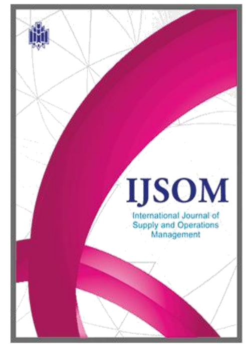
International Journal of Supply and Operations Management

IJSOM

2025, Volume 12, Issue 4, pp. 524-546
ISSN-Print: 2383-1359
ISSN-Online: 2383-2525
www.ijsom.com



Interaction of Flight Scheduling and Ticket Pricing: A Modern Data-Driven Approach Based on Distributionally Robust Optimization and Bi-Level Programming

Ali Papi [a], Mir Saman Pishvaee [a*] and Hadi Sahebi [a]

[a] *School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*

Abstract

In airline planning, tactical decisions related to flight schedule design and fleet assignment play a pivotal role in enhancing operational efficiency and maximizing revenue. On the other hand, ticket pricing, directly influencing market share, is inherently affected by the tactical flight timetable, market uncertainties, and passenger choice behavior. To jointly optimize tactical scheduling decisions and ticket pricing policies, and create optimal interaction between them, this paper proposes a modern data-driven decision-making framework that blends Distributionally Robust Optimization (DRO) with Bi-Level Programming (BLP). In this framework, leveraging historical data and machine learning algorithms, a distributional ambiguity set is first constructed to model uncertainty within the DRO framework. The BLP formulation then captures the interaction between flight scheduling (upper level) and ticket pricing (lower level). Additionally, passengers' choice behavior is incorporated using a Multinomial Logit (MNL) discrete choice model. To address the computational complexity, a column-and-constraint generation (CCG) algorithm is adopted, enabling model decomposition and enhancing computational efficiency. Finally, the proposed model and solution framework are validated through a case study and a series of numerical experiments. Numerical results demonstrate that, compared to classical approaches, the proposed framework significantly improves market share and airline revenue, ensures robustness against uncertainty and passenger behavior variability, and enhances computational tractability.

Keywords: Data-Driven Decision-Making; Distributionally Robust Optimization; Bi-Level Programming; Flight Scheduling; Ticket Pricing; Decomposition Method.

Introduction

Air mobility is a cornerstone of economic development, international trade, human connectivity, and global integration. Due to its exceptional speed and efficiency, the aviation sector plays an indispensable role in meeting origin-destination (O-D) travel demand and fostering a competitive environment for airlines. The enhanced accessibility and the rapidly growing demand for efficient transport services have considerably expanded air operations and improved the industry's profitability. According to reports by the International Air Transport Association (IATA), over 5 billion passengers travel on commercial flights annually (IATA, 2019), and this number is projected to grow at an average annual rate of approximately 4.1% through 2045 (Zhou, Liang, Chou, &

*Corresponding author email address: pishvaee@iust.ac.ir
DOI: 10.22034/ijsom.2025.110846.3431

Chaovalitwongse, 2020). This significant revenue potential has motivated airlines to adopt data-driven optimization methods, particularly those that explicitly address demand uncertainty and passenger preferences in their planning and decision-making processes.

The airline industry is characterized by several distinct characteristics: resource constraints, high capital investment requirements, strict regulatory frameworks, a competitive and uncertain market environment, highly interdependent decisions, and complex planning processes. Therefore, a primary objective for airlines is to maximize the utilization of available resources, such as flight routes, aircraft fleets, crew rostering, and airport slots, by deploying advanced optimization models and leveraging data analytics within the operations research framework (Barnhart & Cohn, 2004; Wandelt et al., 2025).

As shown in Figure 1, the airline planning process can be categorized across three decision-making horizons: strategic, tactical, and operational. For the strategic horizon, long-term decisions such as fleet acquisition and network design are addressed. Subsequently, tactical decisions, including flight scheduling and fleet assignment, are made, followed by operational decisions related to resource allocation. Given their critical influence on efficiency and profitability, this study focuses on optimizing tactical planning decisions, particularly flight timetables and fleet assignment. Tactical flight planning comprises three interconnected subproblems: frequency planning, which determines the number of daily flights on each O-D pair; timetable development, which specifies the departure and arrival times for each flight; and fleet assignment, which involves allocating the appropriate aircraft to each scheduled flight.

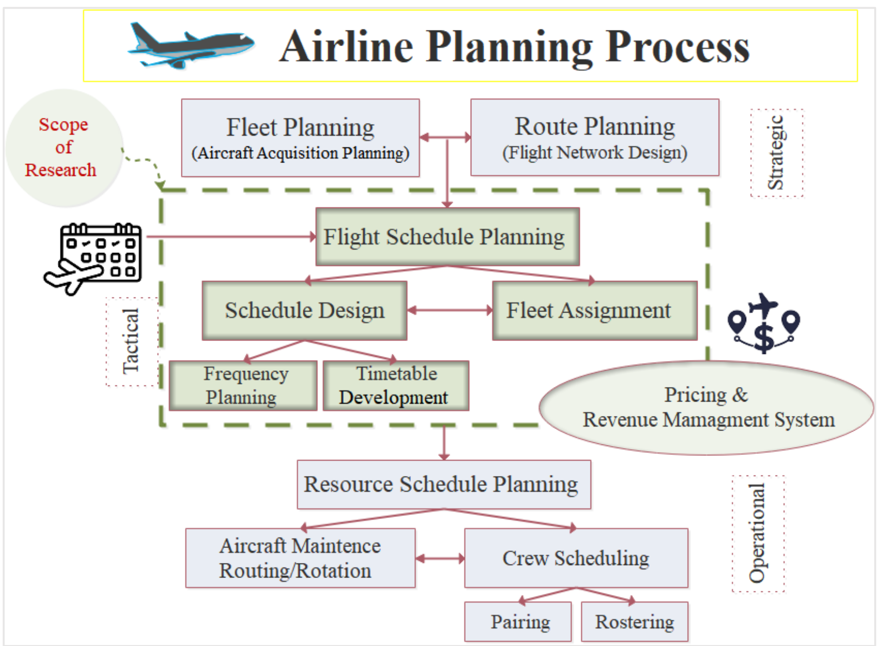


Figure 1. Decision optimization problem within airline planning process

Due to the dynamic nature of the market, intense competition, and sensitivity to passenger behavior, airlines adopt decision-support models only if they meet seven essential criteria: (1) integrated decision-making, especially for flight-scheduling decisions; (2) robustness, adaptability, and recoverability under uncertainty and operational disruptions; (3) competitive awareness, including the consideration of rivals' actions ; (4) customer-orientation, by incorporating passenger preferences to mitigate demand loss; (5) data-driven intelligence, achieved through analytics and machine learning; (6) coordinated pricing and revenue management, ensuring that pricing policies and flight-planning decisions interact effectively to maximize market share and profitability; and (7) computational efficiency, enabled by decomposition techniques and efficient solution algorithms.

Among these features, Decision integration is particularly critical. Modern flight-planning models that address interdependencies between scheduling and pricing avoid the inefficiencies inherent in decoupled approaches and thus enhance profitability (Grosche & Grosche, 2009). Moreover, the inclusion of discrete choice models to capture passenger decision behavior has gained considerable attention in recent research (Cadarso & Vaze, 2023). This study advances that trend by explicitly modeling the interaction between tactical flight planning and ticket pricing, addressing the essential requirements of competitiveness and revenue management coordination.

Uncertainty modeling through data-driven analyses powered by machine-learning algorithms equips contemporary distributionally robust optimization (DRO) approaches with enhanced robustness properties and risk-mitigation capabilities (Bertsimas & Thiele, 2006; Özarık, da Costa, & Florio, 2024). To manage large-scale planning problems, decomposition-based algorithms (C. Yan, Barnhart, & Vaze, 2022) and metaheuristic search strategies have proven effective.

Accordingly, the aim of this paper is to provide a modern data-driven decision-making framework for the integrated optimization of tactical flight planning and ticket pricing. The framework aligns with the seven key features noted above. It employs a machine-learning-enhanced DRO model to manage uncertainty and ensure solution robustness. Bilevel programming (BLP) is used to coordinate hierarchical decision levels, while passenger choice behavior is modeled using an advanced MNL formulation. Lastly, computational efficiency is achieved through the application of a column-and-constraint generation (CCG) decomposition technique.

The rest of this paper is organized as follows. In Section 2, the relevant literature on airline planning and DRO models is reviewed. Section 3 formulates the optimization problem that captures the interaction between flight scheduling and ticket pricing decisions. Section 4 introduces the proposed methodological framework, which consists of data-driven DRO, BLP, MNL model, and CCG decomposition method. Section 5, based on the proposed methodological framework, presents the proposed model and solution method for the problem of joint optimization of flight scheduling and ticker pricing. Section 6 deals with the model implementation and result analysis. Finally, Section 7 concludes the paper and suggests some recommendations for future research directions.

Relevant Literature

This section provides a general overview of the relevant literature on airline planning problems. Since this research centers on airline tactical planning, we primarily focus on related works in this area. For a more comprehensive review of the entire airline planning process, the readers are referred to Eltoukhy, Chan, and Chung (2017) and Zhou et al. (2020). Furthermore, to specify the reason of selecting DRO approach, the various paradigms for optimization under uncertainty are classified. Finally, the research gaps which are filled through our contributions are summarized.

Studies on Airline Planning

The profitability of an airline is closely linked to its capacity to design flight schedules that effectively accommodate passenger demand in competitive markets. Lohatepanont and Barnhart (2004) is widely recognized as one of the pioneering and most influential works in the field of airline tactical planning, particularly addressing decisions related to schedule design and fleet assignment. Extending this research, Barnhart, Farahat, and Lohatepanont (2009) developed a fleet assignment model for specific sub-networks to improve the alignment between fleet capacity and demand. Further advancements were made by (Sherali, Bae, & Haouari, 2013), who proposed an integrated model combining schedule design, fleet assignment, and operational considerations such as aircraft rotation, maintenance, and flight rescheduling.

Several studies, including (Sa, Santos, & Clarke, 2020) and (Birolini, Jacquillat, Cattaneo, & Antunes, 2021), addressed strategic decisions such as fleet planning and network design. One of the earliest and most frequently cited works on tactical decision-making in this domain is (Lohatepanont & Barnhart, 2004), which concentrates on schedule design and fleet assignment. Building on this foundation, Weide, Ryan, and Ehrgott (2010) introduced an iterative optimization approach for the operational scheduling of fleet and crew. Given the emphasis of this paper, the current review primarily focuses on the literature of tactical decision-making.

Understanding passenger choice behavior is essential for aligning tactical flight planning with market demand. As a crucial determinant of realized demand, passenger preferences directly influence the effectiveness of scheduling and fleet assignment decisions. In this regard, S. Yan, Tang, and Lee (2007) proposed a flight scheduling model that incorporates passenger preferences in competitive markets. Similarly, Wei, Vaze, and Jacquillat (2020) introduced an optimization framework that integrates passenger preferences, capturing how schedule design and fleet assignment affect ticket purchase decisions. C. Yan et al. (2022) further refined this approach by incorporating price-based substitution behavior in route and cabin class selection.

