sustainability

Article

Green Airline-Fleet Assignment with Uncertain Passenger Demand and Fuel Price

Ming Liu [1,2], Yueyu Ding [1], Lihua Sun [2,*], Runchun Zhang [2], Yue Dong [2], Zihan Zhao [2], Yiting Wang [2] and Chaoran Liu [2]

[1] Urban Mobility Institute, Tongji University, Shanghai 201804, China
[2] School of Economics & Management, Tongji University, Shanghai 200092, China
* Correspondence: sunlihua@tongji.edu.cn

Abstract: Although air transport contributes to globalization, airline emissions have attracted focus in green logistics. In this work, we investigate the airline-fleet assignment problem from a risk-averse perspective in which uncertain demand and fuel price are considered simultaneously. The objective is to maximise the total profit in a risk-averse fashion, i.e., the weighted sum of the expected profit and the conditional value at risk of profit. An appropriate assignment can reduce fuel use and carbon dioxide emissions. For the problem, a two-stage stochastic programming model is constructed. The first stage consists of assigning aircraft families to flight legs, while the second stage determines specific aircraft deployment with the realized information. To solve the problem, a sample average approximation (SAA) approach is firstly applied. An efficient string-based heuristic is, further, developed. Numerical experiments are conducted and sensitivity analysis is performed. The results show the efficiency of the proposed heuristic and managerial insights are drawn.

Keywords: fleet-assignment problem; green logistics; uncertain passenger demand; uncertain fuel price; risk averse; string-based heuristic

Citation: Liu, M.; Ding, Y.; Sun, L.; Zhang, R.; Dong, Y.; Zhao, Z.; Wang, Y.; Liu, C. Green Airline-Fleet Assignment with Uncertain Passenger Demand and Fuel Price. *Sustainability* **2023**, *15*, 899. https://doi.org/10.3390/su15020899

Academic Editor: António Abreu

Received: 20 November 2022
Revised: 14 December 2022
Accepted: 15 December 2022
Published: 4 January 2023

1. Introduction

Civil aviation is a typical industry with heavy assets, high cost and low profits. Airline management consists of a variety of decision-making steps, such as fleet assignment, crew scheduling, aircraft routing, and pricing management, which affect the airline profitability [1–3]. Recently, governments are advocating for the green and sustainable development of the aviation industry [4,5]. Many airlines have recognized the importance of reducing emissions from a fuel-consumption perspective [6]. In addition, allocating the proper type of aircraft to flights can maintain a competitive edge for airlines [7]. Therefore, the research on airline fleet assignment is necessary, for it takes passenger demand, aircraft seat capacity, and fuel price into account.

There are various uncertainties, such as passenger demand, flight delay, non-cruise time and fuel price, which impact the profit of airline companies. Therefore, appropriate airline management under uncertainty is a great challenge. Airline fleet management is an important issue of airline management, especially under various uncertain factors.

Generally, uncertain demand is a commonly considered random factor in previous studies [8–10], because the assignment decisions of an aircraft family need to be made 10–12 weeks before the departure of flights based on early demand forecasting. Whereas the actual demand of passengers will not be finalized until the last flight departure day, at which point it is hard for decision makers to assign the right types of aircraft to flight legs [11]. In practice, decision makers need to adjust their initial decisions on the fleet assignment by swapping aircraft to appropriate flight legs when more precise demand information is realized [12]. In order to deal with this issue, two-stage stochastic programming models addressing airline fleet assignment with uncertain demand are attracting increasing attention [8,9,13]. In the first stage, aircraft families are assigned to flight legs. In the second stage, a specific aircraft type of the family determined in the first stage is assigned to each flight leg after random demand information is realized. However, these works have not considered fuel-price fluctuation.

In practice, jet fuel cost accounts for about 30% of operational costs [11]; even small changes in its price may cause a significant impact on airline revenue. Ryerson and Hansen [14] evaluate different types of aircraft and find that the fuel price and passenger costs have a significant influence on fleet sizing decisions. Generally, an efficient fleet assignment by deploying different types and capacities of aircraft taking fuel-efficiency considerations into account can save total costs. However, to the best of our knowledge, only the work outlined by Naumann, Suhl, and Friedemann [15] incorporates passenger demand and fuel-price uncertainty together into fleet-assignment management. The authors present a two-stage stochastic programming model where a worst-cases-oriented measure (i.e., CVaR) is also integrated. However, the average-case (or expected) value is not considered, which is a commonly studied objective function in the literature. When solving the problem, a commercial solver is applied in given small-scale scenarios but the commercial solver may lose its power when dealing with large number of scenarios.

Passenger demand and fuel price are susceptible to low frequency and high disruption events. Hsu and Eie [16] agree that fluctuations in passenger demand and fuel price are correlated with the external environment, and these two uncertainties will affect the reliability of the proposed network planning (i.e., types of aircraft and flight frequencies on airline routes). Airline managers may hope to evaluate the impact of uncertain fluctuations from a conservative (or risk-averse) perspective and make decisions accordingly. However, taking risk measures in uncertain situations into consideration is poorly researched in the stochastic fleet-assignment optimization models, and most are typically based on a risk-neutral perspective, i.e., optimizing the expected value.

Motivated by the above observations, in this work, we investigate a new risk-averse fleet-assignment problem with uncertain passenger demand and fuel price. For the studied problem, a two-stage stochastic programming model is proposed. A sample average approximation (SAA) approach is applied, and an efficient string-based heuristic is, further, developed to solve it. The contribution of this work mainly includes:

1. A new airline fleet-assignment problem is addressed in which uncertain passenger demand and fuel-price fluctuation are considered simultaneously.
2. A risk-averse two-stage stochastic programming model is established for the fleet-assignment problem.
3. An efficient string-based heuristic is, further, developed to solve the problem.

The remainder of this paper is organised as follows. In Section 2, a brief literature review is provided. Section 3 describes the problem and presents a two-stage stochastic programming model. In Section 4, the SAA method is applied, and a string-based heuristic is developed to solve the problem. Computational experiments are reported in Section 5. The conclusion is given in Section 6.

2. Literature Review

The deterministic airline fleet-assignment problem is well-studied in early works [17–21], where passenger demand is seen as a fixed parameter under a time–space network. In reality, the efficiency of fleet assignment is impacted by various uncertain factors such as flight departure time, passenger demand and fuel price, etc. Recently, researchers have investigated the airline fleet-assignment problem with uncertain passenger demand.

For a stochastic fleet-assignment problem, Sherali and Zhu [8] develop a two-stage stochastic programming model to maximise the total expected profit, in which several scenarios are used to capture demand uncertainty. The decision makers are allowed to assign aircraft families to flight legs in the first stage and determine an appropriate aircraft type for each leg based on the decisions of the first stage. Similarly, Pilla et al. [22] also establish a two-stage stochastic programming model for the fleet-assignment problem.

