. A future research agenda is proposed using the 5Ws1H and Input–Process–Output frameworks.

Keywords: Airline Scheduling; Optimisation; Bibliometric Analysis; Air Transportation; Systematic Literature Review.

1. Introduction

Air transportation is one of the most operationally complex systems in the transport sector. Airlines, airports, aircraft, passengers and traffic management processes are linked into a single coordinated chain, so a change in any one element can create disruptions elsewhere. Effective use of aircraft, crews, gates, slots and airspace therefore depends on planning decisions that are made well in advance and revised continuously. Optimisation has become fundamental to this task. It provides the mathematical and computational basis for allocating scarce resources in a way that balances cost, efficiency, reliability and service quality. These characteristics make air transportation a natural, and demanding, setting for optimisation research.

Airline scheduling sits at the heart of this planning problem. It covers flight timetabling, fleet assignment, aircraft routing, crew pairing and rostering, and the coordination of airport capacity, and these decisions are strongly interdependent. A change to fleet assignment, for example, can simultaneously affect routing, capacity and scheduling decisions. Scheduling also interacts with the commercial side of airline operations.

Timetables and capacity determine what can be sold, while pricing and revenue management in turn influence which schedules are profitable. For this reason, airline scheduling is a key issue not only in operations research but also in airline management and transport planning.

In recent years, there has been an increasing interest in airline scheduling and optimisation. The dataset analysed in this study shows annual publications rising from 26 articles in 2011 to 113 in 2026, an increase of about 335%, with a peak of 122 articles in 2025. Average annual output grew from 35.6 articles in 2011–2015 to 58 in 2016–2020 and 95.5 in 2021–2026. This growth has been accompanied by an international expansion of the field. The country-level data covers 68 countries, with China (639 records) and the United States (344 records) the largest contributors. Such growth suggests that airline scheduling has become an established and expanding area of scholarly enquiry, and it makes a systematic assessment of the literature both timely and necessary.

Previous research has addressed many parts of the scheduling problem. In airline operations, studies increasingly treat interrelated decisions jointly. Zhu (2026) examines fleet scheduling together with cargo flow allocation, while Chen (2026) integrates aircraft routing and fleet assignment using a decision-unit approach designed to reduce model complexity.

Crew scheduling has received particular attention, with work on crew pairing using hybrid metaheuristics (Devi, 2026), on pilot fatigue (Ong, 2026) and, more broadly, a meta-review of aircrew scheduling and its research challenges (Sun, 2026). A parallel body of work addresses airports and traffic management, including network-level slot allocation (Melder et al., 2026), airport gate assignment (Ding, 2026) and congestion management using reinforcement learning (Qiao et al., 2026). Methodologically, exact approaches such as mixed-integer programming remain central. However, metaheuristics, machine learning and even quantum computing are increasingly used to cope with computational complexity (Hızır et al., 2026; Du, 2026).

Recent studies also reflect a shift in how scheduling problems are framed. Uncertainty and disruption have become prominent concerns, with research on robust scheduling (Guardo-Martinez, 2026), aircraft routing under stochastic delays (Li, 2026a) and integrated recovery under uncertain disruptions (Abdelghany & Abdelghany, 2026; Wu et al., 2026). Objectives have broadened beyond cost to include sustainability, as in the selection of electric-flight airports (Joormann et al., 2026) and trajectory optimisation that accounts for contrails (Demouge et al., 2026). Passenger behaviour is also being incorporated into airline planning models (Papi-RAD et al., 2026). Taken together, these studies suggest that the field is moving towards integrated, adaptive and multi-objective approaches, although these developments are reported in separate strands of work rather than within a single framework.

The literature reviewed for this study also points to a structure of three interconnected clusters. The first brings together air transportation, optimisation and decision-support; the second concentrates on computational and mathematical methods such as integer programming and combinatorial optimisation; and the third covers airport operations and air navigation management. The close links between these clusters indicate that researchers increasingly combine different optimisation approaches across several areas of aviation decision-making. At the same time, major challenges remain, including computational complexity, real-world uncertainty, the data and validation requirements of machine learning, and conflicts between objectives such as cost, fairness and environmental performance.

Despite this growth, the literature has become fragmented. Research is dispersed across sub-problems, methods and journals, ranging from aviation-specific outlets such as the Journal of Air Transport Management to broader operations research and transportation journals. Existing reviews tend to concentrate on a single area, such as aircrew scheduling (Sun, 2026) or airport ground operations (Dahanayaka, 2026). Consequently, the overall intellectual structure of the field remains unclear. It is not well understood which themes are central and which are peripheral, how machine learning relates to the optimisation methods at the core of the field, or how research on airlines, airports and air traffic management connects. To our knowledge, few studies have mapped airline scheduling as a whole using systematic bibliometric methods. A broader and more structured picture is needed to help researchers position new work and identify where coordination across stakeholders, uncertainty and computational complexity remain unresolved.

A systematic literature review is a well-established approach to assessing and synthesising an existing body of research (Paul & Rialp-Criado, 2020). Bibliometric analysis extends this approach by quantifying patterns in publication data, which makes it particularly suitable for a field that is large, growing and spread across many outlets. It allows the evolution of a research area to be traced objectively, and it reveals relationships between authors, countries and themes that a narrative review alone may overlook.

This study therefore reviews the literature on airline scheduling and optimisation using a bibliometric approach, supported by a thematic literature review. It addresses the following research questions:

- **RQ1:** How has publication output on airline scheduling changed over time?
- **RQ2:** Which countries, authors and sources contribute most to the field?
- **RQ3:** What are the dominant, niche and emerging research themes, and how do they relate to one another?
- **RQ4:** What gaps and future research directions does the literature suggest?

To answer these questions, we adopted a structured bibliometric approach based on the PRISMA framework (Moher et al., 2010), which allows the stages of identification, screening, eligibility and inclusion to be reported transparently. Records were retrieved from Scopus using the keywords “airline” AND “optimization”, which returned 2,828 records. These were refined by time span, subject area, document type, language and source type, leaving a final dataset of 1,045 journal articles and review papers. The data were analysed using RStudio and Biblioshiny, covering annual scientific production, country and source productivity, keyword and author relationships through a three-field analysis, and a

thematic map that classifies themes as motor, niche, emerging or declining. A VOSviewer cluster analysis of a separate RIS database complements these results and provides the basis for the thematic literature review.

This study contributes to the field in three ways. First, it offers a systematic overview of how airline scheduling research has developed, showing which publication trends, international research networks and methods have shaped it. Second, it locates established themes, such as fleet assignment, aircraft routing and integer programming, alongside less developed ones, such as machine learning and deep learning, which appear to be less closely connected to the core of the field. Third, it identifies gaps that can guide future work, notably the coordination of decisions among airlines, airports and air traffic management under uncertainty. In this way, the study enhances our understanding of how airline scheduling has evolved from a mainly mathematical optimisation problem into a broader research area spanning aircraft operations, air traffic management, pricing and decision-making. It is hoped that the findings will be useful to researchers seeking to identify promising directions, and to practitioners interested in how optimisation methods are being applied across airline and airport operations.

The scope of the study should be noted. The analysis is based solely on Scopus-indexed, English-language journal articles and review papers, and it excludes other document types such as conference papers. A bibliometric approach also describes patterns in the literature rather than evaluating the quality or performance of individual optimisation methods. A full comparison of algorithms lies beyond the scope of this study. The findings should therefore be read as a map of the literature as indexed in Scopus rather than a complete record of all work on airline scheduling.

Throughout this paper, the term “airline scheduling” is used in its broad sense to refer to the planning and optimisation of airline resources, including flights, aircraft, crews and the associated airport capacity. The term “bibliometric analysis” refers to the quantitative analysis of publication data to identify patterns of growth, authorship, collaboration and themes within a research field.

2. Literature Review Overview of the Thematic Clusters

Air transportation is a multi-faceted system where airlines, airports, aircraft, passengers, and traffic management processes are integrated together into one coordinated chain. This implies that changes in any single element have the potential to create disruptions elsewhere, and optimisation is extremely important to ensure effective resource allocation. The reviewed literature in the provided RIS database demonstrates a trend in the field towards addressing integrated optimisation and decision-support models for air transport rather than individual issues of scheduling or routing. The VOSviewer analysis reflects this situation by illustrating three large, interconnected clusters of publications. The first includes topics of air transportation, optimisation, and decision-support, while the other two are focused on computational/mathematical methods of optimisation and airport operations/air navigation management. The fact that these three clusters are so closely interconnected suggests that researchers are now paying significantly more attention to the combined use of different optimisation approaches and areas of air transportation decision-making processes.

Airline Operations, Crew Scheduling and Disruption Recovery

One of the main themes is the optimisation of airline operations. Airline planning is characterised by interrelated decisions on fleet assignment, aircraft routing, flight scheduling, crew scheduling and cargo allocation. Zhu (2026) for example, studies the joint fleet scheduling and cargo flow allocation in air cargo services while Chen (2026) looks at aircraft routing and fleet assignment via decision-unit approach aimed at reducing the model complexity. These works are indicative of the trends in which decisions related to these areas are increasingly regarded as interrelated decisions that should be optimised jointly. The reason is that in fleet assignment changes in one decision may affect routing, capacity and scheduling decisions simultaneously.

Crew scheduling is another interesting area in which optimisation techniques can be applied. It is known that airline crew planning is a challenging task that must satisfy numerous regulatory and operational constraints while optimising a vast solution space. Devi (2026) suggests a hybrid jumping particle swarm optimisation with local search for airline crew pairing. At the same time, Wang (2026b) explores integrated crew scheduling with fairness considerations based on the crew’s familiarity and experience, while Ong (2026) addresses the issue of pilot fatigue in crew scheduling. Finally, Sun (2026) conducts a meta-review of the aircrew scheduling optimisation research, highlighting the challenges of the problem area. By exploring multiple optimisation issues related to airline crews, these researchers demonstrate that this field has multiple aspects that go beyond mere cost reduction and became a subject of studies addressing fairness, fatigue, and efficiency.

A related development is an increased focus on uncertainty and disruptions. Aviation operations are subject to delays, weather-related events, variance in turnaround-time and unforeseen operational occurrences. Guardo-Martinez (2026)

explores robust airline scheduling under uncertainty while Li (2026a) considers aircraft routing with stochastic delays. Assink (2026) focuses on ground-staff shift and break scheduling in the face of flight delays. These studies highlight the need to develop optimisation models that are capable of maintaining their effectiveness in the face of discrepancies between actual and planned operational conditions.
This issue can be seen particularly in the context of airline recovery research.