Modeling uncertainty is critical for ensuring the robustness and reliability of airline operations. Stochastic programming and robust optimization are two well-established frameworks for decision-making under uncertainty. For instance, Sherali and Zhu (2008) developed a two-stage stochastic model for fleet assignment that explicitly addresses demand uncertainty. Similarly, Kenan, Jebali, and Diabat (2018) proposed a two-stage formulation for flight scheduling and fleet assignment, accounting for uncertainty in both demand and ticket pricing.

More recent studies have explored integrated models that account for multiple planning dimensions. For example, Wei et al. (2020) proposed a unified model for scheduling and fleet assignment, while Kiarashrad, Pasandideh, and Mohammadi (2021) introduced a nonlinear optimization model that integrates scheduling, fleet assignment, and ticket pricing, capturing competitive network effects. Kiarashrad et al. (2021) addressed supply-demand interactions through an integrated model for flight scheduling, fleet assignment, and demand generation, employing a nested logit framework to represent competition between routes. C. Yan et al. (2022) utilized a decomposition approach to manage network complexity in schedule design and fleet assignment. At the operational level, Ben Ahmed, Hryhoryeva, Hvattum, and Haouari (2022) explored the integration of fleet assignment with aircraft rotation and crew scheduling. Lastly, Xu, Adler, Wandelt, and Sun (2024) employed a game-theoretic framework to model competition among airlines and high-speed rail in the context of ticket pricing, schedule design, and fleet assignment.

These studies underscore the evolution of airline planning research from isolated sub problems toward integrated, data-driven, and passenger-centric models. Nevertheless, important research gaps remain, which are addressed in the next section.

Data-Driven Robust Optimization under Uncertainty

The increasing availability of data forms the foundation for designing data-driven decision-making techniques for the uncertain decision problems (Yang, Xu, Gong, & Rekik, 2024). Data-driven methods outperform traditional SP and RO approaches by leveraging prior knowledge for more accurate modeling (Papadimitriou, 2016). Data-driven decision-making is essential when historical data is available, as it provides a realistic representation of uncertainty sets using statistical tools, in contrast to the unrealistic assumptions, e.g., normal distribution, of classical methods (Peng & Delage, 2024). Classical approaches typically rely on fixed, predetermined parameters, such as known means and variances, which may not accurately capture real-world complexity. In contrast, data-driven methods leverage empirical data to estimate these parameters, enhancing model robustness.

As a modern data-driven approach for decision-making under uncertainty, data-driven robust optimization is a modern methodology improving optimality robustness and conservatism level by systematically incorporating historical or real-time data into robust optimization frameworks (Bertsimas & Thiele, 2006). Unlike traditional RO which depends on a priori theoretical assumptions to define uncertainty sets (Gorissen, Yanıkoğlu, & Den Hertog, 2015), DDRO employs statistical analysis, machine learning (ML), or distributionally robust techniques to construct uncertainty/ambiguity sets directly from empirical data. This paradigm enhances models’ adaptability to real-world variability while ensuring robustness against worst-case realization.

While conventional DDRO models primarily focus on constructing uncertainty sets through ML algorithms or statistical hypothesis testing (Bertsimas, Gupta, & Kallus, 2018a; Han, Shang, & Huang, 2021; Shang, Huang, & You, 2017; Zhang, Jia, He, & Chu, 2022), a more advanced paradigm leverages data to generate families of distributions for distributionally robust optimization models. Data-driven DRO approach synthesizes principles from RO and SP, ensuring robust expected performance under worst-case probability distributions while achieving enhanced performance guarantees across both in-sample and out-of-sample scenarios (Van Parys, Esfahani, & Kuhn, 2021).

DRO, as a modern extension of stochastic optimization models based on the distributional robustness characteristic, advances classical SP by optimizing decisions over an ambiguity set of probability distributions, thereby circumventing the risk of relying on a single, potentially misspecified, nominal distribution. The construction of an ambiguity set is a pivotal step in DRO models, as it delineates a family of probability distributions that are consistent with the observed data and are statistically similarly distributed. There exist several alternatives for generating this ambiguity set: (I) Moment-based approaches that utilize known statistical properties, such as means, variances, and higher-order moments (Delage & Ye, 2010); and (II) Distance-based or discrepancy-based methods that leverage various measures, such as metrics (e.g., Wasserstein (Gao, Chen, & Kleywegt, 2024; Mohajerin Esfahani & Kuhn, 2018)), φ-divergences (e.g., Kullback-Leibler (Ben-Tal, Den Hertog, De Waegenaere, Melenberg, & Rennen, 2013)), likelihood or statistical hypothesis test (Bertsimas et al., 2018a; Bertsimas, Gupta, & Kallus, 2018b; W. Liu, Yang, & Yu, 2023), total variation (Sun & Xu, 2016), and ℓ_p-norms (Huang, Qu, Gong, Goh, & Ji, 2020; Jiang & Guan, 2018), to construct a set of distributions that are sufficiently close to a nominal/reference distribution and constrained by a predetermined maximum distance. For a more detailed review of alternative approaches for generating ambiguity sets in DRO models, the reader is referred to Rahimian & Mehrotra (2019), Van Parys et al. (2021), and Zude Wang, Jiang, Lin, Tao, & Chen (2024).

DRO models have found diverse applications in various fields requiring robust decision-making under uncertainty. For instance: in finance, DRO enhances portfolio optimization and risk management (Costa & Kwon, 2022). Healthcare systems employ DRO for adaptive resource allocation and personalized treatment strategies under dynamic patient demands (Y. Wang, Zhang, & Tang, 2024). Transportation systems utilize DRO for network design, tactical planning, and traffic management (L. Liu, Song, Wang, & Zhang, 2022; C. Wang, Liu, Zhang, & Miao, 2024; R. Yan, Liu, & Wang, 2024). Supply chain operations apply DRO for strategic network design, optimal distribution, inventory control, and logistics management (Gilani & Sahebi, 2022; Gong & Zhang, 2022; Li, Liu, & Yang, 2024). Beyond these domains, DRO is increasingly adopted in energy systems (Lu & Zhou, 2024; Mohseni & Pishvaee, 2023; Zechen Wang, Liu, & Huo, 2025), manufacturing and production planning (Metzker, Thevenin, Adulyasak, & Dolgui, 2025; Razavi Al-e-hashem, Papi, Pishvaee, & Rasouli, 2022), pricing and revenue management systems (Chen & Hu, 2023; Qiu, Sun, Zhou, & Sun, 2023), confirming its broad applicability. This applicability motivates our research paper to adopt a novel data-driven DRO model to address uncertainty in the aviation industry.

Research Gaps and Proposed Contributions

An analysis of the existing literature reveals several critical gaps that warrant further investigation: (I) Limited integration across decision-making levels. (II) Inadequate modeling of passenger choice behavior. (III) Underutilization of data-driven optimization. (IV) Insufficient emphasis on computational efficiency. To fill these gaps, this study contributes to the literature on tactical flight planning through the following innovations:

1. Integrated Optimization Framework: The paper develops a comprehensive model that simultaneously optimizes flight frequency, scheduling, and fleet assignment, while coordinating these tactical decisions with operational ticket pricing policies. The framework accommodates strategic and operational constraints, including route structure, fleet availability, slot and gate limitations, and passenger behavioral responses.

2. Data-Driven Machine Learning Integration: A novel DRO approach is proposed, enriched with machine learning techniques to extract statistical insights from historical data. This hybrid approach enables the construction of a data-driven ambiguity set that enhances decision robustness against demand uncertainty.

3. Enhanced Passenger Choice Modelling: Passenger decision-making is incorporated using an advanced MNL formulation, which considers booking decisions. The model supports tractability through appropriate linearization techniques, allowing for realistic yet computationally feasible analysis of passenger behaviour.

4. Computational Tractability: a decomposition-based CCG method is adopted to efficiently solve the final data-driven DRO model.

Problem Description

This section formally describes the concern optimization problem, referred to as the Interaction of Flight Scheduling and Ticket Pricing (IFSTP). In the airline operations context, tactical flight planning serves as a bridge between strategic decisions, such as network design and fleet acquisition, and operational planning tasks, including aircraft and crew assignment. The primary objective at this stage is to maximize profitability by optimizing tactical decisions, namely flight frequency, timetable development, and fleet assignment, while adhering to various operational constraints such as fleet capacity, seating availability, and slot and gate limitations.

In IFSTP problem, these tactical decisions are coordinated with operational ticket pricing strategies to enhance the attractiveness of the flight schedule, accounting for both demand uncertainty and the flexibility to adjust prices. The goal is to increase demand and improve market share through better alignment of schedule planning and price optimization. Consequently, IFSTP problem involves four interrelated decision variables: (i) determining flight frequency for each O-D air travel market across different periods; (ii) developing the flight timetable; (iii) assigning the appropriate fleet type to the selected flight; and (iv) setting operational ticket prices. Beyond satisfying strategic and operational constraints, the model must also ensure consistency in the space–time network, meaning that the arrival and departure of each fleet type at every airport and time period must be properly synchronized.

One of the fundamental challenges considered in the IFSTP problem is the inherent uncertainty in travel demand. To address this, the model must provide a robust solution derived from historical data that not only aims to maximize expected performance but also mitigates risk under adverse demand scenarios. Importantly, tactical flight scheduling decisions are made before the realization of actual demand, whereas operational ticket pricing decisions can be adjusted adaptively once demand information becomes available.

Another critical feature of the model is the explicit incorporation of passenger choice behavior, which significantly influences realized demand. Each flight itinerary is characterized by attributes such as ticket price, departure time, service level, and airport access time or cost. These attributes affect the perceived utility of each option, and their respective weights are used to estimate the probability that passengers will select a particular flight. This behavior is integrated into IFSTP problem to more accurately reflect market dynamics. In cases of mismatched supply and demand, or due to specific passenger preferences, some demand may be lost. However, such demand can potentially be recaptured through adjustments to the flight schedule or pricing strategies.

In addition to demand uncertainty and passenger utility weights, other influential parameters include fleet capacities, airport constraints, flight durations, fixed and variable costs, and others, as summarized in Table 1. IFSTP model captures these factors to ensure practical feasibility and solution robustness.

In IFSTP problem, the concept of interaction between flight scheduling and ticket pricing is represented as a hierarchical, Bi-level framework. At the outer level (the leader), tactical planning decisions, such as flight frequency, timetable, and fleet assignment, are optimized without complete knowledge of final demand or prices. At the inner level (the follower), pricing decisions are made in response to the upper-level decisions and the observed demand, to maximize revenue. The results of this lower-level optimization, including demand and revenue, are fed back into the upper level, which updates the flight schedule accordingly. This iterative process continues until convergence is achieved; i.e., when there is no longer an incentive to modify the flight schedule, yielding an integrated plan that is coherent with both market conditions and operational constraints. Next section provides the proposed solution framework to solve IFSTP problem.

Proposed Data-Driven and Interactive Solution Framework

To solve IFSTP problem, this study proposes a modern, data-driven decision-making framework that integrates advanced techniques from robust optimization, hierarchical modelling, and behavioral analysis. Specifically, the framework incorporates a data-driven DRO approach to capture uncertainty in O-D market demand, a BLP structure to model the interaction between scheduling and pricing decisions, an MNL model to represent passenger choice behavior, and a CCG algorithm to enhance computational efficiency through decomposition. Figure 2 illustrates the overall architecture of the proposed framework, and the key components are described below.