To improve the level of passenger-demand captured, they introduce the concept of crew compatible families (i.e., Boeing 757 and 767 are crew compatible) into their model to maximise the total revenue and statistical computer experiments are applied. However, they do not optimize the approximation function for their model. Based on [22], Pilla et al. [9] further propose a new solution method. They apply the multivariate adaptive regression splines cutting plane (MARS-CP) algorithm which uses the approximation function for their two-stage stochastic programming model, which is faster than the traditional Benders decomposition method in convergence. Liu et al. [10] present a risk-averse two-stage stochastic programming model for addressing the airline fleet assignment optimization problem, where the conditional value-at-risk (CVaR) and the expected value are combined to measure the risk of total revenue. Such a fashion of combining the expected value and CVaR is called a risk-averse perspective in [23]. However, they only consider demand uncertainty and no efficient heuristic is proposed.

To further increase profits, integrated fleet assignment and flight scheduling have gradually started to be considered. For the integrated problem, Sherali, Bae, and Haouari [24] consider itinerary-based demands, optional legs, and multiple fare classes in their two-stage stochastic programming model, to maximise net revenue under various demand scenarios. A Benders decomposition approach is developed to solve the model. Sherali, Bae, and Haouari [25] further extend the integrated problem by simultaneously taking passenger recapture issues, flexible flight times, and flight-schedule balance throughout the day into account, and a Benders decomposition solution approach is also applied to solve the problem. Cadarso and de Celis [26] develop a two-stage nonlinear stochastic programming model by simultaneously considering uncertain demand and stochastic operating conditions (i.e., flight-connection time). They propose an improved Benders decomposition algorithm to solve the airline schedule-planning problem on a large scale. A sample average approximation (SAA) algorithm is used in [27], to solve the two-stage stochastic programming model for integrated flight scheduling and the fleet-assignment problem. They capture the demand and fare uncertainty in the proposed model for the maximisation of the total expected profit.

The above studies all consider passenger-demand fluctuations, while fuel price is another crucial factor of uncertainty involved in operational costs. Naumann, Suhl, and Friedemann [15] present a two-stage stochastic program while simultaneously considering uncertain demand and fuel price for the airline fleet-assignment problem. The results indicate that a previously determined fleet assignment may be significantly changed due to the realized high fuel price. They use CVaR as the objective function, which is worst-case-oriented. However, they do not consider the mean risk measure (i.e., expected value) and their solution method by calling a commercial solver is not suitable for practical large-scale scenarios.

Furthermore, for the fleet-deployment problem in terms of liner shipping, uncertainties are also considered in existing works [28–30]. Specifically, the relevant works that use the two-stage stochastic programming and the SAA method to deal with the liner-shipping problem are reviewed in the following. A joint service-capacity planning and dynamic container-routing problem with uncertain demand is studied in [28]. The problem is formulated by a two-stage stochastic programming model, and SAA, progressive hedging algorithm (PHA) and adapted PHA are proposed to solve their problem. Considering the uncertain container demand, vessel fleet deployment, bunkering and the sailing-speed optimization problem is investigated in [29]. For the problem, the authors design a two-stage stochastic and non-linear programming model, and apply SAA and L-shaped methods. Wang et al. [30] consider a container-slot allocation problem with uncertain shipping demand. They propose a two-stage stochastic mixed-integer non-linear programming model and solve the problem using the SAA method based on Lagrangian relaxation and dual decomposition.

Summing up, two-stage stochastic programming is a popular approach to formulating the stochastic fleet-assignment problem, and SAA is employed to solve the model. However, no one has studied the risk-averse airline fleet-assignment problem with uncertain fuel price and passenger demand, combining the expected value and the CVaR into the objective function. Although CVaR is applied in various research fields, to the best of our knowledge, we are the first to present a risk-averse two-stage stochastic programming model and develop an efficient heuristic algorithm for this problem.

To better understand the current research status, the comparison of the most related works on fleet-assignment problem under uncertainty is reported in Table 1.

Table 1. Comparison of the most related research on stochastic airline fleet-assignment problem.

	Uncertainties		Objective		
Literature	**Passenger Demand**	**Fuel Price**	**Expected Value**	**CVaR**	**Solution Approach**
Sherali and Zhu [8]	✓		✓		Benders decomposition algorithm
Pilla et al. [22]	✓		✓		Statistical computer experiments approach
Pilla et al. [9]	✓		✓		MARS-CP algorithm
Liu et al. [10]	✓		✓	✓	SAA
Naumann, Suhl, and Friedemann [15]	✓	✓		✓	Commercial solver
This paper	✓	✓	✓	✓	**String-based heuristic**

3. Two-Stage Stochastic Fleet Assignment

In this section, we first state the problem under consideration and then a two-stage stochastic programming model is established.

3.1. Problem Description

Similar to [8], our problem is also based on itinerary demands, where an itinerary constitutes a passenger completing a journey from the departure airport to the destination airport using a sequence of consecutive flight legs [1]. For example, if a passenger flies from airport A to B and then to C, the itinerary comprises two flight legs (i.e., A–B and B–C). To address airline scheduling problems, a time–space network is proposed because of its advantage in modelling possible connections between arriving and departing flights [8,17]. The time–space network is also widely used in other research fields, including bus scheduling [31], bicycle sharing systems [32], etc. In our work, the time–space network is used to formulate the fleet-assignment problem within one day. Each event node in the network is associated with the combination of an aircraft type, an arrival or departure time, and an airport, representing a take-off or landing event. Since different aircraft types have different attributes, each aircraft type is associated with one sub-network layer (e.g., Figure 1a,b), and the time–space network is the combination of a series of sub-networks (e.g., Figure 1c). In each sub-network, there are three types of arcs to connect all the nodes, including ground arcs, flight arcs and wraparound arcs [8]. Taking Figure 1a as an example, the inclined arrows represent the flight arcs (or flight legs) of an aircraft of type 1; the vertical dotted-line arrows indicate that the aircraft stays at the airport during the time period, which are the ground arcs; the wraparound arcs connect the last landing event with the first take-off event of the day at each airport by dashed curved arrows. Moreover, in Figure 1a, flight arcs A_1 and C_1 indicate that aircraft type 1 arrives at Airport 1 at different times, flight arcs B_1 and D_1 mean aircraft type 1 departs from Airport 1, and flight arcs E_1 and F_1 represent the arrival and departure from airport 2, respectively. Similarly, flight arcs A_2 to F_2 represent the arrival and departure at different airports of a type-2 aircraft, respectively. The turn-around times of various aircraft types are different. Figure 1c represents the whole time–space network for two types of aircraft with two airports and 12 flight legs.