Abdelghany and Abdelghany (2026) address the complexity of the system of airline irregular operations and integrated schedule recovery. Wu et al. (2026) study integrated airline recovery under uncertain disruptions with fuzzy programming. In addition, the question of using machine learning in making recovery decisions is also being explored. Xiao et al. (2026) identify features relevant to aircraft recovery informed by machine learning, and Hızır, Barnhart and Vaze (2026) combine mixed-integer optimisation and supervised machine learning for large-scale crew recovery. It could be argued that these two areas of research are complementary; that is, machine learning can help find patterns or make predictions while optimisation can be used to find allocations of resources.
The first VOSviewer cluster can be explained by the fact that terms such as “air transportation,” “optimisation,” “decision making,” “airline industry,” “machine learning,” “recovery” and “forecasting” are closely related. This suggests that optimisation is now regarded as an integral part of the decision-making process. It can be seen that prediction, scheduling, and recovery are integrated into a comprehensive approach rather than viewed as separate areas of study.

Computational and Mathematical Optimisation Methods

The second group focuses on computational and mathematical optimisation. Terms such as integer programming, computational efficiency and combinatorial optimisation indicate the level of complexity of aviation issues. Many decisions are discrete: an aircraft is assigned to a flight or not, a crew member to a pair or not and an airport slot is assigned to one flight rather than another. When thousands of such decisions interact, the number of possible solutions is very large.
For its part, mathematical programming remains a key area of research: Hızır et al. (2026) address the problem of crew recovery using mixed-integer optimisation, and Yuan, Zhou, and Ribeiro (2026) formulated the problem of storage assignment in an airport cargo-terminal as a temporal-spatial packing. Melder et al. (2026) investigate the network-level airport slot allocation using a novel formulation and a matheuristic solution approach. Thus, we can see that the usage of the exact and mathematical programming methods allows researchers to find a way to represent their complex real-life problems.

At the same time the increased computational complexity has led to the use of heuristics and meta-heuristics. Devi (2026) uses Particle Swarm Optimisation and a local search to solve the crew pairing problem. Meanwhile, Ding (2026) combines Particle Swarm Optimisation and Tabu Search for the airport gate assignment problem. Jia, Guo and Mei (2026) combine a self-adapting memetic algorithm with reinforcement learning to solve the multi-objective aircraft recovery problem. Demir and Dervis (2026) apply evolutionary and swarm-based algorithms in the synchronisation of arrivals and departures at hub airports.
Such approaches are valuable since they can explore large solution space without having to find an exact solution.
The literature also reflects an increasing interest in improving computational efficiency using innovative technologies. For example, in his article, Du (2026) explores the application of quantum computing in the field of air transportation. Gili, San Sebastian and Blázquez-García (2026) take a similar approach, analysing the use of a quantum approximate optimisation algorithm for solving aircraft tail assignment problems. This indicates a tendency in the field to explore alternative approaches rooted in new emerging technologies and asking if they can offer any advantages in the context of complex optimisation tasks that prove too computationally intensive in other frameworks.

Airports, Air Navigation and Air Traffic Management

The third VOSviewer cluster is concerned with airports, air navigation and air traffic management. Airports represent complex systems in which aircraft movements, gates, baggage, passengers, ground resources, and many other factors have to be coordinated.
Dahanayaka (2026) provides a review of airport ground-operations optimisation, while Wang (2026a) concentrates on integrated airport slot allocation and boarding-gate assignment.

Melder et al. (2026) examine slot allocation at the network level and Bao, Chu and Li (2026) incorporate airport and airline fairness in regional multi-airport slot allocation. Thus, airport optimisation studies increasingly reflect multiple resource and stakeholder considerations.

Congestion and air traffic flow management are also crucial and important. Qiao et al. (2026) develop a multi-agent real-time pricing mechanism for airport congestion management based on distributed reinforcement learning. Wang, Wang and Hansen (2026) investigate the distributional multi-agent reinforcement learning framework for air traffic flow and capacity management of multiple airports. Meanwhile, airspace and trajectory optimisation are also becoming increasingly dynamic. Borndörfer, Do and Hoang (2026) explore the vertical free-flight optimisation, Shihua, Nguyen and Liem (2026) consider aircraft trajectory optimisation in real-time with varying winds. All of these works embody that planning and decision-making of air traffic management systems gradually change from traditional fixed to new adaptive forms.

Widening Objectives and Passenger Behaviour

Another important trend is that of optimising objectives. While traditional approaches to the problems of commercial aviation focused on cost, time or other narrow factors, modern researchers frequently consider a combination of different objectives. Joormann, Kroner and Tillmann (2026) have proposed an optimisation method for hybrid-electric aircraft that chooses airports to serve based on their ability to support electric flights. Gurturk and Celik (2026) examine optimisation of air-cargo routes with a focus on sustainability, while Demouge, Guitart and Delahaye (2026) propose a similar method for general aviation that takes both fuel-burn and climate impact of contrails into account. Clearly, the focus of flight-optimisation research spans beyond mere economy.

In addition to the optimisation of internal resources, researchers have begun to focus more on passenger behaviour as well as market uncertainty. Papi-RAD, Pishvaee and Sahebi (2026) propose an airline planning model that considers both market uncertainty and passenger choice by using distributionally robust optimisation. Meanwhile, Li et al. (2026b) analyse the flight supply-demand relationship taking into account passenger choice as well. These studies aim to shift the emphasis from optimising internal elements to decision-making process associated with passengers' reactions.

Challenges, Research Gap and Synthesis

Across the three clusters, there is therefore also a common pattern: aviation optimisation is becoming more integrated, adaptive and multi-objective. However, there are also major challenges ahead. The sheer computational complexity of large models makes it challenging to find good solutions quickly. Real-world uncertainty makes solutions that work well under particular assumption not necessarily perform well in practice. Machine learning and reinforcement learning approaches raise issues with respect to data requirements, interpretability and validation. Finally, multi-optimisation objectives such as cost, efficiency, fairness, passenger satisfaction and environmental performance can often be at odds with each other.

A further research gap concerning coordination of actions among various aviation stakeholders was identified. While airlines, airports, air traffic management organisations and other entities operate within the same system there are often conflicting goals and levels of information between them. Some studies have begun to address this challenge by considering integrated recovery, network-level slot allocations and collaborative scheduling. Future work could build on these studies by developing system-level models which coordinate decision-making across stakeholders while taking into account factors such as uncertainty, sustainability and computational complexity.

In general, the RIS literature highlights the interdisciplinary nature of the field of transportation optimisation. The VOSviewer clusters emphasise the tight interrelations among decision-making processes in airlines, advanced optimisation methodologies, and airport and traffic management. Mathematical programming remains a dominant force, but its role is increasingly supplemented by metaheuristics, machine learning, and reinforcement learning. Sustainability, uncertainty, passenger behaviour, and equity are also gaining ground. Thus, the research frontier is defined by the design of integrated and adaptive decision-support systems for managing large-scale and complex air transportation networks. Another key finding is that the development of aviation transportation optimisation is shifting toward more data-driven and adaptive solution approaches. The literature does not identify a universal algorithm for optimising air transportation; instead, it emphasises the importance of context-specific algorithms that consider the scope of the problem, sources of uncertainty, objectives, and available resources. This finding underscores the need for further benchmark testing and experimentation in realistic conditions to evaluate the performance of different

optimisation algorithms.

3. Research Methodology

One of the established approaches to assessing and synthesising an existing body of literature is a systematic literature review (Paul & Rialp-Criado, 2020). For this study, we applied a structured bibliometric approach to reviewing the existing research about airline scheduling. We have used the PRISMA framework as a way to provide a systematic and transparent approach to identification and selection of relevant publications (Moher et al., 2010). PRISMA provides a framework for reporting the process which was followed in conducting systematic reviews and allows for the presentation of the stages of identification, screening, eligibility, and inclusion in a clear manner.

The database search was conducted using Scopus. We have used the keywords “airline” AND “optimization” to identify publications relevant to the research topic. The initial search retrieved 2,828 records. The purpose of this stage was to establish an initial pool of potentially relevant publications before applying the inclusion and exclusion criteria. This approach enabled us to begin with a broad dataset and subsequently refine the records according to the scope of the study.

Article Selection, Inclusion and Exclusion Criteria

The first stage of the selection process entailed the application of a time-span criterion. We selected the period between 2011 and 2027. After the application of this criterion, the number of the records was reduced from 2,828 to 2,230. The selected period was intended to provide a defined scope for the analysis, and to include publications falling within the period relevant to the present study.

The next stage of the selection process entailed the application of a subject-area filter to the remaining records. We selected the subject areas which were considered relevant to airline scheduling and optimisation. These included Engineering, Computer Science, Mathematics, Business, Management and Accounting, and Economics, Econometrics and Finance. The inclusion of these subject areas enables us to focus the analysis on the technical, mathematical, managerial and economic dimensions related to airline scheduling. After the application of the subject-area criterion, 2,006 records remained.

The third stage entailed the application of a document type filter. For the purpose of this study, we restricted the dataset to Articles and Review Papers. Other document types such as conference papers were excluded from the analysis. We selected Articles and Review Papers, as these forms of publication were seen to be more appropriate for the objectives of the present study. In particular, we considered that the availability of detailed bibliographic information, including author and researcher information, would be important for the subsequent bibliometric analysis. After the application of the document type filter, the dataset was reduced from 2,006 to 1,161 records.

The next stage of the selection process involved the application of a language criterion. We confined the dataset to publications written in English. This criterion was applied to maintain consistency in the analysis, and to facilitate interpretation and comparison of the selected literature. After the application of the language filter, the number of records was reduced from 1,161 to 1,050.

Finally, we applied a source-type filter. We confined the dataset to journal publications, and excluded trade journals. We considered journal publications to provide more appropriate scholarly information for the objectives of the study. This criterion was also applied to maintain consistency in the type of publications in the final dataset. After the application of the journal filter, the final search size was reduced from 1,050 to 1,045 documents.