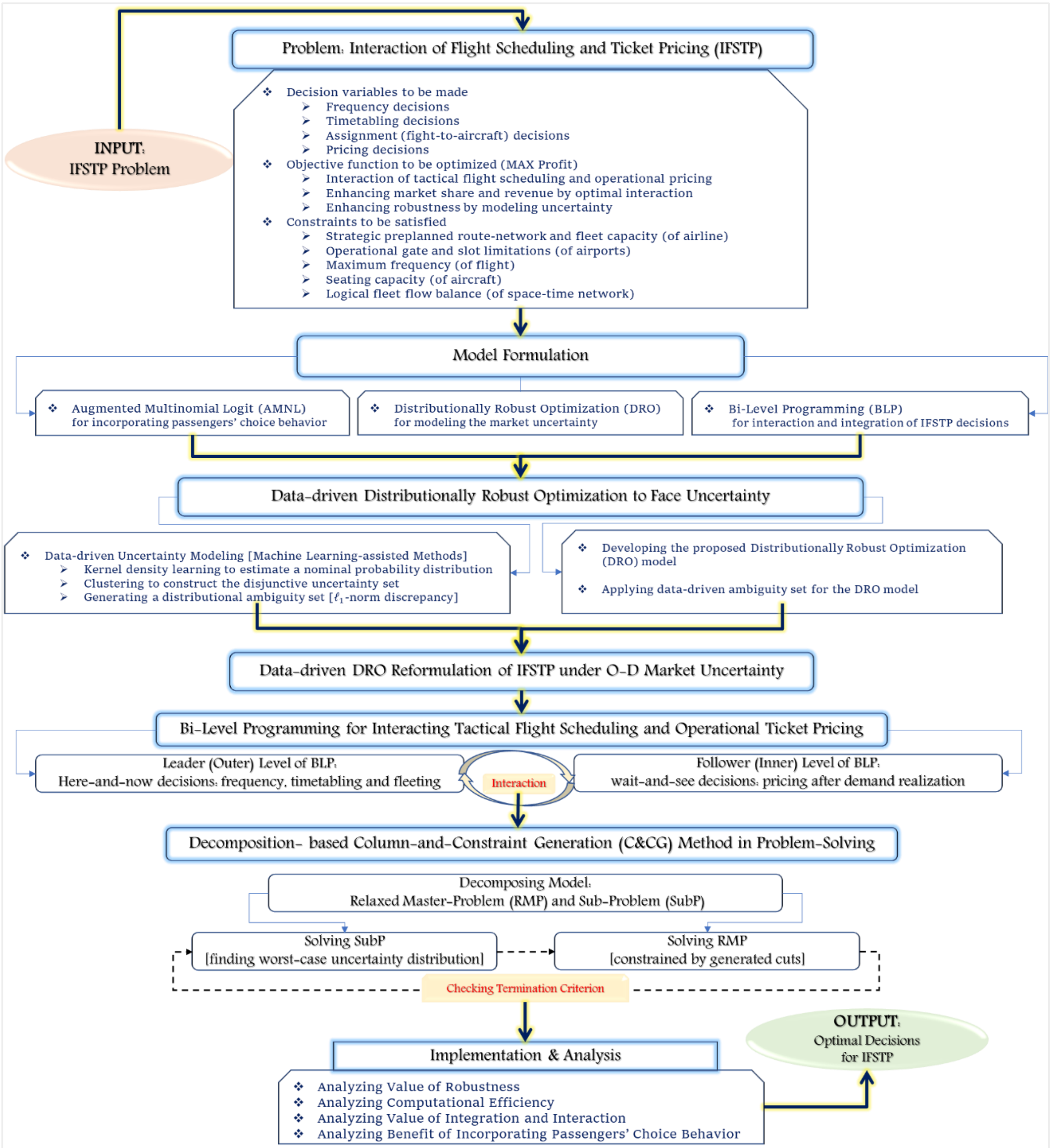


Figure 2. Proposed methodological framework for solving IFSTP problem

Data-Driven DRO for Decision-Making under Uncertainty

In general, an optimization under uncertainty problem is expressed as follows:

$$\min_{x \in \mathbb{X}} F(x) + \mathcal{R}(x;\theta) \quad (1)$$

where $x \in \mathbb{X}$ represents the decision variables, and θ denotes the uncertain parameters. Additionally, $F(x)$ and $\mathcal{R}(x;\theta)$ represent the uncertainty-independent and uncertainty-dependent parts of the objective function, respectively.

If, in addition to x, wait-and-see variables are also included in Model (1), $\mathcal{R}(x;\theta)$ is determined by solving an inner optimization problem, known as the recourse problem, that depends on the first-stage variables x and the observed uncertainty θ.

If a support set Θ for the uncertain parameters θ, along with a specified probability distribution f_Θ is available, SP approach is used, optimizing the expected performance under the distribution f_Θ. If only the support set Θ is considered, and no probability distribution is available, RO approach optimizes the performance under the worst-case scenario within the support set Θ. Finally, in DRO, instead of a single probability distribution, a distributional ambiguity set (DAS) $\mathcal{P}$ is constructed for θ, especially based on historical data, and the expected performance of the system under the worst distribution belonging to $\mathcal{P}$ is optimized. Therefore, the formulation of an uncertain optimization problem using SP, RO, and DRO approaches can be summarized as follows:

$$\min_{x\in\mathbb{X}} F(x)+\mathcal{R}(x;\theta) \; \underbrace{\overrightarrow{\text{Reformulation under Uncertain }\theta}} \begin{cases} \text{SP} \coloneqq \min\limits_{x\in\mathbb{X}} F(x)+\mathbb{E}_{f_\Theta}[\mathcal{R}(x;\theta)] \\ \text{RO} \coloneqq \min\limits_{x\in\mathbb{X}} F(x)+\sup\limits_{\theta\in\Theta}\{\mathcal{R}(x;\theta)\} \\ \text{DRO} \coloneqq \min\limits_{x\in\mathbb{X}} F(x)+\sup\limits_{f_\Theta\in\mathcal{P}}\{\mathbb{E}_{f_\Theta}[\mathcal{R}(x;\theta)]\} \end{cases} \tag{2}$$

To construct the proposed data-driven DAS, we first consider $\mathfrak{D}=\{\theta^o=(\theta_1^o,\theta_2^o,\dots\theta_m^o) \mid o=1,2,\dots,N\}$ as the dataset available for the uncertain parameter θ. Then, the Kernel Density Estimation (KDE) method is employed to define an empirical/nominal probability distribution. As a non-parametric technique, KDE does not require prior assumptions and estimates the probability distribution as follows (Parzen, 1962; Węglarczyk, 2018):

$$\widetilde{f_\Theta}(\theta)=\frac{1}{N(\prod_{m=1}^{M}\hbar_m)}\sum_{o=1}^{N}\left(\prod_{m=1}^{M}\mathcal{K}\left(\frac{\theta_m-\theta_m^o}{\hbar_m}\right)\right)\;;\;\theta=(\theta_1,\theta_2,\theta_3,\dots,\theta_M) \tag{3}$$

where $\mathcal{K}(u)=\frac{1}{\sqrt{2\pi}}e^{-u^2/2}$ is a Gaussian kernel satisfying the conditions $\mathcal{K}(u)\geq 0$ and $\int_{u\in\mathbb{R}}\mathcal{K}(u)du=1$. Furthermore, $\hbar>0$ denotes the smoothing parameter. Empirically, the smoothing parameter is set as $\hbar_m=\widehat{\sigma_m}\left(\frac{4}{3N}\right)^{\frac{1}{5}}$ where $\widehat{\sigma_m}$ represents the estimated standard deviation of the component m in the dataset $\mathfrak{D}$ (Silverman, 2018).

After constructing a nominal distribution from the available dataset, the ℓ_1-norm is employed to measure the distance between probability distributions:

$$\mathrm{d}_{\ell_1}(f_\Theta,g_\Theta)\coloneqq\|f_\Theta(.)-g_\Theta(.)\|_1=\int_{\theta\in\Theta}|f_\Theta(\theta)-g_\Theta(\theta)|\,d\theta \tag{4}$$

Based on this, the following DAS is formulated to generate distributions close to the nominal distribution $\widetilde{f_\Theta}(\theta)$:

$$\mathcal{P}_{\ell_1}\coloneqq\left\{f_\Theta\in\mathfrak{B}(\Theta)\,\middle|\,\left\|f_\Theta(.)-\widetilde{f_\Theta}(.)\right\|_1\leq\delta\;,\int_{\theta\in\Theta}f_\Theta(\theta)\,d\theta=1\right\} \tag{5}$$

where $f_\Theta\in\mathfrak{B}(\Theta)$ denotes a real-valued Borel-measurable function, whose distance from the empirical distribution $\widetilde{f_\Theta}$ does not exceed δ.

To determine the parameter δ in a data-driven manner, Hoeffding's inequality can be utilized, establishing that $\text{Prob}\left(\left\|\widetilde{f_\Theta}-f_\Theta\right\|_1\leq\delta\right)\geq 1-\;2\exp\left(-\frac{2N\delta^2\prod_{m=1}^{M}\hbar_m^2}{\bar{K}^{2M}}\right)$. Thus, it can be concluded that, at a confidence level of $1-\epsilon$, the maximum distance between the true distribution of the uncertain parameters θ and the nominal distribution $\widetilde{f_\Theta}$ can be computed as (Hoeffding, 1994):

$$\delta=\frac{\bar{K}^M}{\prod_{m=1}^{M}\hbar_m}\sqrt{\frac{1}{2N}\ln\left(\frac{2}{\epsilon}\right)} \tag{6}$$

in which $\overline{K} = \sup_{u\in\mathbb{R}} \mathcal{K}(u)$.