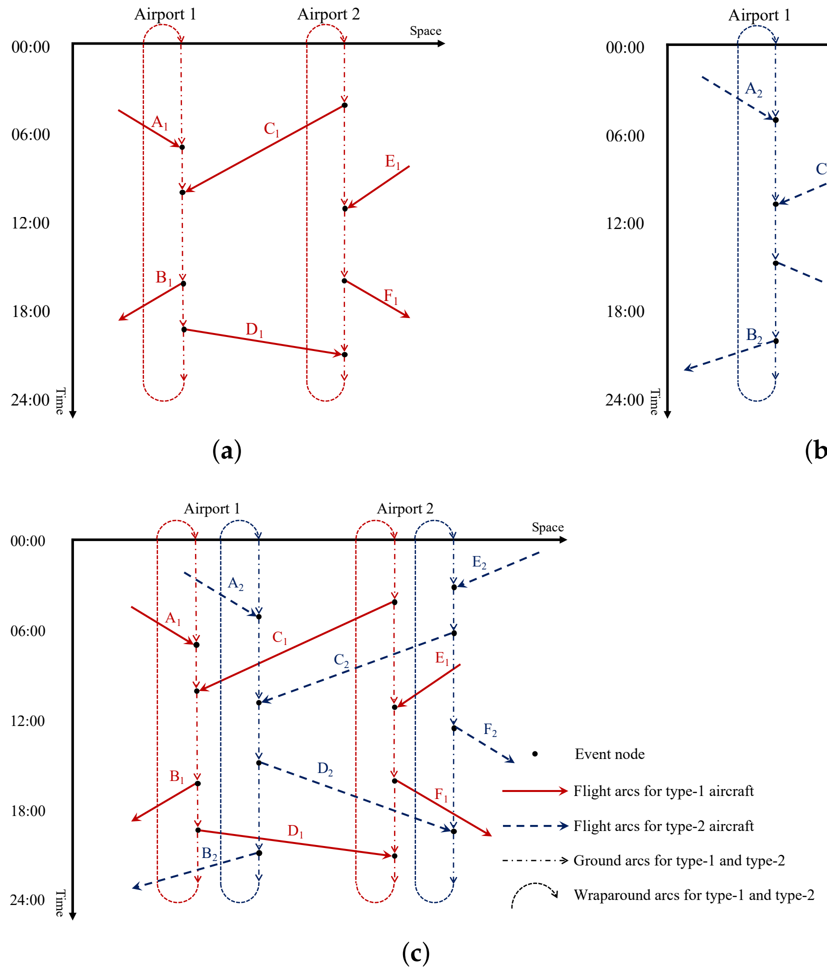

Figure 1. A time–space network (Adapted from [8]). (**a**) A time–space sub-network for type-1 aircraft. (**b**) A time–space sub-network for type-2 aircraft. (**c**) A time–space network for two types of aircraft with two airports and twelve flight legs.

Jet fuel cost occupies an important part of airline operational costs, and the fuel price is highly volatile. Apart from passenger-demand uncertainty, our considered problem also involves fuel-price uncertainty. We use a scenario $s \in S$ to portray the realization of random passenger demand and fuel price where S denotes the set of scenarios. Generally, the fuel efficiency can be measured by the fuel consumption per kilometer of each aircraft. For example, under the same conditions, the small aircraft B787-8 is more fuel-efficient than the large aircraft A330-300. Please see Table 2, where the fuel consumption of a B787-8 is 6.45 L per kilometer (km), while an A330-300 burns about 11.51 L per kilometer. When more fuel-efficient aircraft is assigned to flight legs, an airline transportation company can reduce its total fuel cost. However, there could be a risk of the additional cost of passenger spillover due to insufficient seating capacity [33]. Therefore, we hope to achieve a balance between fuel cost and the cost induced by spilled passengers by rationally arranging different aircraft types to flight legs. Similar to [8], the decision maker determines the assignment of an aircraft family to a scheduled flight leg 10–12 weeks before the start of flight operations. Then, the demand and fuel price are realized. After that, the decision maker can select a specific aircraft type conforming to the aircraft family previously determined (from a selfowned fleet of aircraft or leasing one at a large cost from other airline companies) for the flight leg about 4–6 weeks before its execution.

Table 2. Information of each aircraft type.

Aircraft Family K	**B787**		**A330**		**B737**	
Aircraft Type T	**8**	**9**	**200**	**300**	**Max8**	**700HH**
A_t	5	2	4	6	3	3
$cons_t$	6.45	7.37	11.29	11.51	3.59	3.40
Cap_t	213	289	260	303	176	120
CASK	0.41	0.38	0.39	0.37	0.43	0.50

Note that the information of Cap_t comes from [34], the information of $cons_t$ (fuel consumption of aircraft type t in litres per km) of B737-Max8 comes from [35], and the information of $cons_t$ of other aircraft comes from [36]. We divide the fuel consumption (kilogram/hour) by the average flying speed (kilometer/hour) and fuel density (0.775 kg/L) to obtain the fuel consumption (litre/kilometer) of each aircraft.

This paper investigates a stochastic airline fleet-assignment problem to maximise profit, where the passenger demand and fuel price are uncertain. We make assignment decisions in a risk-averse fashion, i.e., combining the average-case value and the worst-cases-oriented value (i.e., CVaR). For the addressed problem, in the first stage, a set K of aircraft families are to be assigned to a set L of flight legs. In the second stage, after the information of demand and fuel price is realized, a set of specific aircraft types (say, T_k) within a previously determined family (say, the kth family) is assigned to the flight leg. This work is based on the following assumptions:

1. The airline's flight-leg network is given, and the passenger demand and fuel price are uncertain 10–12 weeks before the departure of flights in the first-stage decision;
2. Airlines can no longer recapture the passengers they have lost;
3. Each flight leg l is covered by only one aircraft family k, and only one aircraft type t of the family k can be selected for flight leg l. For example, if flight leg "airport A to airport B" is assigned Airbus A330 family in the first stage, then in the second stage, we can only select one aircraft from set {A330-200, A330-300} to fly this flight leg;
4. Considering that there may exist the violation of the available number A_t of selfowned type-t aircraft in the stochastic environment, this violation can be addressed by leasing an aircraft from other airline companies. Such a violation will also cause a huge cost in the second-stage objective via $\psi_t \cdot P_t^s$, where ψ_t is the leasing cost of one aircraft with type t, and P_t^s records the number of aircraft with type t to be leased in scenario s.

3.2. The Stochastic Programming Model (Minimising Expected Value)

In this subsection, we first present a risk-neutral two-stage stochastic programming model which minimises the expected total profit. Then, we further develop a risk-averse formulation for the concerned problem, by modifying the original mean-risk objective function. The following notation is presented before describing our model.