Data Collection and Analysis

After the selection process, we prepared the final dataset for bibliometric analysis. A number of software tools were used to organise, manage and analyse the selected publications. RStudio and Biblioshiny were used for conducting bibliometric analysis and generating relevant analytical outputs. These tools allowed us to analyse patterns and relationships within the selected body of literature.

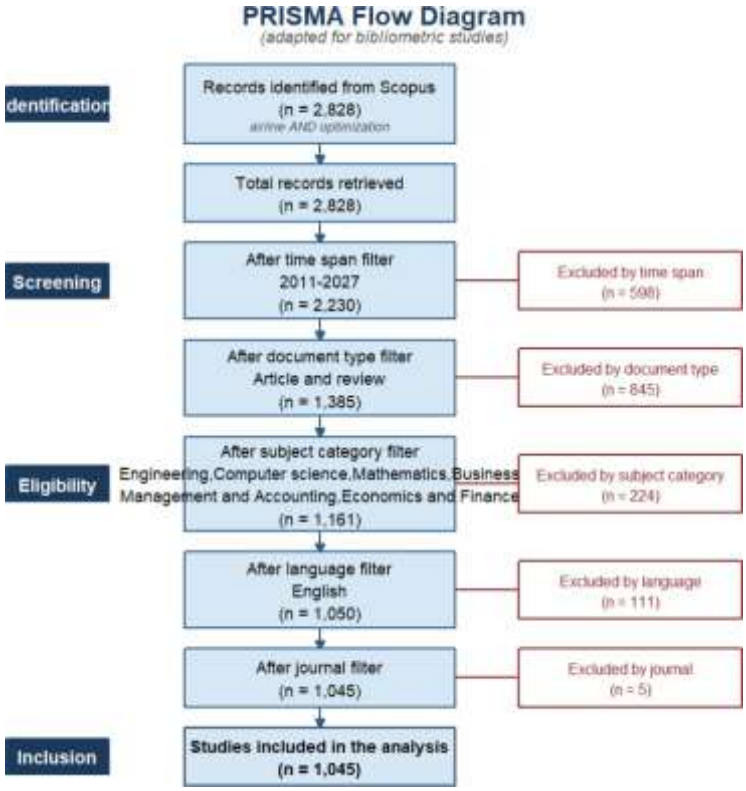
Mendeley was used for reference management and organisation of selected academic publications. Microsoft Excel was used to organise and manage the bibliographic information, while Microsoft Word was used for documentation and preparation of the research report. The use of these tools allowed us to organise the data systematically, and to facilitate the subsequent stages of analysis.

PRISMA-Based Selection Process

The PRISMA flow diagram (Figure 1) was used to provide a clear and transparent representation of the literature-selection process. It documents the number of publications retained at each stage and allows us to demonstrate how the final dataset was constructed. The use of a structured selection process also provides greater clarity regarding the inclusion and exclusion criteria applied throughout the study (Moher et al., 2010).

Figure 1

PRISMA Flow Diagram of the Screening Process (Author's Work)



The entire process may be summarised in the following steps: A total of 2,828 records were identified in the Scopus database using the keywords "airline" AND "optimization." After applying the 2011–2027 time-span criterion, 2,230 records were left. The next step involved another filter, the subject-area, which comprised such fields as Engineering, Computer Science, Mathematics, Business, Management and Accounting, and Economics, Econometrics and Finance. This brought the number of records to 2,006.

The document type screening was subsequently applied, with only Articles and Review Papers being retained. This reduced the dataset to 1,161 records. The language criterion was then applied, restricting the dataset to English-language publications and resulting in 1,050 records. Finally, a source-type criterion was applied that limited the search to journal publications. A total of 1,045 documents were left at this stage and formed the final dataset used for bibliometric analysis.

4. Data Analysis and Results

The data analysis shows that research on airline scheduling has expanded considerably over the period covered by the dataset. The analysis focuses on publication growth, geographical distribution, research themes, major authors, keywords, and the relationship between these areas.

Figure 2 summarises the main information of the final dataset as reported by Biblioshiny.

Figure 2

Main Information of the Final Dataset (Biblioshiny Output)



Publication Growth Over Time

The annual publication data shows a clear long-term increase in research activity related to airline scheduling. In 2011, the dataset recorded 26 articles. This increased to 50 articles by 2014, although publication output fell to 34 articles in 2015. From 2016 onward, the overall trend became stronger.

Publication activity reached 83 articles in 2020 and remained relatively high in the following years. The number increased to 90 articles in 2023 and then rose sharply to 122 articles in 2025, the highest value in the dataset. In 2026, the dataset records 113 articles.

Overall, publications increased from 26 in 2011 to 113 in 2026, representing an increase of about 335%. The average number of publications also rose across the different periods. The average was 35.6 articles per year between 2011 and 2015, compared with 58 articles between 2016 and 2020 and 95.5 articles between 2021 and 2026.

The growth was not completely steady. Several years show declines, particularly 2015, 2018, 2022, 2024, and 2026. The largest year-on-year increase occurred in 2025, when publications increased from 83 to 122. This suggests that research interest in airline scheduling has grown strongly, while still showing short-term fluctuations.

Geographical Distribution of Research

The country-level data contains 2,147 records across 68 countries. China has the highest contribution, with 639 records, followed by the United States with 344. The United Kingdom, Turkey, Germany, India, Iran, France, Australia, and Spain are also among the leading contributors.

China alone accounts for approximately 29.8% of the country-level records. China and the United States together account for about 45.8%. The top five countries account for approximately 59.9% of the total records, showing that research production is concentrated among a relatively small group of countries (Figure 3).

Figure 3

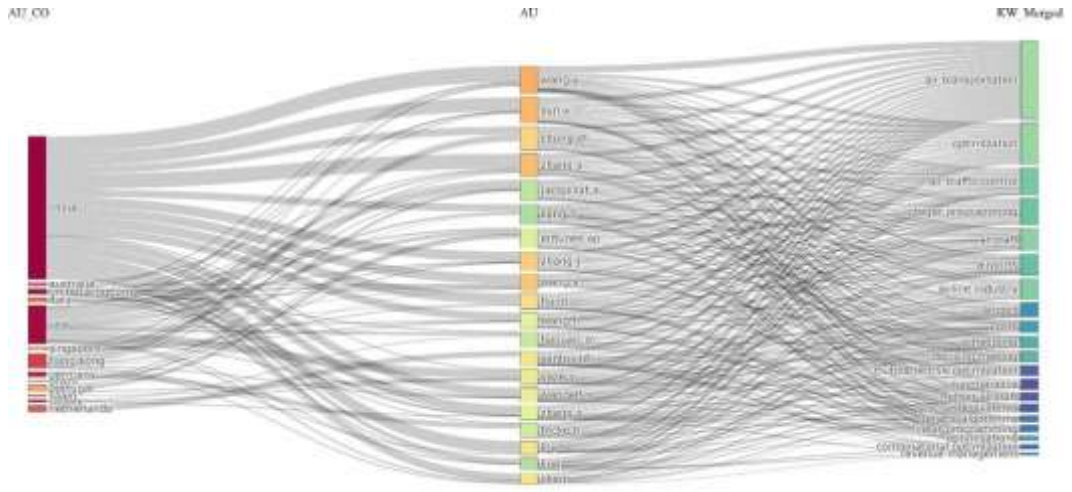
Geographical Distribution of Records by Country (Author's Work)



The three-field analysis (Figure 4) supports this pattern. China has the strongest connection to the author network shown in the dataset, with several Chinese researchers appearing among the prominent authors. Researchers such as Wang Y, Sun X, Chung SH, and Zhang Y appear prominently in the author section. Their work connects with keywords such as air transportation, optimisation, air traffic control, integer programming, aircraft, and scheduling.

Figure 4

Three-Field Plot of Countries (AU_CO), Authors (AU) and Keywords (KW_Merged) (Biblioshiny Output)



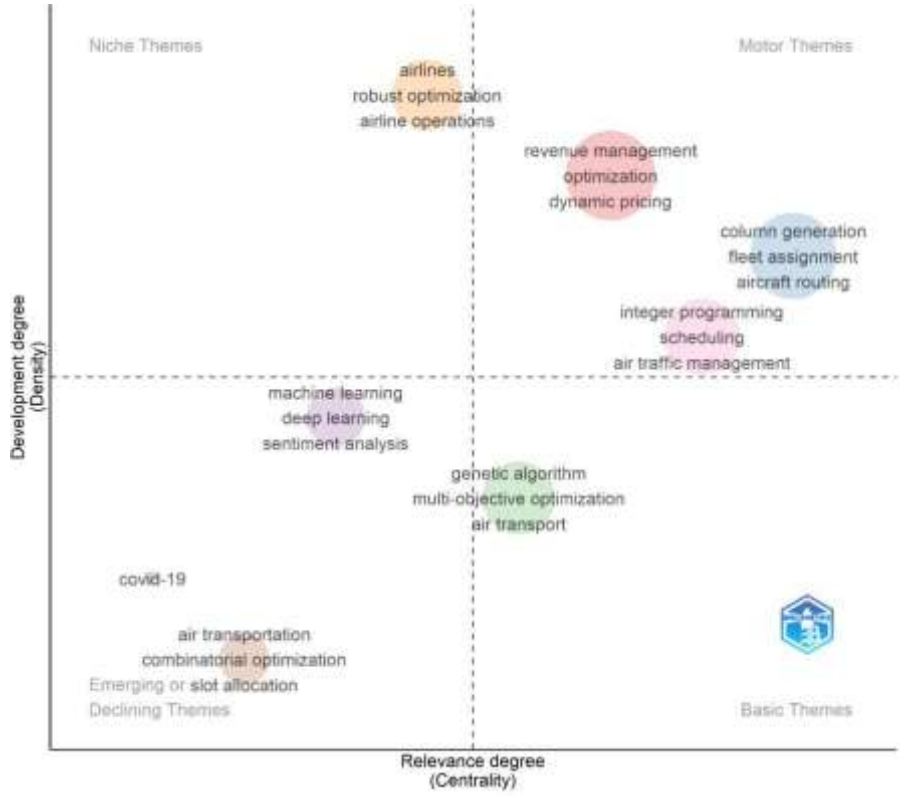
This geographical pattern indicates that airline scheduling research has developed through strong research communities in China and the United States, while contributions also come from Europe, Asia, the Middle East, and other regions.