To account for disjunctions and multiple scenarios in the dataset $\mathfrak{D}$, and to enable second-stage decision variables to adapt to each scenario, the dataset $\mathfrak{D}$ is classified into κ clusters using the Fuzzy C-Means (FCM) algorithm (Bezdek, Ehrlich, & Full, 1984), thereby discretizing DAS defined in Equation (5). The optimal number of clusters κ is determined via a Golden-Section Search that maximizes between-cluster variance (BCV) while minimizing within-cluster variance (WCV). We then define the scenario set $\Xi = \{1,2,\dots,\kappa\}$ and construct the finite support set Θ as follows:

$$\Theta \coloneqq \left\{ \theta_\xi = \frac{\sum_{o=1}^{N} \mu_{o\xi}^* \theta^o}{\sum_{o=1}^{N} \mu_{o\xi}^*} \middle| \xi \in \Xi \right\} \tag{7}$$

where $\mu_{o\xi}^*$ is the output of the FCM algorithm and represents the degree of membership of each observation to each cluster. For each scenario $\xi \in \Xi$, the empirical probability $\pi_\xi^0 = \frac{1}{\sum_{\xi'=1}^{\kappa}\left(\sum_{o=1}^{N} \mu_{o\xi'}^* \widetilde{f_\Theta}(\theta^o)\right)} \sum_{o=1}^{N} \mu_{o\xi}^* \widetilde{f_\Theta}(\theta^o)$ can be estimated using the derived KDE and the membership degrees of data points within the respective cluster. By aggregating these probabilities, the nominal discrete distribution $\boldsymbol{\pi}^0 = [\pi_1^0, \pi_2^0, \dots, \pi_\kappa^0]^\top$ is considered as a substitute for Equation (3). Finally, the proposed DAS based on the norm ℓ_1 is constructed as follows:

$$\mathcal{P}_{\ell_1} = \left\{ \boldsymbol{\pi} = [\pi_\xi]_{\xi\in\Xi} \in \mathbb{R}_+^\kappa \middle| \, \|\boldsymbol{\pi} - \boldsymbol{\pi}^0\|_1 \le \delta \, , \boldsymbol{e}^\top \boldsymbol{\pi} = 1 \right\} \tag{8}$$

where $e = [1,1,\dots,1]^\top$ is the unit vector, and δ represents the data-driven distance derived from Equation (6). Consequently, the proposed DRO model is formulated as follows:

$$\begin{aligned} \text{DRO} \coloneqq \min_{x\in\mathbb{X}} F(x) + \max_{\boldsymbol{\pi}=[\pi_\xi]_{\xi\in\Xi}\in\mathcal{P}_{\ell_1}} \left\{ \sum_{\xi\in\Xi} \pi_\xi \mathcal{R}(x;\theta_\xi) \right\} s.t.\ \mathcal{P}_{\ell_1} \\ = \left\{ \boldsymbol{\pi} = [\pi_\xi]_{\xi\in\Xi} \middle| \begin{array}{c} \sum_{\xi\in\Xi} \upsilon_\xi \le \delta \\ \pi_\xi - \pi_\xi^0 \le \upsilon_\xi \ \forall \langle \xi \in \Xi \rangle \\ \pi_\xi^0 - \pi_\xi \le \upsilon_\xi \ \forall \langle \xi \in \Xi \rangle \\ \sum_{\xi\in\Xi} \pi_\xi = 1 \\ \pi_\xi \ge 0 \ \forall \langle \xi \in \Xi \rangle \\ \upsilon_\xi \ge 0 \ \forall \langle \xi \in \Xi \rangle \end{array} \right\} \end{aligned} \tag{9}$$

Multinomial Logit Model for Incorporating Discrete Choices

Passenger decisions are influenced by various factors, such as ticket price, departure time, trip duration, airport access time/cost, service quality, and flight punctuality. Each factor contributes to the overall utility of a flight option, with associated weights reflecting the relative importance perceived by passengers. Given the set of factor values $\boldsymbol{V}$ and their corresponding weights β, the utility of each flight option is computed through a compensatory–non-compensatory relationship as follows:

$$U = F(V,\beta) \to \widehat{U} = \frac{1}{2}\left(\sum_{i\in\mathcal{J}} \beta_i V_i + \min_{i\in\mathcal{J}} \beta_i V_i \right) \tag{10}$$

where $\mathcal{J}$ is the set of factors, and β_i and V_i represents the weight and the normalized value of each factor, respectively.

After estimating the utility of each available flight option, discrete choice models (DCMs) can be applied to determine the probability of passengers selecting a given alternative. A prominent DCM in air transportation research is the Multinomial Logit Model (MNL), which serves as the foundation for this analysis (Cao, Kleywegt, & Wang, 2022; Wei et al., 2020). This study employs an augmented version of the MNL model, named AMNL, in which the sensitivity to utility is explicitly incorporated. This sensitivity parameter reflects passengers' tendency to favor a high-utility option over alternatives with lower utilities.

Given a set $\mathcal{L}$ of flight options offered by an airline, where $\widehat{U_l}$ denotes the utility of each flight $l \in \mathcal{L}$ and $\widehat{U_0}$ represents the utility of competing alternatives, the probability of selecting each option under the proposed AMNL model is calculated as follows:

$$\wp_l = \frac{e^{\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{U_l}}}{\sum_{l' \in \mathcal{L}} e^{\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{U_{l'}}} + e^{\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{U_0}}} \tag{11}$$

where $0 < \vartheta < 1$ is the augmenting parameter to reflect passengers' sensitivity to utility differences. As ϑ increases, passengers are more likely to choose the option with the highest utility; as it decreases, their selections become more random based on uniform distribution. Using the selection probability of each flight option, the maximum attainable market share for each option can be considered equal to its probability, while the total market share captured by all options sums to one. This property allows AMNL function to be reformulated using linear relationships, which will be used in the proposed optimization model.

Bi-level Programming for Decisions Interaction

In decision-making processes involving interactions between primary/tactical decisions (e.g., flight schedule development) and secondary/operational decisions (e.g., ticket pricing), a nested optimization framework can effectively capture their mutual dependencies. For this purpose, BLP models have been proposed (Kalashnikov, Dempe, Pérez-Valdés, Kalashnykova, & Camacho-Vallejo, 2015). Specifically, BLP models possess a leader-follower structure in which the outer/leader-level decisions must be optimized while considering the optimality of the inner/follower-level decisions, and the inner-level optimization itself is conditioned on the leader's decisions. This action-reaction structure, also known as a Stackelberg game, enables the interaction between flight scheduling and ticket pricing in a way that, first, flight schedule development takes into account ticket pricing optimization and the pursuit of greater market share, and second, ticket pricing decisions are made with the consideration of the designed flight schedule.

In general, BLP models are presented as follows:

$$\begin{cases} \max\limits_{x_1 \in \mathcal{S}_1(x_2^*)} F_1(x_1; x_2^*) \\ \quad s.t. \\ x_2^* = \operatorname*{argmax}\limits_{x_2 \in \mathcal{S}_2(x_1)} F_2(x_2; x_1) \end{cases} \tag{12}$$

In BLP Model (12), the optimal value of the upper-level objective function, $F_1(x_1; x_2^*)$, and its corresponding solution space, $\mathcal{S}_1(x_2^*)$, depend on the optimal response of the lower-level decision variables (x_2^*), whereas the feasible region $\mathcal{S}_2(x_1^*)$ and the optimality status of the lower level are also influenced by the leader's decision (x_1).

One classical approach to solving the aforementioned BLP model involves replacing the optimality constraints of the inner optimization problem, with equivalent Karush-Kuhn-Tucker (KKT) conditions. This reformulation converts the BLP into a Mathematical Programming with Equilibrium Constraints (MPEC). Other solution approaches for BLP models include heuristic search techniques and evolutionary heuristic algorithms (Calvete, Gale, & Mateo, 2008). Moreover, in situations where the lower-level decisions directly affect either the feasible space or the objective value of the upper-level problem, an interactive and iterative solution approach can be employed.

In the proposed interactive solution method, an initial solution for the upper-level problem is fed into the lower-level problem. Given this fixed upper-level solution, the lower-level problem is solved to obtain the corresponding x_2^*. The obtained value of x_2^* is then used to update the upper-level decision if necessary. This iterative process continues until the upper-level decision no longer tends to change, at which point an equilibrium solution is reached.

Proposed Optimization Model and Solution Method

This section presents the mathematical formulation of the IFSTP problem, based on the proposed DRO and BLP frameworks. It also describes the interactive, decomposition-based solution approach used to efficiently solve the resulting model. All mathematical symbols and notation used in the formulation are summarized in Table 1.

Table 1. Symbols used to problem formulation

Sets	
$\mathcal{A}$	Set of airports.
$\mathcal{M}$	Set of air-travel O-D markets that can be served by the airline ($\mathcal{M} \subseteq \mathcal{A} \times \mathcal{A}$).
$\mathcal{P}$	Set of dayparts (Early Morning (4 AM - 8 AM); Morning (8 AM - 11 AM); Midday (11 AM - 2 PM); Afternoon (2 PM - 6 PM); Evening (6 PM - 9 PM); Night (9 PM – 12 AM (Midnight)).
$\mathcal{T}$	Set of all 30-minute time segments in a day ($\mathcal{T} = \{1,2,3, \dots, \lvert\mathcal{T}\rvert\}$).
$\mathcal{T}^p$	Subset of time segments belongs to daypart $p \in \mathcal{P}$ during which a flight leg can depart the origin airport.
$\mathcal{F}$	Set of aircraft fleet types.
$\mathcal{L}_m$	Set of time-based replications of a flight leg can be executed to serve air-travel market $m \in \mathcal{M}$.
$\mathcal{L}$	Set of all potential flight to be executed to serve O-D market ($\mathcal{L} = \bigcup_{m \in \mathcal{M}} \mathcal{L}_m$).
$\mathcal{B}$	Subset of airport for initial location of airline aircrafts.*
Ξ	Uncertainty set.
$\mathcal{P}_{\ell_1}$	Ambiguity Set.
Parameters	
$d_{mp\xi}$	Demand for air-travel market $m \in \mathcal{M}$ during daypart $p \in \mathcal{P}$ (Uncertain).
π	Probability distribution over uncertainty set Ξ
MF_{mp}	Maximum flight frequency can be operated to serve the market $m \in \mathcal{M}$ during daypart $p \in \mathcal{P}$.
ρ_l^0	Fare of flight leg $l \in \mathcal{L}$ to serve one demand of the corresponding market.
OC_{lf}	Operating cost for flight leg $l \in \mathcal{L}$ operated by fleet type $f \in \mathcal{F}$.
SC_m	Setup cost of serving air-travel market $m \in \mathcal{M}$ once.
$(O_l - D_l)$	Origin and Destination airports of flight leg $l \in \mathcal{L}$ serving an O-D market.
δ_l	Departure time of flight leg $l \in \mathcal{L}$.
Δ_l	Duration of flight leg $l \in \mathcal{L}$.
V_f	Seating capacity, i.e., the number of seats, of each aircraft in fleet type $f \in \mathcal{F}$.
NA_f	Number of available aircraft of fleet type $f \in \mathcal{F}$.
SA_{ap}	Slot availability at airport $a \in \mathcal{A}$ during daypart $p \in \mathcal{P}$.
GA_{ap}	Gate availability at airport $a \in \mathcal{A}$ during daypart $p \in \mathcal{P}$.
$\widehat{U}_{lmp}$	Utility estimated for flight leg $l \in \mathcal{L}$ in serving demand $\tilde{d}_{mp}$.
ϑ_m	Adjusting coefficient that represents the intensity of utility sensitivity in market $m \in \mathcal{M}$.
Tactical Decisions – Flight Scheduling and Fleet Assignment (Upper-Level of BLP)	
x_{fl}	Binary variable: if flight leg $l \in \mathcal{L}$ is assigned to fleet type $f \in \mathcal{F}$, 1; otherwise, 0.
n_{mp}	Frequency of flights to serve market $m \in \mathcal{M}$ during daypart $p \in \mathcal{P}$. (dependent on $\boldsymbol{x}$)
g_{fa}^t	Number of ground aircraft belonging to the fleet $f \in \mathcal{F}$ in space-time node $(a, t) \in \mathcal{N}$.
g_{fa}^0	Number of ground aircraft belonging to the fleet $f \in \mathcal{F}$ initially located at airport $a \in \mathcal{A}$.
Operational Decision – Ticket Pricing and Market Share (Lower-Level of BLP)	
$\rho_{l\xi}$	Price of flight leg $l \in \mathcal{L}$ under scenario $\xi \in \Xi$.
s_{lmp}	Share of market can be achieved by flight leg $l \in \mathcal{L}$ from demand $d_{mp\xi}$.
$\zeta_{l\xi}^+$	Auxiliary variable to determine capacity shortage of flight leg $l \in \mathcal{L}$ under scenario $\xi \in \Xi$.
$\mathcal{R}$	Operational revenue obtained by selling tickets.