Input Parameters

K: set of aircraft families in the fleet, indexed by k;
T_k: set of aircraft types in aircraft family k, indexed by t, where $k \in K$;
T: set of aircraft types in the fleet, where $T = \cup_{k \in K} T_k$;
N_t: set of event nodes in aircraft type $t's$ sub-network, indexed by n, where $t \in T$;
G_t: set of ground arcs in aircraft type $t's$ sub-network, indexed by g, where $t \in T$;
L: set of flight legs in the flight schedule, indexed by l;
L_k: set of flight legs assigned to aircraft family k, indexed by l, where $k \in K$;
I: set of all itineraries, indexed by i;
I_l: set of itineraries in I that contain flight leg l, indexed by i, where $l \in L$;
F_t: set of arcs traversing in time through a counting time line (i.e., flight legs l and ground arcs g) in aircraft type $t's$ sub-network, where $t \in T$;
c_{lt}: cost of assigning aircraft type t to flight leg l (without considering jet fuel costs), where $l \in L,\ t \in T$;
Cap_t: seating capacity of aircraft type t, where $t \in T$;

S: set of scenarios corresponding to the realizations of passenger demand and fuel price, indexed by s;
d_i^s: demand realization for itinerary i in scenario s, where $i \in I,\ s \in S$;
f_i^s: estimated average fare for itinerary i in scenario $s, i \in I,\ s \in S$;
p^s: probability for scenario s, where $s \in S$;
A_t: available number of aircraft with type t, where $t \in T$;
δ^s: jet fuel price in scenario s;
$\boldsymbol{\xi}^s$: vector of uncertain demands and fuel price in scenario $s \in S$, and $\boldsymbol{\xi}^s = [d_1^s, d_2^s, \ldots, d_{|I|}^s, \delta^s]$, where $s \in S$;
$cons_t$: fuel consumption of aircraft type t in litres per km;
$dist_l$: distance of flight leg l in km;
ψ_t: an adequately large leasing cost of an aircraft with type t for exceeding the number of available aircraft, where $t \in T$;

$$\alpha_{ln}: \begin{cases} 1 \text{ if flight leg } l \text{ starts at node } n \text{ (in aircraft type } t\text{'s sub-network);} \\ -1 \text{ if flight leg } l \text{ ends at node } n \text{ (in aircraft type } t\text{'s sub-network);} \\ 0 \text{ otherwise, } \forall l \in L,\ n \in N_t,\ t \in T; \end{cases}$$

$$\beta_{gn}: \begin{cases} 1 \text{ if ground arc } g \text{ starts at node } n \text{ (in aircraft type } t\text{'s sub-network);} \\ -1 \text{ if ground arc } g \text{ ends at node } n \text{ (in aircraft type } t\text{'s sub-network);} \\ 0 \text{ otherwise, } \forall g \in G_t,\ n \in N_t,\ t \in T. \end{cases}$$

Decision Variables

z_{lk}: (first-stage variable) =1 if flight leg l is flown by aircraft family k, and =0 otherwise, where $l \in L$ and $k \in K$;
buy^s: (second-stage variable) fuel purchased in litres in scenario s, where $s \in S$;
x_{lt}^s: (second-stage variable) =1 if flight leg l is flown by aircraft type t in scenario s, and =0 otherwise, where $l \in L,\ t \in T,\ s \in S$;
y_g^s: (second-stage variable) number of aircraft on ground arc g in scenario s (in aircraft type $t's$ sub-network), where $g \in G_t,\ t \in T,\ s \in S$;
q_i^s: (second-stage variable) accepted number of passengers travelling on itinerary i in scenario s, where $i \in I,\ s \in S$;
P_t^s: (second-stage variable) number of aircraft type t to be leased from other airline companies in scenario s, where $t \in T,\ s \in S$;
C^s: (second-stage variable) fuel cost in scenario s, where $s \in S$.

Accordingly, we formulate the first-stage problem as follows:

First-stage problem

$$\max \mathbb{E}[Q(\mathbf{z}, \boldsymbol{\xi})] \tag{1}$$

subject to

$$\sum_{k \in K} z_{lk} = 1, \quad l \in L \tag{2}$$

$$\mathbf{z} \in \{0,1\}^{|L| \times |K|} \tag{3}$$

where $\mathbb{E}[Q(\mathbf{z}, \boldsymbol{\xi})] = \sum_{s \in S} p^s \cdot Q^s(\mathbf{z}, \boldsymbol{\xi}^s)$, and $Q^s(\mathbf{z}, \boldsymbol{\xi}^s)$ indicates the profit under scenario s. Then, the second-stage formulation is defined for each scenario $s \in S$ as follows:

Second-stage problem

$$Q^s(\mathbf{z}, \boldsymbol{\xi}^s) = \max \left\{ \sum_{i \in I} f_i^s \cdot q_i^s - \sum_{l \in L} \sum_{t \in T} c_{lt} \cdot x_{lt}^s - \sum_{t \in T} \psi_t \cdot P_t^s - C^s \right\} \tag{4}$$

subject to

$$\sum_{t \in T_k} x_{lt}^s = z_{lk}, \quad l \in L,\ k \in K,\ s \in S \tag{5}$$

$$\sum_{l \in L} \alpha_{ln} \cdot x_{lt}^s + \sum_{g \in G_t} \beta_{gn} \cdot y_g^s = 0, \quad n \in N_t,\ t \in T,\ s \in S \tag{6}$$

$$\sum_{l \in F_t} x_{lt}^s + \sum_{g \in F_f} y_g^s - P_t^s \le A_t, \quad t \in T,\ s \in S \tag{7}$$

$$\sum_{i \in I_l} q_i^s \le \sum_{t \in T} Cap_t \cdot x_{lt}^s, \quad l \in L,\ s \in S \tag{8}$$

$$q_i^s \le d_i^s, \quad i \in I,\ s \in S \tag{9}$$

$$buy^s \ge \sum_{l \in L} \sum_{t \in T} x_{lt}^s \cdot dist_l \cdot cons_t, \quad s \in S \tag{10}$$

$$C^s = buy^s \cdot \delta^s, \quad s \in S \tag{11}$$

$$\mathbf{x}^s \in \{0,1\}^{|L| \times |T|} \tag{12}$$

$$\mathbf{y}^s, \quad \mathbf{q}^s, \quad \mathbf{P}^s : \text{non-negative integer variables}, \quad s \in S \tag{13}$$

$$buy^s \ge 0, \quad s \in S \tag{14}$$

In this formulation, the first-stage objective function (1) coupled with the second-stage objective function (4) is to maximise the expected total profit. Constraint (2) ensures that exactly one aircraft family is assigned to each flight leg l. Constraint (3) gives the domain of first-stage decision variable z. Constraint (5) requires that flight leg l is assigned to only one aircraft type t of the same family k. Constraints (6) and (7) assure the flow-conservation at each node in the time–space network and the available number of aircraft in each aircraft type is not exceeded. Constraints (8) and (9) guarantee that the number of passengers accepted does not exceed the capacity of the aircraft deployed in each flight leg, and also does not exceed the realized demand. Constraint (10) ensures that the amount of fuel purchased is not less than the amount of fuel required. Equation (11) calculates the fuel costs. Constraints (12)–(14) give ranges of second-stage decision variables.