Research Themes and Their Development

The thematic map (Figure 5) provides a clearer picture of what researchers are studying within airline scheduling.

Figure 5

Thematic Map on the Basis of Relevance and Development of Topic (Biblioshiny Output)



Several themes appear as motor themes, meaning they have both strong relevance to the wider research area and relatively high development within the dataset. Revenue management, optimisation, and dynamic pricing form one important cluster. These topics show that airline scheduling research is closely connected with the financial side of airline operations. Scheduling decisions increasingly interact with pricing and revenue management rather than being treated as completely separate problems.

Another important cluster includes column generation, fleet assignment, and aircraft routing. These topics are directly connected to the practical problem of deciding which aircraft should operate which routes and how aircraft should be allocated efficiently. Their position on the thematic map indicates that they have an established connection with the broader airline scheduling research field.

Integer programming, scheduling, and air traffic management form another significant cluster. Integer programming is particularly relevant because airline scheduling involves many decisions that must be represented as discrete choices, such as assigning aircraft, selecting routes, allocating time slots, and constructing feasible schedules.

The thematic map also identifies airlines, robust optimisation, and airline operations as a niche theme. These topics are more specialised but show a relatively high level of development. Robust optimisation is especially relevant to airline scheduling because schedules have to deal with uncertainty caused by delays, weather, aircraft availability, and changes in demand.

Emerging and Less Developed Themes

Machine learning, deep learning, and sentiment analysis appear in the lower-left section of the thematic map. Their position suggests that these topics have a weaker connection to the central airline scheduling research structure in the dataset, compared with established optimisation-based approaches.

This does not mean that machine learning is irrelevant to airline scheduling. Instead, the data suggests that traditional optimisation techniques remain more strongly connected to the core research structure. Machine learning appears as a developing area that could become more integrated with scheduling problems as researchers combine predictive models with optimisation methods.

Genetic algorithms, multi-objective optimisation, and air transport form another cluster. These methods are relevant to scheduling because airline decisions often involve several objectives at the same time, such as reducing operating costs, improving aircraft utilisation, minimising delays, and maintaining passenger service levels.

COVID-19, air transportation, combinatorial optimisation, and slot allocation appear closer to the emerging or declining area of the map. COVID-19 is particularly interesting because the pandemic created major disruptions in aviation, but the thematic position suggests that pandemic-related research does not represent the central long-term structure of airline scheduling research in the dataset.

Keyword and Author Relationships

The three-field analysis (Figure 4) shows a strong connection between countries, authors, and keywords. The largest country contribution comes from China, followed by the United States and several European and Asian countries.

The keyword side of the analysis is dominated by terms such as air transportation, optimisation, air traffic control, integer programming, aircraft, airports, airline industry, scheduling, costs, decision making, and multi-objective optimisation.

This shows that airline scheduling research is largely built around mathematical and operational decision-making techniques. Optimisation appears repeatedly across the thematic map and keyword network, making it one of the central concepts in the research area.

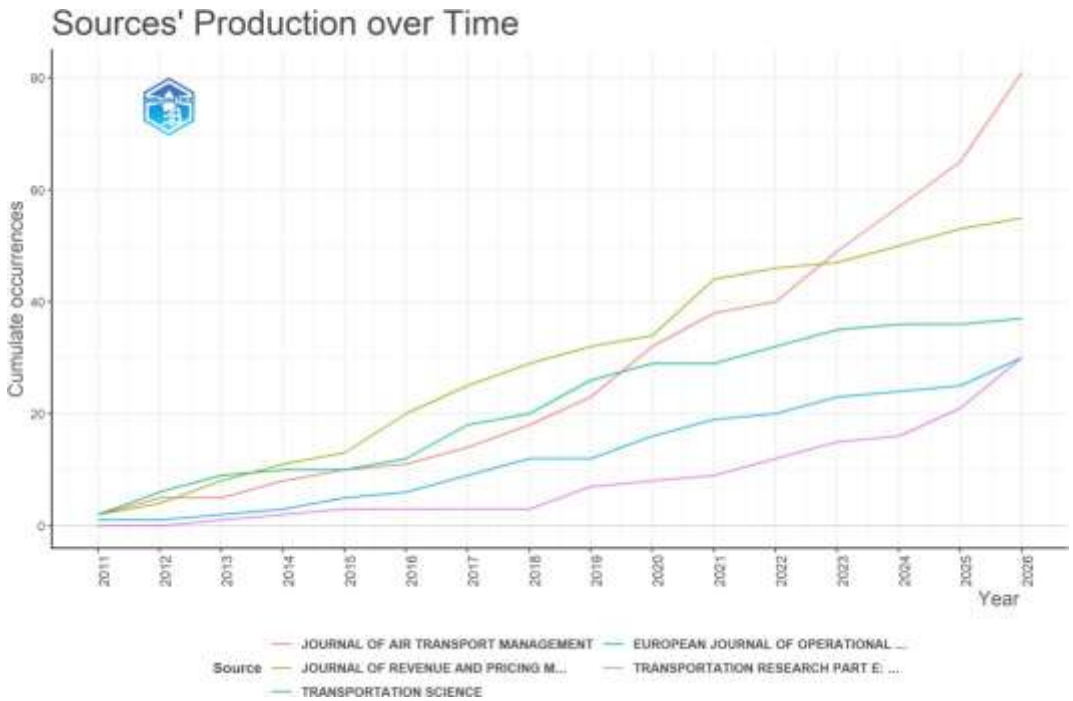
The presence of integer programming, linear programming, combinatorial optimisation, genetic algorithms, and multi-objective optimisation also shows that researchers use a range of mathematical approaches to solve scheduling problems. These methods are applied to different operational areas, including aircraft routing, fleet assignment, airport operations, air traffic control, and revenue management.

Source Production

The source production graph (Figure 6) shows that publication activity has increased across the major journals included in the dataset. The Journal of Air Transport Management shows particularly strong growth toward the later years, reaching approximately 80 cumulative occurrences by 2026.

Figure 6

Sources' Production Over Time (Biblioshiny Output)



The Journal of Revenue and Pricing Management also shows consistent growth, reaching approximately 55 cumulative occurrences. Transportation Science reaches around 37, while the European Journal of Operational Research and Transportation Research Part E reach approximately 30 each.

The growth across these sources indicates that airline scheduling research is spread across both aviation-specific journals

and broader operations research and transportation journals. This supports the thematic findings, where airline scheduling overlaps with optimisation, revenue management, operations research, and transportation planning.

Overall Data Interpretation

The data presents airline scheduling as a research field that has grown substantially over time. Publication activity has increased by more than four times compared with the 2011 level, while the number of countries contributing to the field has also produced a broad international research network.

The central research focus remains strongly connected to optimisation. Aircraft routing, fleet assignment, scheduling, integer programming, revenue management, and dynamic pricing form major parts of the research structure. At the same time, machine learning, deep learning, genetic algorithms, and multi-objective optimisation appear as areas that are becoming increasingly connected with airline and transportation problems.

The country and author analysis shows a strong contribution from China and the United States, while the keyword analysis demonstrates extensive international research around optimisation and airline operations. Taken together, the data suggests that airline scheduling has developed from a mainly mathematical optimisation problem into a broader research area involving aircraft operations, air traffic management, pricing, decision-making, and emerging computational techniques.

5. Future Research Agenda

The bibliometric analysis of research on airline scheduling and optimisation indicates that the field has developed a substantial mathematical and computational foundation, while several questions at the boundaries of established research areas remain less integrated. The dataset analysed in this study contains 1,045 Scopus articles published between 2011 and 2026, drawn from 355 sources and involving 2,371 authors. The reported annual growth rate is 10.29%, while international co-authorship accounts for 27.47%. These indicators suggest a growing field, but publication volume alone does not establish whether research findings are comparable, operationally transferable or useful to decision-makers. A future research agenda should therefore focus not only on proposing new algorithms but also on evaluating existing approaches under consistent conditions and connecting optimisation models with the operational and human consequences of airline decisions.

The thematic map, three-field plot and VOSviewer review provide the starting point for this agenda. Integer programming, fleet assignment and aircraft routing appear as established and well-connected areas. By comparison, machine learning, passenger behaviour, coordination among airlines and airports, and some forms of disruption resilience appear less connected to the central themes. The map also places slot allocation, combinatorial optimisation and COVID-19-related research in less central positions. Such positions should be interpreted carefully: keyword co-occurrence indicates how topics are connected in the selected literature, not the practical importance of a problem or the quality of an individual study. The gaps discussed below are therefore research propositions suggested by the bibliometric evidence, rather than definitive proof that no relevant research exists.

The agenda is organised around two complementary frameworks. The 5Ws1H framework specifies what should be studied, why it matters, who is involved, where and when the research should apply, and how it should be investigated. The Input-Process-Output framework identifies the data required, the analytical process to be applied and the outcomes against which a study should be evaluated. Together, these frameworks translate broad bibliometric patterns into researchable questions. The central proposition is that future work should connect methods to operational problems, test models against credible baselines, report outcomes beyond cost alone and make evidence sufficiently transparent for other researchers to reproduce.

5.1 Basis of the Agenda and Interpretation of Keyword Patterns

The publication trend provides evidence of sustained scholarly attention. The dataset reports an increase from 26 articles in 2011 to a peak of 122 in 2025. China contributes 639 country records and the United States 344, indicating that a large share of the mapped research is associated with these two countries. Five countries account for 59.9% of country records. These figures point to both a mature international research base and a potential geographical concentration. Future work should examine whether methods developed using a limited range of airline networks, regulatory systems and airport structures perform equally well in other operating environments.

The keyword evidence suggests that the field is frequently described through methods and broad domain labels. Terms

such as air transportation, airline industry, airports and aircraft identify the sector but do not specify the decision being optimised. Method-related terms include optimisation, integer programming, linear programming, combinatorial optimisation, column generation, genetic algorithms and multi-objective optimisation.

Problem-related terms include scheduling, fleet assignment, aircraft routing, slot allocation, maintenance, revenue management and dynamic pricing. Air traffic control, air traffic management and airline operations connect airline decisions with the wider transport system. However, the outcome vocabulary visible in the three-field plot is comparatively narrow: costs and decision-making are apparent, while passenger experience, fairness and emissions do not appear among the plotted terms.