Proposed DRO-BLP Model for IFSTP Problem

In this section, the proposed BLP model, incorporating the DRO approach to handle uncertainty, is presented. Constraints (13)-(23) represent the upper-level of BLP model, where the tactical decision-making for the flight schedule is optimized. Constraints (24)-(29) formulate the lower-level problem of BLP structure, focusing on the optimization of pricing decisions. It should be noted that the tactical flight scheduling decisions are treated as fixed parameters in the lower-level model, while the market share outcomes derived from pricing decisions in the lower-level model are considered fixed parameters in the upper-level model.

$$\textbf{MAX} \;\; \overline{\mathcal{R}} - \left(\sum_{m \in \mathcal{M}} \sum_{p \in \mathcal{P}} SC_m n_{mp} + \sum_{f \in \mathcal{F}} \sum_{l \in \mathcal{L}} OC_{fl} x_{fl} \right) \tag{13}$$

$$\sum_{f \in \mathcal{F}} x_{fl} \le 1 \quad \forall \langle l \in \mathcal{L} \rangle \tag{14}$$

$$\sum_{a \in \mathcal{B}} g_{fa}^0 = NA_f \quad \forall \langle f \in \mathcal{F} \rangle \tag{15}$$

$$\sum_{f \in \mathcal{F}} \sum_{a \in \mathcal{A} \setminus \mathcal{B}} g_{fa}^0 = 0 \tag{16}$$

$$g_{fa}^t = g_{fa}^{t-1} + \sum_{l \in \mathcal{L}:\, (\delta_l = t - \Delta_l \,\wedge\, D_l = a)} x_{fl} - \sum_{l \in \mathcal{L}:\, (\delta_l = t \,\wedge\, O_l = a)} x_{f,l} \quad \forall \langle f \in \mathcal{F}, a \in \mathcal{A}, t \in \mathcal{T} \rangle \tag{17}$$

$$\sum_{f \in \mathcal{F}} \sum_{l \in \mathcal{L}:\, (\delta_l \in \mathcal{T}^p \,\wedge\, O_l = a)} x_{fl} + \sum_{f \in \mathcal{F}} \sum_{l \in \mathcal{L}:\, (\delta_l + \Delta_l \in \mathcal{T}^p \,\wedge\, D_l = a)} x_{fl} \le SA_{ap} \quad \forall \langle a \in \mathcal{A}, p \in \mathcal{P} \rangle \tag{18}$$

$$\sum_{f \in \mathcal{F}} \sum_{l \in \mathcal{L}:\, (\delta_l \in \mathcal{T}^p \,\wedge\, D_l = a)} x_{fl} \le GA_{ap} \quad \forall \langle a \in \mathcal{A}, p \in \mathcal{P} \rangle \tag{19}$$

$$n_{mp} = \sum_{f \in \mathcal{F}} \sum_{l \in \mathcal{L}_m : \delta_l \in \mathcal{T}^p} x_{fl} \quad \forall \langle m \in \mathcal{M}, p \in \mathcal{P} \rangle \tag{20}$$

$$n_{mp} \le MF_{mp} \quad \forall \langle m \in \mathcal{M}, p \in \mathcal{P} \rangle \tag{21}$$

$$\sum_{f \in \mathcal{F}} V_f x_{fl} \ge \sum_{m \in \mathcal{M}} \sum_{p \in \mathcal{P}} d_{mp\xi} \overline{s}_{lmp\xi} \quad \forall \langle l \in \mathcal{L}, \xi \in \Xi \rangle \tag{22}$$

$$\begin{cases} x_{fl} \in \{0,1\} \\ n_{mp} \in \mathbb{Z}_+, g_{fa}^t \in \mathbb{Z}_+, g_{fa}^0 \in \mathbb{Z}_+ \end{cases} \tag{23}$$

$$\textbf{MAX} \;\; \mathcal{R} - \sum_{\xi \in \Xi} \pi_\xi^0 \sum_{l \in \mathcal{L}} \rho_{l\xi} \zeta_{l\xi}^+ \tag{24}$$

$$\mathcal{R} = \min_{\pi \in \mathcal{P}_{\ell_1}} \sum_{\xi \in \Xi} \pi_\xi \sum_{l \in \mathcal{L}} \rho_{l\xi} \sum_{m \in \mathcal{M}} \sum_{p \in \mathcal{P}} d_{mp\xi} s_{lmp\xi} \tag{25}$$

$$\mathcal{s}_{lmp\xi} \leq \frac{e^{\left(\frac{\vartheta_m}{1-\vartheta}\right)\widehat{U_l}}}{\sum_{l'\in\mathcal{L}} e^{\left(\frac{\vartheta_m}{1-\vartheta_m}\right)\widehat{U_{l'}}} + e^{\left(\frac{\vartheta_m}{1-\vartheta_m}\right)\widehat{U_0}}}\left(1+\frac{\rho_{l\xi}-\rho_{0_{l\xi}}}{\rho_{0_{l\xi}}}\right) \quad \forall \langle l\in\mathcal{L}, m\in\mathcal{M}, p\in\mathcal{P}, \xi\in\Xi\rangle \tag{26}$$

$$\sum_{l\in\mathcal{L}_m} \mathcal{s}_{lmp\xi} \leq 1 \quad \forall \langle m\in\mathcal{M}, p\in\mathcal{P}, \xi\in\Xi\rangle \tag{27}$$

$$\sum_{m\in\mathcal{M}}\sum_{p\in\mathcal{P}} d_{mp\xi}\mathcal{s}_{lmp\xi} \leq \sum_{f\in\mathcal{F}} V_f\overline{x}_{fl} + \zeta_{l\xi}^{+} \; \forall \langle l\in\mathcal{L}, \xi\in\Xi\rangle \tag{28}$$

$$\{\rho_{l\xi}\in\mathbb{R}_+, \mathcal{s}_{lmp\xi}\in\mathbb{R}_+, \zeta_{l\xi}^{+}\in\mathbb{R}_+ \tag{29}$$

Objective (13) aims to maximize profit, where $\overline{\mathcal{R}}$ denotes the total ticket revenue obtained from the lower-level optimization. Eq. (14) ensures that no more than one fleet type is assigned to any given flight. Constraints (15) and (16) regulate the initial positioning of aircraft at their designated home bases. Constraint (17) enforces network flow balance by managing the inflow and outflow of aircraft at each airport over time. Constraints (18) and (19) impose airport slot and gate availability limits across different time intervals. Constraints (20) and (21) control the number of flights operated in each O-D market and enforce upper bounds on flight frequency. Constraint (22) links the tactical and operational layers by requiring that fleet assignment decisions accommodate the estimated passenger demand derived from the pricing model. Finally, Constraint (23) defines the feasible domains of the tactical decision variables.

At the lower-level of BLP model, Constraint (24) formulates the operational objective function associated with the average revenue generated from ticket sales, wherein the opportunity cost due to capacity shortages has also been deducted. Constraint (25), based on DRO approach, calculate the expected profit under worst-case probability distribution of uncertain demand. Passenger choice behavior is incorporated in Constraint (26) using the proposed AMNL model, which estimates maximum market shares based on perceived utility and adjusts them according to price differences relative to competitors ($\frac{\rho_{l\xi}-\rho_{0_{l\xi}}}{\rho_{0_{l\xi}}}$). Constraint (27) ensures that the cumulative market share of all flights within each O-D market does not exceed one. Constraint (28) accounts for flight capacity limitations and introduces an auxiliary variable $\zeta_{l\xi}^{+}$ to capture shortages. Finally, the feasible domains for the lower-level decision variables (prices, market shares, and capacity shortages) are defined in Constraint (29). In the next section, an interactive and decomposition-based solution method is presented to efficiently solve the proposed model.

Decomposition-based Column-and-Constraint Generation Method

To improve computational efficiency, particularly in large-scale planning problems, a decomposition-based solution approach is adopted using CCG method. This technique has proven effective in solving robust optimization problems, especially two-stage DRO models (Zeng & Zhao, 2013). Recent studies demonstrated that CCG algorithm often outperforms traditional methods, such as Benders decomposition, in terms of convergence properties and efficiency (Zujian Wang & Qi, 2020).

In CCG method for solving DRO models, the main problem is decomposed into a Relaxed Master Problem (RMP) and one or more Sub problems (SubP). In RMP, the first-stage variables are optimized, and by solving the SubP, new constraints (rows) corresponding to pessimistic uncertainty scenarios, along with optimal values for the second-stage variables (columns), are generated and added to the RMP. It should be noted that in DRO models, the pessimistic probability distributions are also generated iteratively and incorporated as new constraints in the evaluation of the expected performance under the worst-case distribution.

The proposed DRO-BLP model can be compactly formulated as follows:

$$\begin{cases} \max_{x\in\mathcal{S}_{s}^{1}} \bar{\mathcal{R}} - c^{\top}x \\ s.t. \\ \max_{\rho,s,\zeta\in\mathcal{S}_{x}^{2}} \bar{\mathcal{R}} - P(\zeta;\theta_{\xi}) \\ \bar{\mathcal{R}} = \min_{\boldsymbol{\pi}\in\mathcal{P}_{\ell_1}} \left[\sum_{\xi\in\Xi} \pi_{\xi} \mathcal{R}(\rho, s; \theta_{\xi}) \right] \end{cases} \tag{30}$$

where $\mathcal{S}_{s}^{1}$ denotes the set of constraints (14)-(23) at the upper-level, which also depends on the decisions s obtained from the lower-level, and $\mathcal{S}_{x}^{2}$ represents the feasible region resulting from constraints (25)-(29) at the lower-level, which in turn depends on the flight schedule design decisions x from the upper level. Furthermore, the expressions $\bar{\mathcal{R}} - c^{\top}x$ and $\bar{\mathcal{R}} - P(\zeta;\theta_{\xi})$ are the objective function values at the upper and lower levels, respectively, where $P(\zeta;\theta_{\xi})$ captures the opportunity cost arising from insufficient flight capacity. Algorithm 1 shows the proposed decomposition-based solution method based on CCG, coordinating the interactions between the upper-level and lower-level decisions.

Algorithm 1. Proposed Column-and-Constraint Generation (CCG) Method

Input: DRO-BLP Model (30)
$\mathcal{C} = \{0\}, \boldsymbol{\pi^0}$, = 1, $\varphi = -\infty$, $s = 0$, $\bar{\mathcal{R}} = 0$, OPT=NO.

Output: Optimal decision (x^*, ρ^*, s^*) , Optimal profit φ^*.