3.3. Risk-Averse Model (Minimising the Weighted Sum of Expected Value and CVaR)

The risk-averse two-stage stochastic programming problem in the context of cost minimisation can be typically expressed as

$$\min_{\mathbf{z} \in \mathbf{Z}} \{\mathbb{E}[f(\mathbf{z}, \zeta)] + \rho \cdot \mathrm{CVaR}_\alpha(f(\mathbf{z}, \zeta))\}, \tag{15}$$

where $\mathbf{z}$ is the vector of the first-stage decision variables; ζ is the random variable; $f(\mathbf{z}, \zeta)$ is the cost function of the second-stage problem and $\mathbf{Z}$ is the domain of $\mathbf{z}$; the non-negative trade-off coefficient ρ represents the weight of the conditional value-at-risk; and $\mathrm{CVaR}_{\mathrm{ff}}$ denotes the conditional value-at-risk at confidence level α. Considering that our goal in this paper is to maximise the total profit (i.e., to minimise the negative of the total profit), the risk-averse objective function in this work can be rewritten as

$$\min_{\mathbf{z} \in \mathbf{Z}} \{\mathbb{E}[-Q(\mathbf{z}, \boldsymbol{\xi})] + \rho \cdot \mathrm{CVaR}_\alpha(-Q(\mathbf{z}, \boldsymbol{\xi}))\}, \tag{16}$$

where $\mathbf{Z} \doteq \{0,1\}^{|L| \times |K|}$. This risk-averse function incorporates both the expected value and the CVaR.

The relevant definitions of risk measure CVaR are presented as follows.

Definition 1 ([37]). *Let $F_\zeta(\lambda)$ represent the cumulative distribution function of a random variable ζ. The value at risk (VaR) at confidence level α is defined as*

$$\mathrm{VaR}_\alpha(\zeta) = \min_{\lambda \in \mathbb{R}} \{\lambda \mid F_\zeta(\lambda) \geq \alpha\}, \tag{17}$$

where $\alpha \in (0,1)$; $\mathrm{VaR}_\alpha(\zeta)$ is the α quantile of the distribution of the random variable.

Then, the CVaR, also called mean excess loss (or tail VaR), can be defined as follows.

$$\mathrm{CVaR}_\alpha(\zeta) = \mathbb{E}[\zeta \mid \zeta \geq \mathrm{VaR}_\alpha(\zeta)], \tag{18}$$

where the $\mathrm{CVaR}_\alpha(\zeta)$ is the conditional expected value of ζ exceeding the $\mathrm{VaR}_\alpha(\zeta)$.

To describe the function of $\mathrm{CVaR}_\alpha(\zeta)$ more precisely, Rockafellar and Uryasev [37] give an alternative definition of it, as shown in Definition 2.

Definition 2 ([37])**.** *The CVaR at confidence level α can be defined as*

$$\mathrm{CVaR}_\alpha(\zeta) = \min_{\lambda \in \mathbb{R}} \left\{ \lambda + \frac{1}{1-\alpha} \mathbb{E}[(\zeta - \lambda)_+] \right\}, \tag{19}$$

where $(\zeta - \lambda)_+ \doteq \max\{0, \zeta - \lambda\}, \alpha \in (0,1)$. To linearise the term, $(\zeta - \lambda)_+$, (19) can be rewritten as

$$\mathrm{CVaR}_\alpha(\zeta) = \min_{\lambda \in \mathbb{R}} \left\{ \lambda + \frac{1}{1-\alpha} \mathbb{E}[v] \right\} \tag{20}$$

subject to

$$\begin{aligned} \zeta - \lambda &\leq v \\ 0 &\leq v \end{aligned} \tag{21}$$

According to the above definitions and the assumption of finite probability space, we present a mathematical formulation for the stochastic fleet-assignment problem considering risk aversion as follows.

Risk-averse formulation

$$\min \left\{ \sum_{s \in S} p^s \cdot \left(-\sum_{i \in I} f_i^s \cdot q_i^s + \sum_{l \in L} \sum_{t \in T} c_{lt} \cdot x_{lt}^s + \sum_{t \in T} \psi_t \cdot P_t^s + C^s \right) + \rho \left(\lambda + \frac{1}{1-\alpha} \sum_{s \in S} p^s \cdot v^s \right) \right\} \tag{22}$$

subject to Constraints (2), (3) and (5)–(14)

$$\left(-\sum_{i \in I} f_i^s \cdot q_i^s + \sum_{l \in L} \sum_{t \in T} c_{lt} \cdot x_{lt}^s + \sum_{t \in T} \psi_t \cdot P_t^s + C^s \right) - \lambda \leq v^s, \quad s \in S \tag{23}$$

$$0 \leq v^s, \quad s \in S \tag{24}$$

$$\lambda \in \mathbb{R} \tag{25}$$

Note that the variable λ can be interpreted as a first-stage variable, and the variables v^s can be seen as the second-stage variables.

4. Solution Approaches

Since there exists a large number of scenarios in real-life situations, commercial solvers are too time-consuming to solve such a problem. Therefore, similar to [10], we introduce a sample average approximation (SAA) algorithm to overcome this challenge. However, SAA only works well with small-scale problems. In order to address large-scale problems more efficiently, we further develop a string-based heuristic, built on the novel concept of flight-leg string, which we will explain later.

4.1. SAA Method

SAA is a well-known approach for solving stochastic programs based on Monte-Carlo simulation [38–41]. In this work, a set of scenarios is randomly generated, corresponding to

Ω realisations of uncertain vector $\boldsymbol{\xi}$. The original expected second-stage objective function is approximated by the sample average function $\frac{1}{\Omega}\sum_{\omega=1}^{\Omega} H^{\omega}(\mathbf{z},\lambda,\boldsymbol{\xi}^{\omega})$, where we define $H^{\omega}(\mathbf{z},\lambda,\boldsymbol{\xi}^{\omega})$ as

$$H^{\omega}(\mathbf{z},\lambda,\boldsymbol{\xi}^{\omega}) = \min\left\{-\sum_{i\in I} f_i^{\omega}\cdot q_i^{\omega} + \sum_{l\in L}\sum_{t\in T} c_{lt}\cdot x_{lt}^{\omega} + \sum_{t\in T}\psi_t\cdot P_t^{\omega} + C^{\omega} + \rho\cdot\frac{1}{1-\alpha}\cdot v^{\omega}\right\} \tag{26}$$

subject to Constraints (2), (3), (5)–(14) and (23)–(25)
where s is replaced by ω, and S is replaced by $\{1,\cdots,\Omega\}$.

Accordingly, the original problem approximated by SAA can be specifically written as follows.

$$\min\left\{\frac{1}{\Omega}\sum_{\omega=1}^{\Omega}\left(-\sum_{i\in I} f_i^{\omega}\cdot q_i^{\omega} + \sum_{l\in L}\sum_{t\in T} c_{lt}\cdot x_{lt}^{\omega} + \sum_{t\in T}\psi_t\cdot P_t^{\omega} + C^{\omega} + \rho\cdot\frac{1}{1-\alpha}\cdot v^{\omega}\right) + \rho\cdot\lambda\right\} \tag{27}$$

subject to Constraints (2), (3), (5)–(14) and (23)–(25)
where s is replaced by ω, and S is replaced by $\{1,\cdots,\Omega\}$. The objective value will converge to the true optimal value when $\Omega\to\infty$ [42].