This imbalance has practical consequences for how research questions are formulated.

If a paper is labelled primarily by its algorithm, it may be difficult to identify whether it addresses crew recovery, fleet reassignment, passenger disruption or airport coordination. Future bibliometric studies should therefore classify publications using at least three dimensions: the operational problem, the method and the outcome. This would make it easier to distinguish a new algorithm from a new application, and to determine whether the field is making progress on outcomes that matter to airlines, airports, employees and passengers. It would also allow researchers to compare whether certain methods are consistently used for particular problem types and whether those associations change over time.

The keyword results require an additional qualification. The groupings reported in this study are based on node size and labels in the available maps rather than on raw author-keyword frequencies. The dataset contains 3,122 distinct author keywords, and spelling variants such as “optimization” and “optimisation” require normalisation. Exact frequencies should therefore be extracted from Biblioshiny and reported in future work so that the bibliometric claims can be verified. A time-sliced comparison covering 2011–2015, 2016–2020 and 2021–2026 could help determine whether learning-related terms are gaining prominence, whether traditional optimisation themes remain stable, and whether topics such as resilience and sustainability are becoming more connected to the central research structure.

The search strategy is another boundary on interpretation. The search string used in this study was built around “airline” and “optimization”. A search strategy that depends on these terms may fail to retrieve relevant work indexed under recovery, disruption management, crew scheduling, passenger flow, airport capacity or resilience without using the word optimisation. Future reviews should test a broader search vocabulary, include both British and American spellings, and document the effects of each search-term expansion. The purpose is not to increase the dataset indiscriminately, but to assess whether the current map adequately represents the wider research landscape. These steps would strengthen the evidence on which the following agenda is based.

5.2 Research Gaps and Future Directions

5.2.1 Integrating Machine Learning with Mathematical Optimisation

Machine learning and deep learning appear weakly connected to the central scheduling themes in the thematic map, even though the reviewed papers show that integration is beginning to occur. Hızır, Barnhart and Vaze (2026) combine mixed-integer optimisation with supervised learning for crew recovery. Xiao et al. (2026) examine features relevant to machine-learning-enhanced aircraft recovery, while Jia, Guo and Mei (2026) combine a memetic algorithm with reinforcement learning. These examples suggest that learning methods can contribute forecasts, policies or search guidance, while mathematical optimisation can preserve explicit operational constraints. However, the existence of individual hybrid approaches does not establish that they consistently outperform established methods.

A priority question is: how much does a learned prediction of delay or cancellation improve the final scheduling decision compared with the same optimisation model without that prediction? A fair test should use identical instances, the same computational time limit and a strong exact or established heuristic baseline. It should report not only prediction accuracy but also the downstream effect on recovery cost, delay minutes, feasibility and the number of disrupted flights restored. Forecasting performance and decision performance are related but not interchangeable. A model can predict delays accurately without changing the decision, while a modest forecast improvement may be valuable if it changes a critical assignment. Studies should make this distinction explicit.

Reinforcement learning raises additional questions about generalisation and interpretability. Policies learned from historical or simulated traffic may perform poorly when demand, weather, airport capacity or disruption patterns change.

Future studies should evaluate policies on out-of-sample periods and under deliberately difficult scenarios. They should also explain which operational features influence the chosen action and identify conditions under which a human planner should override it. Qiao et al. (2026) and Wang, Wang and Hansen (2026), indicate applications of reinforcement learning to congestion pricing and flow management. The next step is to evaluate robustness, constraint compliance and decision quality under comparable conditions rather than treating the use of learning itself as evidence of improvement.

5.2.2 Measuring Passenger Consequences and Incorporating Passenger Information

Passenger sentiment analysis appears in the mapped research but is not visibly connected to scheduling decisions in the reviewed set. Reviews, complaint records and public posts can reveal recurring disruption concerns, but these sources are noisy and may over- represent passengers who choose to comment. Future studies should test whether text-based evidence can complement operational data rather than assuming that sentiment scores directly measure passenger welfare. A useful design would link disruption-related text to booking, cancellation, rebooking and missed-connection records where access and privacy rules permit. The analysis should distinguish different passenger groups and circumstances, because the consequences of a delay may vary according to connection status, journey purpose or available alternatives.

The central research question is whether passenger costs used in scheduling and recovery models can be estimated more realistically from observed behaviour. Many models require a penalty for delay or cancellation; if that penalty is assumed rather than empirically supported, the resulting schedule may optimise a proxy that does not reflect actual passenger consequences. Passenger-choice models, including the work by Papi-RAD, Pishvaee and Sahebi (2026), offer a possible connection. Further research could compare assumed passenger costs with estimates derived from observed rebooking, cancellation or itinerary-choice behaviour. Outcomes should include missed connections, rebooking uptake, passenger delay and the distribution of disruption across passenger groups, alongside the usual operating-cost measures.

Passenger information should be used with clear safeguards. Text data can contain personal information, inconsistent descriptions and strong selection effects. Studies should state how data are anonymised, how sentiment categories are validated and how missing or biased observations are handled. Where direct passenger-level data are unavailable, researchers should use clearly labelled proxies and test whether conclusions change under alternative assumptions. The goal is not to replace operational measures with sentiment analysis, but to assess whether passenger evidence improves decisions and makes trade-offs more transparent.

5.2.3 Designing Schedules That Remain Resilient Under Major Shocks

The small COVID-19 keyword cluster suggests that pandemic-related research is not strongly integrated into the broader thematic structure captured by the dataset. This does not mean that disruption resilience has received no attention. Rather, the map raises the question of whether lessons about sudden demand and capacity shocks have been incorporated into routine scheduling models. COVID-19 demonstrated the possibility of abrupt network contraction followed by uneven recovery. Future research should investigate which schedule characteristics help airlines maintain essential connectivity, reallocate aircraft and crews, and restore services when demand or operational capacity changes sharply.

Robust and stochastic optimisation provide relevant approaches, but they should be assessed against clearly specified disruption scenarios. Researchers could model shocks affecting demand, airport capacity, crew availability or fleet utilisation, then compare schedules designed for expected conditions with schedules that include spare capacity or flexible resource pools. The evaluation should report recovery time, revenue retained, cancellation and delay levels, and the costs of idle aircraft and crews. A model that performs well under one assumed scenario may be fragile under another, so scenario selection and sensitivity analysis should be reported transparently. The objective is not to eliminate all unused capacity, since reserve resources may have option value when disruptions occur, but to estimate the cost and benefit of resilience under plausible conditions.

A further direction is to examine whether resilience investments are worthwhile across different types of airline and network. A hub-and-spoke carrier, a low-cost carrier and an airline operating a thin regional network may have different opportunities to substitute aircraft, crews and routes. Research should therefore compare multiple network structures rather than drawing general conclusions from one case. Historical disruptions can support retrospective testing, while simulation can explore rare but severe scenarios. The limitations

of simulated data should be acknowledged, and the assumptions should be calibrated wherever observed evidence is available.

5.2.4 Strengthening Slot Allocation and Coordination Across Organisations

Slot allocation appears in a less central area of the thematic map, while the reviewed recent papers indicate continued work on the problem. Melder et al. (2026) formulate network-level slot allocation; Bao, Chu and Li (2026) incorporate fairness between airports and airlines; and Wang (2026a) links slot allocation with boarding gates. These directions suggest that the problem cannot always be treated as an isolated allocation exercise. A slot decision affects aircraft rotations, gate use, passenger connections and the feasible schedules of multiple airlines. Research should determine when a network-level model produces better outcomes than sequential decisions made by separate organisations.

Coordination is difficult because airlines, airports and air traffic management bodies do not necessarily share the same objectives or information. Airlines may prioritise network connectivity and operating cost, airports may seek efficient use of capacity, and traffic managers may need to maintain safe and manageable flows. A joint optimisation model may appear superior if it assumes complete information and cooperation, yet those assumptions may be unrealistic. Future studies should state what information each participant possesses, what information can be shared, and what incentives would encourage participation. Game-theoretic models, negotiated allocation mechanisms and incentive-compatible rules could be compared with centralised optimisation. Outcomes should include total system delay and cost, allocation efficiency, fairness across participants and the computational burden of coordination.

Fairness also requires a precise definition. It may refer to equal treatment of airlines, proportional treatment based on historical use, protection of smaller operators or minimisation of passenger harm. These definitions can produce different allocations.

Researchers should report the fairness measure used and show how it changes the solution relative to an efficiency-only model. Multi-airport systems and congested hubs are useful settings because local decisions can have network-wide effects. Evidence from different regulatory environments would help establish whether proposed coordination mechanisms are transferable.

5.2.5 Connecting Revenue Management with Operational Recovery

Revenue management and dynamic pricing appear among the established themes, yet the reviewed literature suggests that pricing is not consistently linked with disruption recovery. This separation may limit the practical value of models that optimise seat inventory or fares under normal operations but do not account for the consequences of cancellations, aircraft substitutions and schedule changes. Future research should investigate how fares, seat availability and rebooking policies should respond while a disrupted schedule is being repaired. The question is not simply whether prices should rise or fall, but how revenue decisions interact with passenger reaccommodation, network connectivity and the availability of seats on alternative flights.

A suitable model would connect schedule recovery decisions with seat inventory, passenger demand and rebooking constraints. Researchers could compare sequential approaches, in which revenue management and recovery are optimised separately, with integrated approaches that account for their mutual effects. Evaluation should include revenue retained, recovery cost, passenger disruption, load factor and the number of passengers successfully rebooked. It should also test different disruption severities and demand conditions. Since pricing decisions can create distributional consequences, studies should distinguish revenue gains from improvements in overall system performance and discuss any passenger groups that bear a disproportionate burden.

5.2.6 Comparing Metaheuristics with Exact and Hybrid Methods

Genetic, swarm and memetic algorithms are useful when airline scheduling problems become too large or complex for direct exact solution within operational time limits. The thematic map places genetic algorithms and multi-objective optimisation near the centre but below the density line, suggesting that these methods are present without being fully integrated into a mature comparative theme. Studies by Devi (2026), Ding (2026) and Jia, Guo and Mei (2026) are relevant examples. The main gap is not necessarily the absence of algorithms; it is the shortage of controlled evidence about the conditions under which one method is preferable to another.