- Define RMP:

RMP $:= \max_{x\in\mathcal{S}_{s}^{1}} \bar{\mathcal{R}} - c^{\top}x$

- Define SubP (Column/Variable s Generator)

SubP $:= \max_{\rho,s,\zeta\in\mathcal{S}_{x}^{2}} \bar{\mathcal{R}} - P(\zeta;\theta_{\xi})\ \ s.t.\ \bar{\mathcal{R}} \le \sum_{\xi\in\Xi} \bar{\pi}_{\xi}^{c} \mathcal{R}(\bar{\rho}, \bar{s}; \theta_{\xi})\ \forall \langle c \in \mathcal{C}\rangle$

- Define Row/Constraint Generator

$$\min_{\boldsymbol{\pi}\in\mathbb{R}_{+}^{\kappa}, \boldsymbol{\upsilon}\in\mathbb{R}_{+}^{\kappa}} \sum_{\xi\in\Xi} \pi_{\xi} \left[\sum_{l\in\mathcal{L}} \bar{\rho}_{l\xi} \sum_{m\in\mathcal{M}} \sum_{p\in\mathcal{P}} d_{mp\xi} \bar{s}_{lmp\xi} \right]$$
$$s.t.$$
$$\sum_{\xi\in\Xi} \upsilon_{\xi} \le \delta\ \forall$$
$$\pi_{\xi} - \pi_{\xi}^{0} \le \upsilon_{\xi}\ \forall \langle \xi \in \Xi \rangle$$
$$\pi_{\xi}^{0} - \pi_{\xi} \le \upsilon_{\xi}\ \forall \langle \xi \in \Xi \rangle$$
$$\sum_{\xi\in\Xi} \pi_{\xi} = 1$$

While OPT = NO **Do**

Solve RMP;
$\bar{x} \hookleftarrow$ argmax RMP ;

Solve SubP under fixed $\bar{x}$;
$(\bar{\rho}, \bar{s}, \bar{\zeta}) \hookleftarrow$ argmax SubP ;

Solve Constraint Generator under fixed $\bar{\rho}, \bar{s}$;
$\bar{\boldsymbol{\pi}}^{k} \hookleftarrow \operatorname{argmin}_{\boldsymbol{\pi}\in\mathbb{R}_{+}^{\kappa}, \boldsymbol{\upsilon}\in\mathbb{R}_{+}^{\kappa}} \sum_{\xi\in\Xi} \pi_{\xi} \left[\sum_{l\in\mathcal{L}} \bar{\rho}_{l\xi} \sum_{m\in\mathcal{M}} \sum_{p\in\mathcal{P}} d_{mp\xi} \bar{s}_{lmp\xi} \right]$;

If $\bar{\mathcal{R}} - c^{\top}\bar{x} > \varphi$ **Then**

$\varphi \hookleftarrow \bar{\mathcal{R}} - c^{\top}\bar{x}$;
$\mathcal{C} \hookleftarrow \mathcal{C} \cup \{k\}$;
$k \hookleftarrow k + 1$;

Else OPT = YES;

End

End

$x^* \hookleftarrow \bar{x}$; $(\rho^*, s^*) \hookleftarrow (\bar{\rho}, \bar{s})$; $\varphi^* \hookleftarrow \varphi$;
Return $(x^*, \rho^*, s^*, \varphi^*)$.

Computational Result and Analysis

In this section, to investigate the applicability of the proposed DRO-BLP model for the IFSTP problem and to evaluate the performance of the interactive and decomposition-based solution approach using CCG, we implement the model on a case study of Iranian airlines as well as several randomly generated experimental instances.

Experimental Setup

Air tour airline (iranairtour.ir) is one of the Iranian carriers whose flight network is illustrated in Figure 3. The airline's fleet consists of 4 Airbus A300-600 aircraft with a seating capacity of 224–273 passengers, 4 Airbus A310 aircraft with a capacity of 245 passengers, 2 Airbus A320 aircraft with a capacity of 174–180 passengers, and 6 McDonnell Douglas MD-82 aircraft with a capacity of 158–163 passengers. In addition to this real case, 10 random test instances is generated, whose scales are specified in Table 2. For all experimental instances, the earliest and latest departure times are set to 4:00 AM and 12:00 AM (Midnight), respectively. Given that the day is divided into 30-minute time slots, the total number of periods is $|\mathcal{T}| = 40$. For each of the six dayparts, early morning (4 AM – 8 AM), morning (8 AM – 11 AM), noon (11 AM – 2 PM), afternoon (2 PM – 6 PM), evening (6 PM – 9 PM), and night (9 PM – 12 AM), the maximum flight frequency (MF_{mp}) per O-D market is set between 2 and 4 flights. The slot access constraints (SA_{ap}) and gate access constraints (GA_{ap}) are set within the range of 5 to 10.

Considering that all flights are domestic, the maximum flight duration (Δ_l) is assumed to be 2 hours, equivalent to four time intervals of 30 minutes. The operational cost (OC_{lf}) and setup cost (SC_m) for each flight are randomly assigned within the ranges of 20–30 USD and 4–6 USD, respectively. The ticket price for each flight (ρ_l) is also randomly assigned within the range of 30–50 USD. Finally, the sensitivity-to-utility parameter (ϑ_m) is set to 0.5 for the case study, while for the experimental instances it is randomly generated within the range [0, 1].

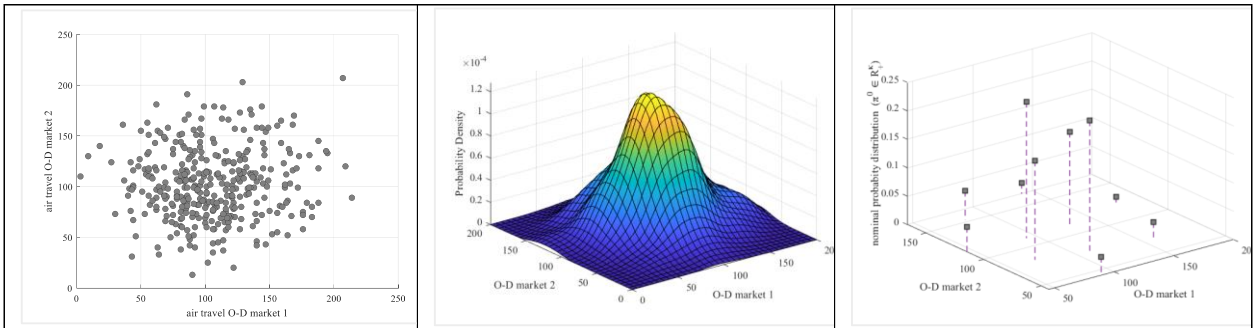
Table 2. The scale of test instances

Instances	Scales				
	$\|\mathcal{A}\|$	$\|\mathcal{M}\|$	$\|\mathcal{L}\|$	$\|\mathcal{F}\|)$	$(\sum_{f\in\mathcal{F}} NA_f)$
1	10	30	1200	3	20
2	12	35	1400	3	25
3	15	40	1600	4	30
4	17	40	1600	4	35
5	20	50	2000	4	40
6	25	60	2400	5	45
7	30	70	2800	5	50
8	35	80	3200	6	60
9	40	90	3600	7	70
10	50	100	4000	8	80

Figure 3. Domestic route network of Air tour airline (iranairtour.ir)

To construct DAS required by the DRO model, a dataset of 365 observations is randomly generated for each experimental instance. In the first phase of the proposed data-driven DRO approach, ML algorithm previously described is applied to process these data and produce the initial nominal distribution via KDE and uncertainty clustering using FCM. Figure 4 illustrates a simple 2-D example of this process with κ = 10. Finally, the maximum risk level is set to $\varepsilon = 0.05$ when constructing DAS, ensuring greater confidence that the true probability distribution is covered.

To execute the experimental instances, a personal computer equipped with a Windows operating system, an Intel(R) Core (TM) i7 CPU @ 2.50 GHz, and 16 GB of RAM is utilized. The optimization approach is implemented using the academic version of the Gurobi solver in Python (GurobiPy).



(a). Dataset (b). KDE to Estimate Nominal Distribution (c). Reference Distribution by KDE-FCM

Figure 4. Constructing a reference distribution using the proposed KDE-FCM method for DRO model

Evaluating the Robustness of Proposed DRO model

To evaluate the performance of the proposed data-driven DRO approach in the presence of uncertainty, we compare it with two alternative methods: (i) the classical SP approach that relies solely on the nominal/empirical probability distribution of uncertain parameters (SP-N), and (ii) the Expected Value Replacement (EVR) approach, which replaces the uncertain parameters with their average expected values.

To this end, we simulate the realization of uncertain parameters in 15 independent scenarios; 10 of which are in-sample and the remaining 5 are out-of-sample. In each realization, the profit is computed using each of DRO, SP-N, and EVR approaches. The value of uncertainty modeling is assessed based on the following evaluation metrics:

$$\begin{cases} \mathrm{VSP} = \dfrac{1}{15}\sum_{i=1}^{15} \mathrm{VSP}_i & ;\ \mathrm{VSP}_i = \dfrac{\mathcal{Z}_i^{\mathrm{SP-N}} - \mathcal{Z}_i^{\mathrm{EVR}}}{\mathcal{Z}_i^{\mathrm{EVR}}} \\ \left[\begin{array}{l} \mathrm{VDR}^{\mathrm{avg}} = \dfrac{1}{15}\sum_{i=1}^{15} \mathrm{VDR}_i^{\mathrm{avg}} \\ \mathrm{VDR}^{\mathrm{max}} = \max\limits_{i=1:15} \mathrm{VDR}_i \end{array}\right. & ;\ \mathrm{VDR}_i = \dfrac{\mathcal{Z}_i^{\mathrm{DRO}} - \mathcal{Z}_i^{\mathrm{SP-N}}}{\mathcal{Z}_i^{\mathrm{SP-N}}} \\ \mathrm{VPI} = \dfrac{1}{15}\sum_{i=1}^{15} \mathrm{VPI}_i & ;\ \mathrm{VPI}_i = \dfrac{\mathcal{Z}_i^{*} - \mathcal{Z}_i^{\mathrm{DRO}}}{\mathcal{Z}_i^{\mathrm{DRO}}} \end{cases} \tag{31}$$

where $\mathcal{Z}_i^{\mathrm{EVR}}$, $\mathcal{Z}_i^{\mathrm{SP-N}}$ and $\mathcal{Z}_i^{\mathrm{DRO}}$ denote the profits obtained from the EVR, SP-N, and DRO approaches, respectively, in each simulation of uncertain parameters, and $\mathcal{Z}_i^{*}$ represents the maximum profit achievable under perfect information.

The Value of Stochastic Programming (VSP) quantifies the average relative improvement in profit obtained by adopting the SP-N approach over the EVR method. Next, the Distributional Robustness Values ($\mathrm{VDR}^{\mathrm{avg}}$ and $\mathrm{VDR}^{\mathrm{max}}$) capture the relative performance gain achieved by introducing distributional robustness through the proposed DRO approach as compared to the classical SP-N method. These indicators compute, respectively, the average and the maximum achievable relative profit improvements due to DRO. Finally, the Value of Perfect Information (VPI) compares the average profit obtained from the DRO solution with the profit under perfect information. This metric indicates to what extent acquiring more information/data about the uncertain parameters or their tendency toward being deterministic can be beneficial.