The steps of the SAA algorithm are as follows:

- Step 1. Generate a sample of size Ω, and solve the SAA model repeatedly, M times.

$$r_{\Omega}^{m} = \min_{\mathbf{z}\in\mathbf{Z},\lambda\in\mathbb{R}}\left\{\frac{1}{\Omega}\sum_{\omega=1}^{\Omega} H^{\omega}(\mathbf{z},\lambda,\boldsymbol{\xi}^{\omega}) + \rho\cdot\lambda\right\} \tag{28}$$

  Store the optimal objective value r_{Ω}^{m} and optimal solutions $\{\hat{\mathbf{z}}_{\Omega}^{m},\hat{\lambda}_{\Omega}^{m}\}$, where $m = 1,\ldots,M$.
- Step 2. Calculate $\bar{r}_{\Omega}^{M}$ and estimate the corresponding variance $\sigma^2_{\bar{r}_{\Omega}^{M}}$.

$$\bar{r}_{\Omega}^{M} = \frac{1}{M}\sum_{m=1}^{M} r_{\Omega}^{m} \tag{29}$$

$$\sigma^2_{\bar{r}_{\Omega}^{M}} = \frac{1}{M(M-1)}\sum_{m=1}^{M}\left(r_{\omega}^{m} - \bar{r}_{\Omega}^{M}\right)^2 \tag{30}$$

  where $\bar{r}_{\Omega}^{M}$ denotes the statistical upper bound of the objective value because statistically it is greater than or equal to the original value.
- Step 3. Select one of the first-stage solutions denoted by $\{\hat{\mathbf{z}},\hat{\lambda}\}$. Generate a sample of size $\Omega' >> \Omega$, and estimate the true objective function $r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})$ and its variance $\sigma^2_{r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})}$.

$$r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda}) = \frac{1}{\Omega'}\sum_{\omega=1}^{\Omega'} H^{\omega}(\hat{\mathbf{z}},\hat{\lambda},\boldsymbol{\xi}^{\omega}) + \rho\cdot\hat{\lambda} \tag{31}$$

$$\sigma^2_{r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})} = \frac{1}{\Omega'(\Omega'-1)}\sum_{\omega=1}^{\Omega'}\left(H^{\omega}(\hat{\mathbf{z}},\hat{\lambda},\boldsymbol{\xi}^{\omega}) + \rho\cdot\hat{\lambda} - r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})\right)^2 \tag{32}$$

  where $r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})$ denotes the statistical lower bound of the objective value which is an unbiased estimator of $\mathbb{E}[H(\hat{\mathbf{z}},\hat{\lambda},\boldsymbol{\xi})] + \rho\cdot\hat{\lambda}$.
- Step 4. Compute the optimality gap and the corresponding estimate variance.

$$gap_{\Omega,M,\Omega'}(\hat{\mathbf{z}},\hat{\lambda}) = \frac{\bar{r}_{\Omega}^{M} - r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})}{\bar{r}_{\Omega}^{M}} \tag{33}$$

$$\sigma^2_{\text{gap}} = \frac{\sigma^2_{\bar{r}_{\Omega}^{M}} + \sigma^2_{r_{\Omega'}(\hat{\mathbf{z}},\hat{\lambda})}}{\left(\bar{r}_{\Omega}^{M}\right)^2} \tag{34}$$

The confidence interval for the optimality gap at confidence level α can be calculated as

$$\left[gap_{\Omega,M,\Omega'}(\hat{\mathbf{z}},\hat{\lambda}) - z_\alpha(\sigma_{\text{gap}}),\quad gap_{\Omega,M,\Omega'}(\hat{\mathbf{z}},\hat{\lambda}) + z_\alpha(\sigma_{\text{gap}})\right] \tag{35}$$

with $z_\alpha = \Phi^{-1}(1-\alpha)$, where $\Phi(z)$ is the cumulative distribution function of the standard normal distribution.

4.2. The String-Based Heuristic Algorithm

In practice, the number of flight legs and fleet size are very large. However, because the SAA is only suitable for small-scale problems, it is difficult to obtain the result of the model in a large scale and short time. Therefore, in order to efficiently solve the problem, we further develop a string-based heuristic algorithm. We first introduce the novel concept of "flight-leg string" because each aircraft can be responsible for the execution of multiple flight legs every day. That is, the flight legs that meet the connection requirements will be linked together and different aircraft types will be assigned to the flight-leg strings. Connection requirements mean that flight legs can only form a flight-leg string if they meet a set of time and space constraints. These connection requirements will be described in detail in Section 4.2.1. The benefit of using flight-leg strings is that it can significantly reduce the number of aircraft required under the condition of satisfying the flow equilibrium.

As the same in Figure 1a, there are six flight legs, denoted by A to F, in the time–space network in Figure 2 (the take-off and landing times, and the place of departure and destination of each flight leg are known). However, the difference is that the aircraft type of each flight leg in Figure 2 is unknown; thus, we need to deploy specific aircraft types to these flight legs. First, we need to generate all feasible flight-leg strings. For example, in Figure 2, if flight arcs (or flight legs) C and D are connected as a flight-leg string, it means that tasks C and D are executed by the same specific aircraft type. When the aircraft arrives at airport 1, it will stay for a period of time until the departure time of flight leg D, and then fly flight leg D to the destination of airport 2. After that, we assign aircraft types to each flight-leg string instead of assigning aircraft type for each leg in the previous section. The string-based heuristic algorithm's outline is as follows:

1. According to a given flight leg set, a set of feasible flight-leg strings is generated by the DFS (Depth First Search) algorithm, which meets the given conditions;
2. With the goal of minimising the number of strings, the set of flight-leg strings covering all flight legs is filtered out through an integer programming model;
3. A new two-stage stochastic programming model is established to efficiently assign aircraft types to each string.

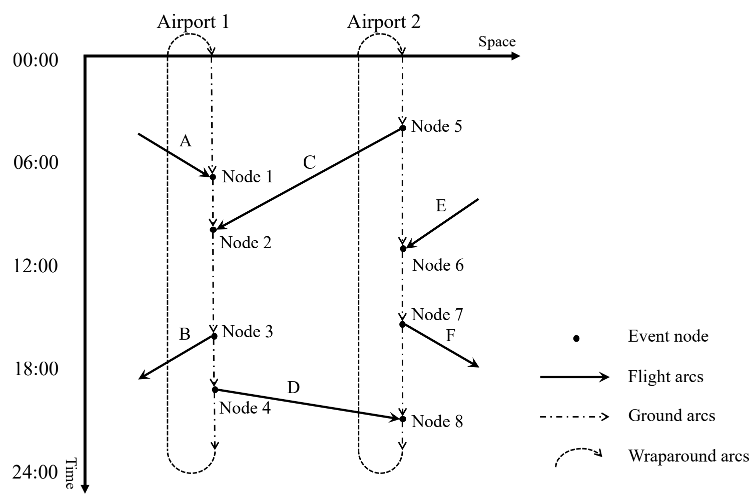

Figure 2. A time–space network with two airports and six flight legs.