Future comparisons should use common instances, consistent hardware and equal time budgets. Exact methods should be used where feasible to establish optimality or a meaningful bound; for larger instances, researchers should report the

best-known solution, optimality gap where available, runtime, memory or computational requirements, and repeatability across random seeds. Algorithm parameters and stopping criteria should be documented. Without such controls, apparent improvements may reflect different test settings rather than a better method. Researchers should also report feasibility rates and the quality of the final operational schedule, not only the objective value returned by the algorithm.

Multi-objective problems require more than a single weighted score. Airline schedules can trade off operating cost, delay, emissions, fairness and passenger inconvenience. Future work should report Pareto-front quality and explain the trade-offs among objectives. Where a weighted sum is used, the rationale for the weights should be stated and sensitivity to alternative weights should be tested. Comparisons between exact methods, metaheuristics and hybrid approaches should determine not only which produces a better solution, but how solution quality and computational effort change as the network grows and as additional objectives are included.

5.2.7 Treating Sustainability as an Integrated Scheduling Objective

Sustainability does not appear as a distinct cluster in the mapped evidence, although individual studies address related topics. Joormann, Kroner and Tillmann (2026) address electric-flight airport selection, Demouge, Guitart and Delahaye (2026) contrail-aware trajectories, and Gurturk and Celik (2026) sustainable cargo routing. These examples suggest that environmental considerations are being studied, but often through particular applications. A broader research direction is to integrate emissions and other environmental measures with fleet assignment, routing, slot allocation and disruption recovery rather than treating sustainability as a separate planning exercise.

The practical challenge is that environmental and operational objectives may conflict. A route or schedule that reduces emissions may increase flight time, cost or disruption risk; an alternative aircraft assignment may reduce fuel use but create crew or maintenance constraints. Researchers should therefore specify the emissions metric, its system boundary and the assumptions used to estimate it. They should compare cost-only solutions with multi-objective solutions and report the additional cost associated with an environmental improvement. Where possible, studies should distinguish direct operational emissions from other environmental indicators and avoid combining incomparable measures into a single score without justification.

Sustainability objectives should also be tested under disruption. A schedule that appears efficient under normal conditions may require extensive recovery actions that increase emissions when delays cascade through the network. Conversely, some reserve capacity or flexibility may reduce the need for inefficient last-minute substitutions.

Integrating environmental measures with resilience could reveal trade-offs that are not visible when the two objectives are analysed separately. The aim should be to provide decision- makers with transparent choices, not to assume in advance that one objective must always dominate.

5.2.8 Assessing the Realistic Role of Quantum and Quantum-Inspired Methods

Quantum computing and quantum-inspired optimisation do not form a prominent cluster in the mapped research, although Du (2026) and Gili, San Sebastian and Blázquez-García (2026) are examples of emerging work. At present, the appropriate research question is narrow and comparative: for which problem structures and instance sizes, if any, do quantum or quantum-inspired methods provide an advantage over well-tuned classical approaches? Small demonstrations alone cannot establish practical superiority for airline scheduling, where constraints, data preparation and solution validation can dominate computational performance.

Future studies should begin with small, controlled benchmark problems that can be solved or bounded using classical methods. They should disclose the problem encoding, preprocessing, hardware or simulator conditions, number of runs and total computational effort. The comparison should include the time required to translate the airline problem into the solver's representation and to convert the output into a feasible schedule. Researchers should avoid claims of advantage based solely on the performance of a solver component if the full workflow is slower or produces less reliable solutions. Quantum approaches should be treated as an exploratory direction until repeatable comparisons demonstrate an advantage on meaningful instances.

5.2.9 Testing Transferability Across Countries, Carriers and Airport Systems

The geographic concentration in the dataset creates a question about external validity.

The dataset contains 639 country records for China and 344 for the United States, while India, Turkey and Iran appear among countries with smaller contributions. Research concentrated in a small number of settings may reflect their

network structures, data availability, regulatory arrangements or research priorities. Models calibrated for one environment may not perform equally well in another. Future work should explicitly test transferability rather than assume that a model validated on one network can be applied without adjustment elsewhere.

A useful design would compare hub-and-spoke and point-to-point networks, full-service and low-cost carriers, and airports with different capacity constraints. Emerging-market settings may reveal different combinations of congestion, infrastructure limitations, data quality and coordination challenges. Researchers should identify which model components are general and which depend on local conditions. They should also describe the data required for replication and explain any differences in regulatory or operational assumptions. International collaboration with airlines, airports and research groups outside the leading contributors could broaden the evidence base, provided that the resulting studies maintain common definitions and comparable evaluation procedures.

5.2.10 Improving Bibliometric Measurement and Reproducibility

A final research direction concerns the evidence used to describe the field itself. The current thematic interpretation relies on a Scopus dataset and keyword co-occurrence. Cited-reference information was not available in the Biblioshiny output used in this study, which limited the ability to conduct co-citation or reference-based analysis. This should be treated as a methodological limitation, not as a finding about airline scheduling.

Future bibliometric studies should preserve a reproducible record of the search string, database, retrieval date, filters, duplicate removal and exclusion decisions. They should normalise spelling variants and related keyword forms, publish frequency tables, and distinguish author keywords from database-assigned terms. Combining thematic evolution with citation, co-citation and bibliographic coupling analyses could show whether a topic is conceptually central, intellectually influential or newly emerging. A broader database search could also test whether the apparent absence of a topic reflects a genuine gap or simply the boundaries of the original search strategy.

5.3 Prioritising the Research Programme

The ten directions differ in the resources and data they require. A practical first phase would prioritise work that can be tested using public or simulated instances, especially comparisons between learning-enhanced models, exact methods and metaheuristics. These studies can establish common evaluation standards and identify where improvements are repeatable. A second phase could focus on passenger behaviour, integrated revenue management and resilience, because these questions require richer operational, passenger or scenario data. A third phase could develop coordination mechanisms across organisations and conduct controlled tests of quantum or quantum-inspired methods, where access, governance or technical limitations may make validation more difficult. Sustainability and geographic transferability can be incorporated throughout rather than postponed to a final phase.

This sequence is a suggested order of work, not a claim that one topic is inherently more important than another. Researchers should select priorities based on data access, operational relevance and the decision the study intends to improve. Each project should define its baseline before selecting a new method, and it should specify what evidence would count as a meaningful improvement. A study should be considered stronger when other researchers can access the benchmark instances, reproduce the main comparison and understand the conditions under which the proposed method fails as well as succeeds.

A minimum reporting standard would include the problem definition, data source, assumptions, constraints, benchmark instances, baseline methods, computational environment, stopping rules and evaluation metrics. For studies involving learning, the report should distinguish predictive error from decision impact and should include out-of-sample testing. For multi-objective studies, it should state the objective definitions and show trade- offs rather than relying only on one aggregate score. For collaborative models, it should specify the information and incentives required from each stakeholder. These standards would make results easier to compare and reduce the risk that the literature accumulates algorithms without accumulating reliable evidence.

5.4 The 5Ws1H Framework

The 5Ws1H framework turns the identified gaps into a coherent programme. What should be studied is integrated, uncertainty-aware airline scheduling that links relevant decisions across fleet assignment, aircraft routing, crew allocation, slot use and operational recovery. The focus should be on connected decision chains where one choice

changes the feasibility or quality of another, rather than assuming that each decision can always be optimised independently. The framework also includes passenger outcomes, fairness and environmental impact where these are relevant to the operational question.

Why this work is needed follows from the pattern of the bibliometric evidence.

Established mathematical methods are well represented, while links between optimisation, learning, passenger information, resilience and coordination appear less developed in the map. Cost is more visible than passenger, fairness or environmental outcomes. These patterns support a research agenda centred on integration and evaluation. They do not, by themselves, prove that integrated models will always be better; the purpose of future work is to test when the additional complexity improves decisions enough to justify its computational and organisational costs.

Who is involved includes airlines, airports and air traffic management bodies as decision-makers, with crews, ground staff and passengers as affected groups. Researchers should identify which actor controls each decision and who bears the costs when a schedule changes. Where several organisations participate, the model should reflect the limits of information sharing and the possibility of conflicting objectives. Research partnerships across countries and organisation types would also help test whether findings are robust beyond the most frequently represented settings in the dataset.

Where the research should be conducted includes major hubs, multi-airport systems and networks with differing operating models. Emerging-market airports and carriers are relevant because transferability cannot be established from a single national or regulatory context. The choice of setting should be justified by the mechanism being studied: slot coordination requires meaningful capacity interactions, while passenger recovery research requires reliable observations of passenger decisions. When research is undertaken, it should span three operational horizons: strategic schedule design, tactical fleet and crew assignment, and day-of-operations recovery. Longer-term bibliometric updates can track whether the themes become more central as evidence accumulates.

How the agenda should be pursued depends on the problem. Mixed-integer programming and decomposition provide a basis for representing explicit operational constraints. Metaheuristics may be useful for larger instances, while supervised learning or reinforcement learning can be tested as prediction or decision-support components. Hybrid learn-then-optimise designs should be compared with methods that do not use learned inputs. Robust and stochastic approaches are appropriate when uncertainty is central to the question. Quantum or quantum-inspired methods should initially be tested on small, controlled cases. Across all approaches, shared benchmark instances, rolling-horizon simulation, out-of-sample testing and transparent reporting should be treated as core evaluation requirements.

5.5 The Input–Process–Output Framework

The Input-Process-Output framework complements 5Ws1H by requiring every research question to specify the evidence entering the model, the analytical procedure and the outcomes used to judge success. For learning-enhanced optimisation, inputs may include historical delay and cancellation records, aircraft and crew status, and estimates of disruption probability. The process may use supervised learning or reinforcement learning to generate forecasts or candidate actions that feed a mixed-integer or constraint-based model. Outputs should include recovery cost, delay minutes, feasibility, solution time, prediction error and the effect of that error on the final decision. A predictive metric alone is insufficient if the research claim concerns better scheduling.

For passenger-centred models, inputs may include disruption-related reviews, complaint logs, social posts, booking records and rebooking information, subject to appropriate privacy and access safeguards. Sentiment or topic analysis can be used to develop passenger-impact indicators, which may then enter a passenger-choice or recovery model.

Outputs should measure passenger delay, missed connections, cancellation or rebooking behaviour and differences between passenger groups. The study should test whether these measures change the schedule selected by the model and whether the resulting change improves passenger outcomes without creating unacceptable operational costs.