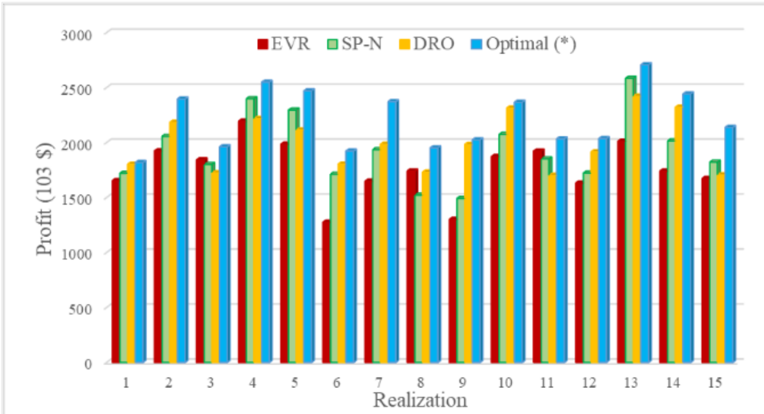


Figure 5. Comparing the profit after 15 realization of uncertain parameters

As illustrated in Figure 5, SP-N approach generally outperforms EVR, and although DRO does not consistently improve on SP-N, it typically yields higher profits with lower risk. In Table 3, the positive values of the VSS and VDR metrics demonstrate the value of accounting for uncertainty and modeling it through the proposed DRO approach. It is also observed that the DRO approach generally performs better under pessimistic conditions compared to the classical SP approach, while their average performances are relatively close. Furthermore, the positive value of VPI indicates that, despite the use of the data-driven DRO approach, increasing the amount of available data and reducing the parameters' uncertainty can enhance the average profit by 12.36%. The superior performance of DRO approach is also validated across 10 additional test instances, as shown in Figure 6, where its higher robustness is confirmed in terms of higher average profit and lower profit variance.

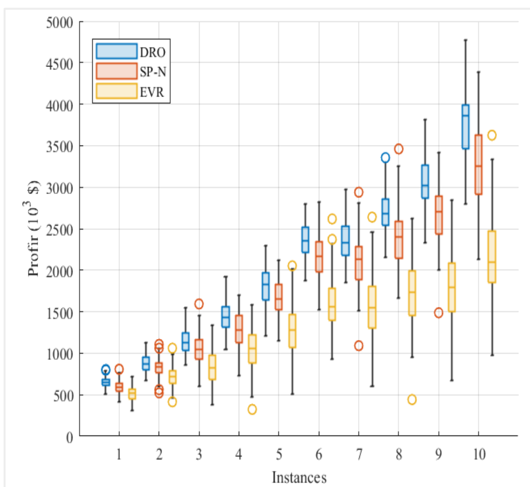


Figure 6. Robustness performance of uncertainty programming paradigms

Table 3. Evaluation criteria to determine the value of robustness in uncertainty

Criteria			
VSP	VDR^{avg}	VDR^{max}	VPI
10.20	2.86	15.22	12.36

Finally, a sensitivity analysis of the risk level parameter ϵ on the average profit of the proposed DRO approach has been conducted. As seen in Figure 7, by varying the risk level within the range of 14% to 16%, more than a 25% increase in the average profit can be achieved under the DRO approach, which can be considered an optimal calibration of the risk level in constructing the ambiguity set for the DRO model.

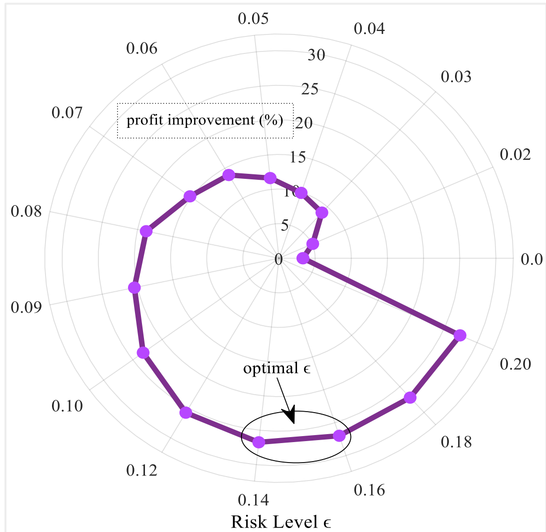


Figure 7. Optimal risk level in constructing ambiguity set for DRO model

The Benefits from Interaction Between Flight Scheduling and Pricing Decisions

This subsection examines the benefits reached by explicitly modeling the interaction between tactical flight scheduling and operational ticket pricing decisions, particularly as influenced by passengers' sensitivity to utility (ϑ_m). To this end, the proposed model, which integrates AMNL and BLP components, is compared with non-interactive models that ignore passengers' choice behavior.

Table 4 summarizes the results across all test instances and different levels of ϑ_m. As demonstrated, the proposed interactive model consistently generates higher profits than non-interactive models, particularly when passengers show greater sensitivity to flight utility. This improvement is achieved through the interaction of pricing and scheduling decisions and the incorporation of passenger choice behavior. These results confirm the value of the proposed approach in facilitating decision interaction and embedding choice behavior within the model.

Table 4. Comparison of the profit obtained by proposed interactive model relative to non-interactive model

Instance	Profit obtained by iterative and non-initiative models								
	($\vartheta_m < 0.35$)			($0.35 \le \vartheta_m \le 0.7$)			($\vartheta_m > 0.7$)		
	Non-interactive profit	Interactive profit (proposed mode)	Improve (%)	Non-interactive profit	Interactive profit (proposed mode)	Improve (%)	Non-interactive profit	Interactive profit (proposed mode)	Improve (%)
1	713.17	744.62	4.41	681.52	751.24	10.23	662.70	720.94	8.79
2	895.88	923.42	3.07	895.18	976.49	9.08	882.94	920.41	4.24
3	1284.20	1394.87	8.62	1205.44	1428.87	18.54	1154.65	1262.48	9.34
4	1731.88	1844.87	6.52	1560.96	2006.36	28.53	1498.77	1761.28	17.52
5	1804.55	2059.10	14.11	1702.84	2252.10	32.26	1619.53	1986.52	22.66
6	1864.98	2217.91	18.92	1819.25	2373.37	30.46	1808.84	2160.96	19.47
7	2003.16	2193.43	9.50	1836.64	2322.14	26.43	1768.60	2104.17	18.97
8	2155.75	2384.51	10.61	2120.92	2523.97	19.00	2035.18	2361.60	16.04
9	2782.24	3135.04	12.68	2579.37	3306.75	28.20	2558.51	3111.96	21.63
10	3734.13	4383.29	17.38	3466.62	4580.57	32.13	3316.54	4221.34	27.28

Computational Efficiency of Proposed CCG Method

As previously mentioned, the proposed DRO-BLP model can be solved by two methods. In the classical method, a single-level model can be obtained through MPEC approach, which can be solved using linear programming solvers such as Gurobi. In this section, the computational efficiency of the proposed interactive and decomposition-based approach using CCG is evaluated and compared with the traditional MPEC method.

Both methods are applied to the test instances, with a time limit of 300 minutes for each run. Table 5 shows the resulting objective values and computational times. For small-sized instances, both approaches yield similar results, although CCG method exhibits lower runtime. As problem size increases, the classical MPEC method not only requires significantly more time, but also the quality of its objective value considerably deteriorates in comparison with the proposed CCG-based method.

Table 5. Comparing the performance of proposed CCG relative to the classical MPEC for solving DRO-BLP model

Instance	MPEC-GurobiPy		Proposed CCG		Profit improvement (%)
	Profit (10^3)	Run-time (min)	Profit (10^3)	Run-time (min)	
1	744.62	19.41	744.62	5.12	0.00
2	974.41	47.39	953.42	8.31	-2.15
3	1402.44	73.51	1394.87	19.41	-0.54
4	1844.87	149.41	1844.87	38.94	0.00
5	2141.19	235.41	2059.10	53.48	-3.83
6	2017.94	300	2217.91	79.57	9.91
7	1941.34	300	2193.43	92.64	12.99
8	2089.04	300	2384.51	107.28	14.14
9	2351.41	300	3135.04	114.74	33.33
10	3052.81	300	4383.29	134.86	43.58

Conclusion

Flight scheduling and fleet assignment represent critical tactical decisions in airline planning. When integrated with operational decisions such as ticket pricing, they play a pivotal role in determining overall airline revenue and market performance. This study investigates the interaction of flight scheduling and ticket pricing (IFSTP) problem, with the objective of maximizing airline profitability under uncertain demand and passenger choice behavior.

To model passenger choice behavior, we developed an augmented multinomial logit (AMNL) model and applied a bi-level programming (BLP) framework to coordinate scheduling and pricing decisions. To address demand uncertainty, we proposed a data-driven distributionally robust optimization (DRO) approach, where the distributional ambiguity set (DAS) was constructed using an ℓ_1-norm-based structure. This set was derived from historical demand data using a machine learning algorithm that integrates kernel density estimation (KDE) and fuzzy C-Means clustering (FCM).

The numerical results demonstrate the critical role of demand uncertainty in the IFSTP model. Ignoring uncertainty not only increases profit volatility but also reduces average profitability significantly. A comparison between the proposed data-driven DRO model and classical approaches shows that the proposed model outperforms traditional stochastic programming methods in terms of profit robustness, particularly in out-of-sample scenarios. Furthermore, assessing the interaction between scheduling and pricing decisions reveals that the proposed model, which accounts for passenger choice behavior, consistently yields higher profit than non-integrated models, especially in markets with greater sensitivity to flight utility. The evaluation of the proposed CCG solution method confirms its computational efficiency, particularly for solving large-scale problem instances.

From a theoretical perspective, this research contributes a novel data-driven DRO framework and an interactive decomposition-based solution methodology for solving bi-level programming problems, with potential applications in other competitive decision-making contexts within transportation systems. In addition, the proposed AMNL model captures passenger choice behavior and sensitivity to flight attributes. From a practical perspective, the proposed DRO-BLP model serves as a robust and modern decision-support tool for airlines, enabling the integration of tactical flight scheduling and operational pricing decisions while directly incorporating parameter uncertainty to derive robust solutions. The CCG algorithm further ensures the computational tractability of large-scale instances.

For future research, one promising direction is to incorporate strategic-level fleet planning decisions and examine their impact on schedule design. This would allow airlines to increase market share through optimal fleet investments and flight schedule adjustments. Additionally, considering weather-induced disruptions in the flight schedule and developing real-time rescheduling strategies could further enhance the practical applicability and attractiveness of the proposed model for airline companies.

References

Barnhart, C., & Cohn, A. (2004). Airline schedule planning: Accomplishments and opportunities. *Manufacturing & service operations management*, *6*(1), 3-22.

Barnhart, C., Farahat, A., & Lohatepanont, M. (2009). Airline fleet assignment with enhanced revenue modeling. *Operations Research*, *57*(1), 231-244.

Ben-Tal, A., Den Hertog, D., De Waegenaere, A., Melenberg, B., & Rennen, G. (2013). Robust solutions of optimization problems affected by uncertain probabilities. *Management Science, 59*(2), 341-357.

Ben Ahmed, M., Hryhoryeva, M., Hvattum, L. M., & Haouari, M. (2022). A matheuristic for the robust integrated airline fleet assignment, aircraft routing, and crew pairing problem. *Computers & Operations Research, 137*, 105551. doi:https://doi.org/10.1016/j.cor.2021.105551

Bertsimas, D., Gupta, V., & Kallus, N. (2018a). Data-driven robust optimization. *Mathematical Programming*, *167*, 235-292.