Each step is detailed as follows.

4.2.1. Generate Feasible Flight-Leg Strings by DFS

First, all feasible flight-leg strings are searched through the DFS algorithm, whose pseudo-code is given in Algorithm 1. The connection between flight legs need to meet the following conditions:

1. Each aircraft to be selected should meet the minimum turn-around time requirement. That is, there is a certain time interval between the arrival time of the previous flight leg and the departure time of the next flight leg of the same aircraft;
2. For two adjacent flight legs, the destination airport of the previous flight leg should be the same as the origin airport of the next flight leg;
3. If the arrival and departure times of two adjacent flight legs are not on the same day, the two flight legs may form an overnight connection;
4. The aircraft must return to the overnight airport before 24:00 every day.

The specific algorithm (Algorithm 1) of finding feasible flight-leg strings works as follows: first, given a set of flight legs ($Flight$), pick a seed flight leg ($fliseed \in Flight$) or called root, and put it into the search path ($Path$). When the search path is not empty, put the last flight leg in the path into an empty set ($fliahead$). Then, search for all flight legs ($Flitailsetall$) from the initial set of flight legs ($Flight$) that can connect with $fliahead$ under four connection conditions mentioned above. After that, find all feasible flight legs ($Flitailset$) from the $Flitailsetall$ which is not in the path. If the feasible flight leg set ($Flitailset$) is not empty, then select a flight leg from it and add it to the last position of the path ($Path$). If the feasible flight leg set ($Flitailset$) is empty, save the existing path ($Path$) and delete the last flight leg from the path. Jump out of the second loop when the path ($Path$) is empty, then traverse the next seed flight leg until all flight legs are traversed.

Figure 3 illustrates the search process of DFS for feasible flight-leg strings. Based on Figure 2, there are a set of six flight legs (i.e., $A, B, \ldots, F$). We first select flight leg A as a root into the path. By searching all the flight legs except A, flight legs B and D are found which can be connected to A. Flight legs B and D are put into the $Flitailset$. Then, flight leg B is selected from the $Flitailset$ and placed behind A in the path. At this time, there is no flight leg that can be connected to B in the remaining flight leg set, so flight legs A and B form a flight-leg string $A \rightarrow B$. After that, flight leg B is deleted and D put into the path. It can be found that there are no more flight legs that can be connected to D in the remaining flight legs. Therefore, flight legs A and D also form a flight-leg string $A \rightarrow D$. Then, flight leg D is deleted from the path. Apparently, there is no flight leg that can connect with A any more. Thus, flight leg A is recorded as a flight-leg string, and then A deleted from the path. At this time, all traversals with A as the root are over, and flight legs $B, C, \ldots, F$ are selected as the root flight leg iteratively to traverse to find all feasible flight-leg strings. The algorithm of DFS is shown in Algorithm 1.

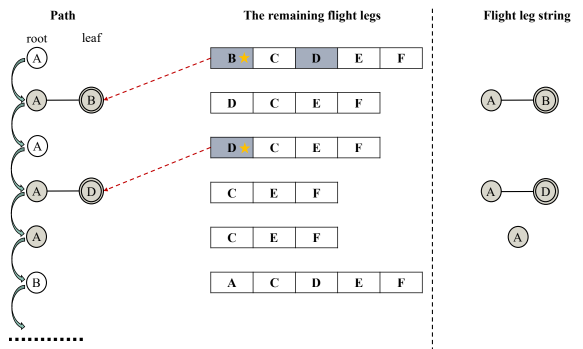

Figure 3. The formation of flight-leg strings under DFS algorithm (the rectangle in shade represents the flight leg that meets the connection requirements, the star means the selected flight leg, and the circle in shade represents the flight leg string).

Algorithm 1: DFS for feasible flight-leg string generation.

Input: A set of flight legs: $Flight$.
Output: A set of feasible flight-leg strings: $Flightstring$.

$Flightstring \leftarrow \varnothing, Path \leftarrow \varnothing$;
for $flight \in Flight$ **do**
    $fliseed \leftarrow flight$; % Select a seed flight leg from $Flight$ in turn
    $Flirest \leftarrow Flight \backslash fliseed$; % Set of remaining optional flight legs
    $Path \leftarrow Path \cup fliseed$; % Put the seed flight leg into the $Path$
    while $Path \neq \varnothing$ **do**
        $fliahead \leftarrow Path(end)$; % Put the last flight leg in the $Path$ into $fliahead$
        $Flitailsetall \leftarrow$ Search for all flight legs that can connect with $fliahead$ under four connection conditions mentioned above;
        $Flitailset \leftarrow Flitailsetall \cap Flirest$; % Exclude flight legs that are already in the $Path$
        if $Flitailset = \varnothing$ **then**
            $Flightstring \leftarrow Flightstring \cup Path$; % Update $Flightstring$
            $Path \leftarrow Path \backslash Path(end)$; % Delete the last flight leg in the $Path$
        else
            $flitail \leftarrow$ Select the flight leg in the $Flitailset$ with smallest index;
            $Path \leftarrow Path \cup flitail$; % Put $flitail$ into the $Path$
            $Flightstring \leftarrow Flightstring \cup Path$; % Update $Flightstring$
            $Flirest \leftarrow Flirest \backslash flitail$; % Update $Flirest$ by deleting the flight leg put in the $Path$
        end
    end
end
$Flightstring \leftarrow$ A set of non-repeated feasible $Flightstring$.

4.2.2. Filter Feasible Flight Strings

There are many feasible flight-leg strings obtained by the DFS algorithm, so it is necessary to filter the feasible flight-leg strings for computational efficiency. The flight-leg strings to be filtered out need to cover all flight legs, and we ensure that the number of aircraft needed is as small as possible. Therefore, we develop an integer programming model to minimise the number of flight-leg strings. Different strings contain different numbers of flight legs. In an extreme case, if all the strings we filtered out each contains only one leg, we are back to the original mathematical model (22)–(25). In another extreme case, if we filter out a set of strings such that each string contains as many legs as possible, then we placed too many restrictions on the original model (22)–(25). Clearly, the length of strings is a key parameter to tune the restrictions put on the original model. Therefore, we introduce such a parameter. In addition, we also prefer short strings, as they resemble the single legs used in the original model. Therefore, we also introduce a weight parameter to penalize a long string if selected. The input parameters, decision variables and new (restricted) mathematical model are as follows.

Input parameters

L: set of flight legs in the schedule, indexed by l, and $|L|$ denotes the number of flight arcs;
Str: set of feasible flight-leg strings, indexed by r;
β_r: non-negative weight for each flight-leg string r;
b_r: the number of flight legs per flight-leg string r;
a_{lr}: =1 if flight leg l is covered by feasible flight-leg string r; =0, otherwise.