For resilience research, inputs should include demand and capacity shocks, fleet and crew limits, and alternative disruption scenarios. Robust or stochastic optimisation can then be used to compare schedules under expected conditions and under stress. Outputs should include the time needed to restore network operation, cancellations, revenue retained, delays and the cost of idle aircraft or crews. For slot allocation and coordination, inputs should include airport capacity, airline requests, fairness rules and the information each participant is willing or able to share. The process may compare centralised optimisation with coordinated, negotiated or incentive-based allocation. Outputs should cover allocation efficiency, system delay, participant-level costs, fairness and the burden of coordination.

For multi-objective optimisation, inputs should include clearly defined cost, delay, emissions, fairness and passenger-impact measures, as relevant to the problem. The process should compare exact methods, metaheuristics and hybrid approaches under equivalent conditions and evaluate the quality of the Pareto front. Outputs should include solution quality, computational effort, scalability and the trade-offs among objectives. For bibliometric research, inputs should include reproducible database exports, documented search strategies, cited references and normalised keyword lists. The process should combine thematic maps with appropriate citation and network analyses. Outputs should include verified keyword frequencies, time-based thematic changes and transparent limitations.

Across all these areas, the IPO principle is the same: a study must report measurable results that correspond directly to its research question.

5.6 Implications for Researchers and Airline Planners

For researchers, the main implication is to define the operational decision and the outcome before selecting a method. A study of crew recovery, for example, should report more than the objective value of a mathematical model; it should also report feasibility, delay, computational time and, where relevant, passenger or staff consequences. Shared instances and clear baselines would allow subsequent work to determine whether an apparent improvement is robust. Where decisions are integrated, authors should specify which decisions are coupled and how that integration affects computation, data needs and implementation.

For airline planners, the agenda suggests that method selection should be conditional on the problem, the level of uncertainty, the available data and the time available for a decision. No single algorithm should be assumed to be best across fleet assignment, slot allocation, revenue management and disruption recovery. Learning-based tools should be assessed using operationally relevant tests and should be monitored when conditions differ from their training data. Planners should also examine whether a proposed model requires information or cooperation that the organisation cannot realistically obtain. A theoretically strong allocation model may be unusable if its success depends on information that other participants will not share.

For airports and air traffic management bodies, the key implication is that local efficiency and system-wide efficiency may differ. A decision that improves one airline's schedule can impose costs on another airline or on airport operations. Research should therefore make stakeholder objectives explicit and report both aggregate and participant-level outcomes. Fairness should not be added as a vague label; it should be operationalised through a stated measure and tested against efficiency. Environmental and passenger outcomes should be included when they are material to the decision, with transparent explanations of the associated trade-offs.

For the research community, a wider geographical and institutional base would improve the evidence on transferability. International collaboration should be used to compare methods under common definitions, not merely to increase the number of author affiliations. Bibliometric updates can identify shifts in topic attention, while operational studies are needed to determine whether those topics produce usable improvements. The most valuable next contribution may therefore be a rigorous comparative study or a reproducible benchmark, rather than another algorithm tested on a unique instance that cannot be independently assessed.

5.7 Summary and Limitations of the Agenda

The future of airline scheduling and optimisation research depends on strengthening the connections between a mature mathematical core and the operational contexts in which schedules are designed, disrupted and recovered. The bibliometric evidence points to established themes in integer programming, fleet assignment and aircraft routing, alongside less integrated work on learning methods, passenger behaviour, resilience, coordination, sustainability and emerging computational approaches. These patterns suggest that the field needs stronger comparative evidence, clearer outcome measures and broader validation across different airline and airport settings.

The immediate priority is to compare methods on shared instances and under equivalent computational conditions. Learning-enhanced models should be judged by their effect on decisions as well as by predictive accuracy. Passenger information should be tested as a source of measured disruption costs rather than treated as an automatic substitute for operational evidence. Resilience models should evaluate recovery under multiple demand and capacity shocks. Coordination models should state what information and incentives are required, while multi-objective models should report trade-offs among cost, delay, fairness, passenger outcomes and emissions. Emerging methods such as quantum optimisation should be assessed through controlled benchmarks against strong classical alternatives.

The agenda has limitations. It is based on the bibliometric dataset and thematic interpretation reported in this study, which uses Scopus records and keyword co-occurrence. The search string may omit relevant studies that use different terminology, and the keyword groupings require verification against exact frequency data. In addition, comments about individual studies reflect the reviewed reference set and should not be interpreted as a complete assessment of every publication in the field. The research gaps are consequently hypotheses for further investigation, not proof that no prior work exists.

Overall, progress should be measured by the quality and transferability of evidence, not simply by the number of proposed techniques. Studies that publish their assumptions, instances, evaluation procedures and outcome measures will allow the community to determine which methods work, for which problems, and under what conditions. Such work can help move airline scheduling research from a collection of specialised methods toward a more integrated evidence base that supports reliable decisions for airlines, airports, employees and passengers.

6. Conclusion and Limitations

This study set out to examine how research on airline scheduling and optimisation has developed, using a bibliometric review supported by a thematic literature review.

Specifically, it aimed to determine how publication output has changed over time, which countries, authors and sources contribute most to the field, which research themes are dominant, niche or emerging, and what gaps and future directions the literature suggests. Following the PRISMA framework (Moher et al., 2010), an initial 2,828 Scopus records were refined by time span, subject area, document type, language and source type to a final dataset of 1,045 journal articles and review papers. These were analysed using RStudio and Biblioshiny and complemented by a VOSviewer cluster analysis of a separate RIS database.

The analysis has identified a clear and sustained expansion of research activity.

Annual publications rose from 26 articles in 2011 to 113 in 2026, an increase of about 335%, with the highest value of 122 articles recorded in 2025. Average annual output increased from 35.6 articles in 2011–2015 to 58 in 2016–2020 and 95.5 in 2021–2026. This growth was not entirely steady, since declines occurred in several years, including 2015, 2018, 2022, 2024 and 2026. Nevertheless, the overall pattern shows that airline scheduling has become an established and expanding area of scholarly enquiry.

The research has also shown that production is geographically concentrated. The country-level data contains 2,147 records across 68 countries, of which China accounts for 639 (about 29.8%) and the United States for 344. Together, these two countries represent about 45.8% of the records, and the top five countries account for approximately 59.9%. The three-field analysis links prominent authors such as Wang Y, Sun X, Chung SH and Zhang Y to keywords including air transportation, optimisation, air traffic control and integer programming. Source production further indicates that the field is spread across aviation-specific and broader operations research outlets. The Journal of Air Transport Management reached approximately 80 cumulative occurrences by 2026, followed by the Journal of Revenue and Pricing Management (about 55), Transportation Science (about 37), and the European Journal of Operational Research and Transportation Research Part E (about 30 each).

The thematic map revealed that the core of the field remains firmly based on optimisation. Revenue management, optimisation and dynamic pricing; column generation, fleet assignment and aircraft routing; and integer programming, scheduling and air traffic management emerged as motor themes. Airlines, robust optimisation and airline operations appeared as a specialised niche. By contrast, machine learning, deep learning and sentiment analysis were positioned in the less developed section of the map, which suggests a weaker connection to the central research structure. Genetic algorithms and multi-objective optimisation formed a further cluster, while COVID-19, combinatorial optimisation and slot allocation appeared closer to the emerging or declining area.

The thematic literature review reinforced these results. The VOSviewer analysis produced three interconnected clusters: air transportation, optimisation and decision-support; computational and mathematical optimisation methods; and airport operations and air navigation management. The close links between them indicate that researchers increasingly combine different optimisation approaches across several areas of aviation decision-making. The reviewed studies show that mathematical programming remains dominant, as in crew recovery using mixed-integer optimisation (Hızır et al., 2026) and network-level slot allocation (Melder et al., 2026), but is increasingly supplemented by metaheuristics, machine learning, reinforcement learning and, more recently, quantum computing (Du, 2026; Gili et al., 2026).

Taken together, these findings suggest that airline scheduling has evolved from a mainly mathematical optimisation problem into a broader research area involving aircraft operations, air traffic management, pricing and decision-making. Three implications follow. First, decisions that were once treated separately, such as fleet assignment, routing, crew scheduling and recovery, are now increasingly optimised jointly. Second, uncertainty and disruption have become central concerns, as shown by work on integrated schedule recovery (Abdelghany & Abdelghany, 2026) and recovery under uncertain disruptions (Wu et al., 2026). Third, objectives are widening beyond cost to include fairness, fatigue, sustainability and passenger behaviour (Papi-RAD et al., 2026). The evidence also suggests that no universal algorithm exists for optimising air transportation. Instead, the literature emphasises context-specific methods that reflect the scope of the problem, the sources of uncertainty, the objectives and the available resources.

The study contributes to our understanding of the field in several ways. First, it provides a systematic overview of publication trends, international research networks, influential authors and sources, which has previously been difficult to obtain from reviews that concentrate on a single sub-area, such as aircrew scheduling (Sun, 2026) or airport ground operations (Dahanayaka, 2026). Second, it locates established, niche and less developed themes within a single structure, showing where optimisation remains dominant and where machine learning is still emerging. Third, it identifies challenges and gaps that can help researchers position future work. The findings should be of interest to academics studying airline and airport operations, and to practitioners concerned with how optimisation methods are being applied in planning and decision-support.

Several limitations should be acknowledged. The study is limited to Scopus-indexed, English-language journal articles and review papers, and it excludes conference papers and other document types. The search used the keywords “airline” AND “optimization”, so relevant studies that use different terminology may not have been captured. The bibliometric approach also describes patterns in the literature rather than evaluating the quality or performance of individual algorithms, and the interpretation of the thematic map depends on how themes are positioned and labelled. In addition, the thematic literature review draws on a separate RIS database, and a full comparison of optimisation methods lies beyond the scope of this study. The findings should therefore be read as a map of the literature as indexed in Scopus rather than a complete record of all work on airline scheduling.

Notwithstanding these limitations, the study offers valuable insights into the structure and direction of airline scheduling research. The consistency between the bibliometric results and the thematic review, in particular the central role of optimisation and the growing role of data-driven methods, strengthens confidence in the main conclusions.