Bertsimas, D., Gupta, V., & Kallus, N. (2018b). Robust sample average approximation. *Mathematical Programming*, *171*, 217-282.

Bertsimas, D., & Thiele, A. (2006). Robust and data-driven optimization: modern decision making under uncertainty. In *Models, methods, and applications for innovative decision making* (pp. 95-122): INFORMS.

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & geosciences, 10*(2-3), 191-203.

Birolini, S., Jacquillat, A., Cattaneo, M., & Antunes, A. P. (2021). Airline Network Planning: Mixed-integer non-convex optimization with demand–supply interactions. *Transportation Research Part B: Methodological, 154*, 100-124.

Cadarso, L., & Vaze, V. (2023). Passenger-centric integrated airline schedule and aircraft recovery. *Transportation Science*, *57*(3), 813-837.

Calvete, H. I., Gale, C., & Mateo, P. M. (2008). A new approach for solving linear bilevel problems using genetic algorithms. *European journal of operational research*, *188*(1), 14-28.

Cao, Y., Kleywegt, A. J., & Wang, H. (2022). Network revenue management under a spiked multinomial logit choice model. *Operations Research, 70*(4), 2237-2253.

Chen, N., & Hu, M. (2023). Frontiers in Service Science: Data-Driven Revenue Management: The Interplay of Data, Model, and Decisions. *Service Science, 15*(2), 79-91.

Costa, G., & Kwon, R. H. (2022). Data-driven distributionally robust risk parity portfolio optimization. *Optimization Methods and Software, 37*(5), 1876-1911.

Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, *58*(3), 595-612.

Eltoukhy, A. E. E., Chan, F. T. S., & Chung, S. H. (2017). Airline schedule planning: a review and future directions. *Industrial Management & Data Systems*, *117*(6), 1201-1243.

Gao, R., Chen, X., & Kleywegt, A. J. (2024). Wasserstein distributionally robust optimization and variation regularization. *Operations Research, 72*(3), 1177-1191.

Gilani, H., & Sahebi, H. (2022). A data-driven robust optimization model by cutting hyperplanes on vaccine access uncertainty in COVID-19 vaccine supply chain. *Omega, 110*, 102637.

Gong, H., & Zhang, Z.-H. (2022). Benders decomposition for the distributionally robust optimization of pricing and reverse logistics network design in remanufacturing systems. *European journal of operational research, 297*(2), 496-510.

Gorissen, B. L., Yanıkoğlu, İ., & Den Hertog, D. (2015). A practical guide to robust optimization. *Omega, 53*, 124-137.

Grosche, T., & Grosche, T. (2009). Integrated airline scheduling. *Computational Intelligence in Integrated Airline Scheduling*, 59-171.

Han, B., Shang, C., & Huang, D. J. E. J. o. O. R. (2021). Multiple kernel learning-aided robust optimization: Learning algorithm, computational tractability, and usage in multi-stage decision-making. *292*(3), 1004-1018.

Hoeffding, W. (1994). Probability inequalities for sums of bounded random variables. *The collected works of Wassily Hoeffding*, 409-426.

Huang, R., Qu, S., Gong, Z., Goh, M., & Ji, Y. (2020). Data-driven two-stage distributionally robust optimization with risk aversion. *Applied Soft Computing, 87*, 105978.

IATA. (2019). Economic performance of the airline industry. *https://www.iata.org/*.

Jiang, R., & Guan, Y. (2018). Risk-averse two-stage stochastic program with distributional ambiguity. *Operations Research, 66*(5), 1390-1405.

Kalashnikov, V. V., Dempe, S., Pérez-Valdés, G. A., Kalashnykova, N. I., & Camacho-Vallejo, J.-F. (2015). Bilevel programming and applications. *Mathematical Problems in Engineering, 2015*(1), 310301.

Kenan, N., Jebali, A., & Diabat, A. (2018). An integrated flight scheduling and fleet assignment problem under uncertainty. *Computers & Operations Research*, *100*, 333-342.

Kiarashrad, M., Pasandideh, S. H. R., & Mohammadi, M. (2021). A mixed-integer nonlinear optimization model for integrated flight scheduling, fleet assignment, and ticket pricing in competitive market. *Journal of Revenue and Pricing Management*, 1-12.

Li, J., Liu, Y., & Yang, G. (2024). Two-stage distributionally robust optimization model for a pharmaceutical cold supply chain network design problem. *International Transactions in Operational Research, 31*(5), 3459-3493.

Liu, L., Song, S., Wang, Z., & Zhang, Y. (2022). Data-driven distributionally robust optimization for railway timetabling problem. *IEEE Transactions on Automation Science and Engineering, 21*(1), 810-826.

Liu, W., Yang, L., & Yu, B. (2023). Distributionally Robust Optimization Based on Kernel Density Estimation and Mean-Entropic Value-at-Risk. *INFORMS Journal on Optimization*, *5*(1), 68-91.

Lohatepanont, M., & Barnhart, C. (2004). Airline schedule planning: Integrated models and algorithms for schedule design and fleet assignment. *Transportation Science*, *38*(1), 19-32.

Lu, X., & Zhou, K. (2024). A distributionally robust optimization approach for optimal load dispatch of energy hub considering multiple energy storage units and demand response programs. *Journal of Energy Storage, 78*, 110085.

Metzker, P., Thevenin, S., Adulyasak, Y., & Dolgui, A. (2025). Distributionally robust optimization for the multi-period multi-item lot-sizing problems under yield uncertainty. *IEEE Transactions on Automation Science and Engineering*.

Mohajerin Esfahani, P., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming, 171*(1), 115-166.

Mohseni, S., & Pishvaee, M. S. (2023). Energy trading and scheduling in networked microgrids using fuzzy bargaining game theory and distributionally robust optimization. *Applied Energy*, *350*, 121748.

Özarık, S. S., da Costa, P., & Florio, A. M. (2024). Machine learning for data-driven last-mile delivery optimization. *Transportation Science, 58*(1), 27-44.

Papadimitriou, D. (2016). *New challenges in network optimization*. Paper presented at the 2016 IEEE 17th International Conference on High Performance Switching and Routing (HPSR).

Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics, 33*(3), 1065-1076.

Peng, C., & Delage, E. (2024). Data-driven optimization with distributionally robust second order stochastic dominance constraints. *Operations research*, *72*(3), 1298-1316.

Qiu, R., Sun, Y., Zhou, H., & Sun, M. (2023). Dynamic pricing and quick response of a retailer in the presence of strategic consumers: A distributionally robust optimization approach. *European journal of operational research, 307*(3), 1270-1298.

Rahimian, H., & Mehrotra, S. (2019). Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*.

Razavi Al-e-hashem, S. A., Papi, A., Pishvaee, M. S., & Rasouli, M. (2022). Robust maintenance planning and scheduling for multi-factory production networks considering disruption cost: a bi-objective optimization model and a metaheuristic solution method. *Operational Research*, *22*(5), 4999-5034.

Sa, C. A. A., Santos, B. F., & Clarke, J.-P. B. (2020). Portfolio-based airline fleet planning under stochastic demand. *Omega*, *97*, 102101.

Shang, C., Huang, X., & You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers & Chemical Engineering*, *106*, 464-479.

Sherali, H. D., Bae, K.-H., & Haouari, M. (2013). An integrated approach for airline flight selection and timing, fleet assignment, and aircraft routing. *Transportation Science, 47*(4), 455-476.

Sherali, H. D., & Zhu, X. (2008). Two-stage fleet assignment model considering stochastic passenger demands. *Operations Research, 56*(2), 383-399.

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*: Routledge.

Sun, H., & Xu, H. (2016). Convergence analysis for distributionally robust optimization and equilibrium problems. *Mathematics of Operations Research, 41*(2), 377-401.

Van Parys, B. P. G., Esfahani, P. M., & Kuhn, D. (2021). From data to decisions: Distributionally robust optimization is optimal. *Management Science, 67*(6), 3387-3402.

Wandelt, S., Signori, A., Chang, S., Wang, S., Du, Z., & Sun, X. (2025). Unleashing the potential of operations research in air transport: A review of applications, methods, and challenges. *Journal of Air Transport Management, 124*, 102747.

Wang, C., Liu, K., Zhang, C., & Miao, L. (2024). Distributionally robust chance-constrained optimization for the integrated berth allocation and quay crane assignment problem. *Transportation Research Part B: Methodological*, *182*, 102923.

Wang, Y., Zhang, Y., & Tang, J. (2024). Wasserstein distributionally robust surgery scheduling with elective and emergency patients. *European journal of operational research, 314*(2), 509-522.

Wang, Z., Jiang, S., Lin, Y., Tao, Y., & Chen, C. (2024). A review of algorithms for distributionally robust optimization using statistical distances. *Journal of Industrial and Management Optimization*, 0-0.

Wang, Z., Liu, Z., & Huo, Y. (2025). A distributionally robust optimization approach of multi-park integrated energy systems considering shared energy storage and Uncertainty of Demand Response. *Electric Power Systems Research*, *238*, 111116.

Wang, Z., & Qi, M. J. T. S. (2020). Robust service network design under demand uncertainty. *54*(3), 676-689.

Węglarczyk, S. (2018). *Kernel density estimation and its application.* Paper presented at the ITM web of conferences.

Wei, K., Vaze, V., & Jacquillat, A. (2020). Airline timetable development and fleet assignment incorporating passenger choice. *Transportation Science*, *54*(1), 139-163.

Weide, O., Ryan, D., & Ehrgott, M. (2010). An iterative approach to robust and integrated aircraft routing and crew scheduling. *Computers & Operations Research, 37*(5), 833-844.

Xu, Y., Adler, N., Wandelt, S., & Sun, X. (2024). Competitive integrated airline schedule design and fleet assignment. *European journal of operational research, 314*(1), 32-50.

Yan, C., Barnhart, C., & Vaze, V. (2022). Choice-based airline schedule design and fleet assignment: A decomposition approach. *Transportation Science*, *56*(6), 1410-1431.

Yan, R., Liu, Y., & Wang, S. (2024). A data-driven optimization approach to improving maritime transport efficiency. *Transportation Research Part B: Methodological, 180*, 102887.

Yan, S., Tang, C.-H., & Lee, M.-C. (2007). A flight scheduling model for Taiwan airlines under market competitions. *Omega*, *35*(1), 61-74.

Yang, B., Xu, X., Gong, Y., & Rekik, Y. (2024). Data-driven optimization models for inventory and financing decisions in online retailing platforms. *Annals of Operations Research, 339*(1), 741-764.

Zeng, B., & Zhao, L. (2013). Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters, 41*(5), 457-461.

Zhang, S., Jia, R., He, D., & Chu, F. (2022). Data-driven robust optimization based on principle component analysis and cutting plane methods. *Industrial & Engineering Chemistry Research*, *61*(5), 2167-2182.

Zhou, L., Liang, Z., Chou, C.-A., & Chaovalitwongse, W. A. (2020). Airline planning and scheduling: Models and solution methodologies. *Frontiers of Engineering Management*, *7*(1), 1-26.