Several questions still remain to be answered, and further research could usefully explore them. First, the most significant gap concerns coordination among airlines, airports, air traffic management organisations and other stakeholders. These actors operate within the same system but often hold conflicting goals and different levels of information. Future work should develop system-level models that coordinate decision-making across stakeholders while accounting for uncertainty, sustainability and computational complexity. Second, machine learning and reinforcement learning raise issues of data requirements, interpretability and validation, so further studies are needed to establish how predictive models can be integrated more effectively with optimisation methods. Third, benchmark testing under realistic operating conditions is required to evaluate the performance of different algorithms. Finally, future bibliometric studies could extend the dataset to additional databases such as Web of Science, include conference papers, and use broader search terms to test whether the patterns reported here hold more widely.

The findings also have a number of practical implications. Airlines and airports considering optimisation tools should select methods according to the specific problem, its uncertainty and its objectives, rather than expecting a single approach to perform well in all settings. Investment in integrated decision-support systems that link scheduling, recovery and forecasting appears worthwhile, given the close connections identified between these areas.

Greater efforts are also needed to make sure that data-driven methods are validated before they are used in operational decisions.

In conclusion, this bibliometric review shows that airline scheduling research is growing, internationally distributed and firmly rooted in optimisation, while becoming more integrated, adaptive and multi-objective. The research frontier is defined by the design of integrated and adaptive decision-support systems for managing large-scale and complex air transportation networks. Progress will depend on closer coordination between stakeholders, more rigorous testing under realistic conditions, and the careful integration of machine learning with established mathematical methods.

Declarations

Conflict of interest: The authors declare no conflict of interest.

Funding: This research received no external funding.

Use of generative AI: Claude (Anthropic) was used to assist with merging the section drafts, formatting and reference formatting. The authors reviewed and edited the content and take full responsibility for the publication.

References

1. Abdelghany, A. and Abdelghany, K. (2026) 'Addressing the systemic complexity of airline irregular operations: Toward integrated schedule recovery', *Journal of Air Transport Management*. Available at: https://doi.org/10.1016/j.jairtraman.2025.102925.

2. Assink, J. (2026) 'Integrated shift and break scheduling for airline ground staff under flight delay uncertainty', *European Journal of Operational Research*. Available at: https://doi.org/10.1016/j.ejor.2026.07.008.

3. Bao, D., Chu, J. and Li, Y. (2026) 'A slot allocation optimization method for regional multi-airport integrating airport fairness and airline fairness', *Transportation Research Part E: Logistics and Transportation Review*. Available at: https://doi.org/10.1016/j.tre.2026.104886.

4. Borndörfer, R., Do, V.H. and Hoang, N.D. (2026) 'An efficient algorithm for vertical free-flight optimization in air transportation', *European Journal of Operational Research*. Available at: https://doi.org/10.1016/j.ejor.2026.02.003.

5. Chen, L. (2026) 'Optimising aircraft routing and fleet assignment with flight unit loops: A novel decision unit approach for reducing model complexity', *Journal of the Operational Research Society*. Available at: https://doi.org/10.1080/01605682.2025.2612142.

6. Dahanayaka, M. (2026) 'From gate to runway: A systematic review of airport ground operations optimization', *Journal of Air Transport Management*. Available at: https://doi.org/10.1016/j.jairtraman.2026.103013.

7. Demir, H.I. and Dervis, S. (2026) 'Enhancing flight connectivity via synchronization of arrivals and departures in hub airports with evolutionary and swarm-based metaheuristics', *Biomimetics*, 11(1). Available at: https://doi.org/10.3390/biomimetics11010006.

8. Demouge, C., Guitart, A. and Delahaye, D. (2026) 'Balancing fuel efficiency and environmental impact: 4D trajectory optimization through Fast Marching Tree for transatlantic flights considering contrails', *Journal of Air Transport Management*. Available at: https://doi.org/10.1016/j.jairtraman.2026.103032.

9. Devi, V. (2026) 'A hybrid jumping particle swarm optimization with local search for airline crew pairing: Theoretical and empirical insights', *National Academy Science Letters*. Available at: https://doi.org/10.1007/s40009-026-02103-z.

10. Ding, K. (2026) 'Sustainable optimization in air transport: Hybrid particle swarm and Tabu Search algorithm for the multi-objective airport gate assignment problem', *Sustainability*, 18(7). Available at: https://doi.org/10.3390/su18073331.

11. Du, Z. (2026) 'Overcoming computational challenges in air transportation: A quantum computing perspective of the status quo and future applicability', *Transportation Research Part C: Emerging Technologies*. Available at: https://doi.org/10.1016/j.trc.2025.105505.

12. Gili, M., San Sebastian, P. and Blázquez-García, A. (2026) 'Optimization of flight routes: Quantum approximate optimization algorithm for the tail assignment problem', *Quantum Machine Intelligence*. Available at: https://doi.org/10.1007/s42484-026-00387-4.

13. Guardo-Martinez, E. (2026) 'Robust airline scheduling with turnaround under uncertainty: Towards collaborative airline scheduling', *Transportation Research Part E: Logistics and Transportation Review*. Available at:

https://doi.org/10.1016/j.tre.2025.104440.

14. Gurturk, M. and Celik, E. (2026) 'Sustainability takes flight: A hybrid approach to optimizing air cargo routes', *Expert Systems with Applications*. Available at: https://doi.org/10.1016/j.eswa.2025.129366.

15. Hızır, A.E., Barnhart, C. and Vaze, V. (2026) 'Large-scale airline crew recovery using mixed-integer optimization and supervised machine learning', *Transportation Science*. Available at: https://doi.org/10.1287/trsc.2025.0105.

16. Jia, X., Guo, T. and Mei, Y. (2026) 'A self-adapting memetic algorithm with reinforcement learning for multi-objective aircraft recovery problems', *Applied Soft Computing*. Available at: https://doi.org/10.1016/j.asoc.2026.115550.

17. Joormann, I., Kroner, F. and Tillmann, A.M. (2026) 'A two-stage optimization approach for selecting electric-flight airports', *Networks*. Available at: https://doi.org/10.1002/net.70066.

18. Li, R. (2026a) 'Integrated aircraft routing problem under stochastic delays', *Journal of Advanced Transportation*. Available at: https://doi.org/10.1155/atr/7019633.

19. Li, Y., Zhao, K., Zhou, Z., Yang, J. and Zhang, S. (2026b) 'Study of flight supply-demand alignment optimization considering passenger choice behavior', *Journal of Beijing Jiaotong University*. Available at: https://doi.org/10.11860/j.issn.1673-0291.20250189.

20. Melder, D., Wang, S., Mansoor, M.H., Drake, J.H. and Burke, E.K. (2026) 'The network level airport slot allocation problem: A new formulation and matheuristic solution approach', *Transportation Research Part B: Methodological*. Available at: https://doi.org/10.1016/j.trb.2026.103428.

21. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. and The PRISMA Group (2010) 'Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement', *International Journal of Surgery*, 8(5), pp. 336–341. Available at: https://doi.org/10.1016/j.ijsu.2010.02.007.

22. Ong, J.D. (2026) 'Airline crew scheduling using fatigue-score-based optimization model for minimizing pilot fatigue', *Engineering Proceedings*, 128(1). Available at: https://doi.org/10.3390/engproc2026128028.

23. Papi-RAD, A., Pishvaee, M.S. and Sahebi, H. (2026) 'Airline planning under market uncertainty and passengers' choice: Data-driven distributionally robust decomposition- based optimization framework', *Transportation Research Part E: Logistics and Transportation Review*. Available at: https://doi.org/10.1016/j.tre.2026.105158.

24. Paul, J. and Rialp-Criado, A. (2020) 'The art of writing literature review: What do we know and what do we need to know?', *International Business Review*, 29(4), 101717. Available at: https://doi.org/10.1016/j.ibusrev.2020.101717.

25. Qiao, P., Hu, M., Yin, J., Su, J. and He, Y. (2026) 'Multi-agent real-time pricing mechanism for airport congestion management with online distributed reinforcement learning', *Transportation Research Part C: Emerging Technologies*. Available at: https://doi.org/10.1016/j.trc.2026.105628.

26. Shihua, J.M., Nguyen, C.H.C. and Liem, R.P. (2026) 'Real-time bi-level aircraft trajectory optimisation under the presence of unsteady wind', *Optimization and Engineering*. Available at: https://doi.org/10.1007/s11081-025-10005-0.

27. Sun, X. (2026) 'AIRCREWSCHED: A meta review of aircrew scheduling and its research challenges', *Transportation Research Part E: Logistics and Transportation Review*. Available at: https://doi.org/10.1016/j.tre.2025.104525.

28. Wang, N. (2026a) 'Integrated optimization of airport slot allocation and boarding gate assignment', *Journal of Air Transport Management*. Available at: https://doi.org/10.1016/j.jairtraman.2025.102930.

29. Wang, T. (2026b) 'Improving fairness in airline integrated crew scheduling: Colleague familiarity and team experience distribution considerations', *IEEE Transactions on Engineering Management*, 73, pp. 3529–3542. Available at: https://doi.org/10.1109/TEM.2026.3693257.

30. Wang, Z., Wang, Y. and Hansen, M. (2026) 'Distributional multi-agent reinforcement learning for air traffic flow and capacity management in a multiple-airport system', *Transportation Research Part E: Logistics and Transportation Review*. Available at: https://doi.org/10.1016/j.tre.2026.104787.

31. Wu, S., Jia, Y., Chen, X. and Qu, D. (2026) 'Integrated airline recovery under uncertain disruptions: A fuzzy programming approach', *Applied Sciences*, 16(8). Available at: https://doi.org/10.3390/app16083667.

32. Xiao, F., Liao, Y., Chen, C. and Liang, Z. (2026) 'Effective features for machine learning-enhanced aircraft recovery', *Transportation Research Part B: Methodological*. Available at: https://doi.org/10.1016/j.trb.2026.103566.

33. Yuan, M., Zhou, C. and Ribeiro, N.A. (2026) 'An exact method for storage space assignment at airport cargo terminal: A temporal-spatial packing problem', *European Journal of Operational Research*. Available at: https://doi.org/10.1016/j.ejor.2026.01.033.

34. Zhu, L. (2026) 'Joint fleet scheduling and cargo flow allocation for air cargo services', *Transportation Research Part B: Methodological*. Available at: https://doi.org/10.1016/j.trb.2026.103469.