Decision variables

x_r: =1 if feasible flight string r is selected, and =0 otherwise.

$$\min \sum_{r \in Str} \beta_r \cdot b_r \cdot x_r \tag{36}$$

subject to

$$\sum_{r \in Str} a_{lr} x_r = 1, \quad l \in L \tag{37}$$

$$x_r \in \{0,1\}, \quad r \in Str \tag{38}$$

The objective function (36) is to minimise the weighted sum of all flight-leg strings selected. Note that we can output different flight string sets by adjusting the weight β_r. We select an appropriate one using previous parameter tuning, and then use the following model (in Section 4.2.3) to evaluate the performance. Constraint (37) ensures that each flight leg l is covered by one and only one flight-leg string r. Constraint (38) gives the domain of x. Note that the integer programming model (36)–(38) output a set of flight-leg strings which can cover all the flight legs exactly once. In other words, all flight legs are covered and no flight leg can be covered more than once. This is because our objective function (36) prefers a shorter string, if constraints are satisfied.

4.2.3. String-Based Stochastic Two-Stage Programming Model

After filtering out the flight-leg strings, it is necessary to assign a specific aircraft type to each flight-leg string. Therefore, this subsection establishes a new two-stage stochastic programming model, based on the model in (22)–(25), to efficiently assign aircraft types to flight-leg strings. Only new input parameters and decision variables are introduced in the following.

Input Parameters

Str: set of flight-leg strings in the flight schedule, indexed by r;
a_{lr}: =1 if flight leg l is covered by flight-leg string r, and =0 otherwise;
$\bar{M}$: a large enough number.

Decision Variables

$\bar{z}_{rk}$: (first-stage variable) =1 if flight-leg string r is flown by aircraft family k, and =0 otherwise, where $r \in Str$ and $k \in K$.

Accordingly, we formulate the model based on the model (22)–(25) as follows:

$$\text{Objective function (22)} \tag{39}$$

subject to Constraints (2), (3), (5)–(14), (23)–(25)

$$\sum_{k \in K} \bar{z}_{rk} = 1, \quad r \in Str \tag{40}$$

$$z_{lk} \geqslant \bar{z}_{rk} - \bar{M}(1 - a_{lr}), \quad l \in L, r \in Str, k \in K \tag{41}$$

$$z_{lk} \leqslant \bar{z}_{rk} + \bar{M}(1 - a_{lr}), \quad l \in L, r \in Str, k \in K \tag{42}$$

$$\bar{\mathbf{z}} \in \{0,1\}^{|Str| \times |K|} \tag{43}$$

Constraint (40) ensures that exactly one aircraft family is assigned to each flight-leg string r. Constraints (41) and (42) establish the connect between $\bar{\mathbf{z}}$ and $\mathbf{z}$, which means if $a_{lr} = 1$, then $\bar{\mathbf{z}} = \mathbf{z}$, expressing that only one aircraft family can be assigned to each flight leg. Constraint (43) gives the domain of decision variable $\bar{z}_{rk}$.

5. Numerical Experiment

In this section, numerical experiments are conducted to evaluate the performance of our proposed solution approaches. The stochastic programming model is implemented by the off-the-shelf commercial optimisation solver CPLEX 12.8. Numerical experiments are run on a PC with Core I7 3.4 GHz processor and 8 GB RAM under Windows 10 Operating System.

5.1. Data Description

The flight-leg network in our computational experiments originates from [43], which covers 24 flight legs. Suppose the set of aircraft families is $K = \{1,2,3\}$ and the set of aircraft types is $T = \{1,2,3,4,5,6\}$ corresponding to B787-8, B787-9, A330-200, A330-300, B737-Max8 and B737-700HH, respectively [10]. The fuel consumption $cons_t$, available number of type-t aircraft A_t, seating capacity Cap_t, and cost per available seat kilometre (CASK) of each aircraft type are given in Table 2.

Since the passenger demand is confidential enterprise information, truncated normal distribution is adopted to randomly generate the passenger demand by Monte-Carlo simulation for each itinerary [8,44]. In order to better reflect the relationship between ticket price and demand, we regard ticket price as a linear function of passenger demand. The basic ticket price is generated according to the flight leg distance of the itinerary and CASK, and then different ticket prices under different scenarios are calculated, so as to estimate the target revenue correctly. Similarly, by means of historical fuel prices, the prices of jet fuel are produced in a truncated normal distribution with a standard deviation equal to 100% of its average. Considering the complexity of the problem, the running time of the SAA is limited to 3600 s.

For ease of profit analysis, we only use the negative of the objective value in (22), and do the same for the SAA and the string-based heuristic.

5.2. Computational Results

In general, a smaller gap and more accurate solutions can be found with a larger sample size Ω. In order to explore the relationship between upper and lower bounds with the change of sample size Ω in SAA, we set the sample size Ω with fixed number of samples $M = 5, \Omega' = 300, \rho = 0.5, \alpha = 0.95$. It can be observed from Table 3, the statistical upper bound (denoted by UB) of the objective decreases rapidly and converges gradually with the increase of sample size Ω. On the contrary, the statistical lower bound (denoted by LB) shows a gradual rise. Moreover, as shown in Table 3, the optimality gap drops from 37.98% to 1.59%, and 95% confidence interval narrows from [28.74, 47.23] to [−1.04, 4.22] when Ω increases from 5 to 100. This indicates that a solution closer to the true optimal value can be obtained with a larger sample size Ω. This can be explained by the fact that a larger sample size can more effectively capture changes in passenger demand and fuel price, so as to obtain a more realistic optimal solution. Therefore, when a larger sample size is used, a smaller optimality gap and a tighter optimality confidence interval can be found. Figure 4 illustrates the trends.

Table 3. Impacts of various sample size Ω on results for SAA ($M = 5, S' = 300$).

Ω	LB	UB	Gap (%)	95% Confidence Interval
5	1,140,886.30	1,839,616.95	37.98	[28.74, 47.23]
10	1,384,242.37	1,728,597.22	19.92	[11.83, 28.02]
20	1,515,947.98	1,736,227.64	12.69	[5.42, 19.96]
50	1,566,319.57	1,641,263.24	4.57	[−1.66, 10.80]
80	1,571,783.13	1,651,733.70	4.84	[0.88, 8.80]
100	1,592,381.59	1,618,120.93	1.59	[−1.04, 4.22]

Gap is calculated by $\frac{\text{UB} - \text{LB}}{\text{UB}} \times 100\%$.

The performance of the string-based heuristic by comparing it with the SAA method is shown in Table 4 with fixed $\rho = 0.5, \alpha = 0.95$. The first column in Table 4 represents the number of scenarios. The second and third columns report the average value and computational time of SAA. The fourth and fifth columns give the average value and computational time of our string-based heuristic. The gap is calculated by $\frac{\text{SAA objective} - \text{String-based heuristic objective}}{\text{SAA objective}} \times 100\%$, presented in the sixth column.

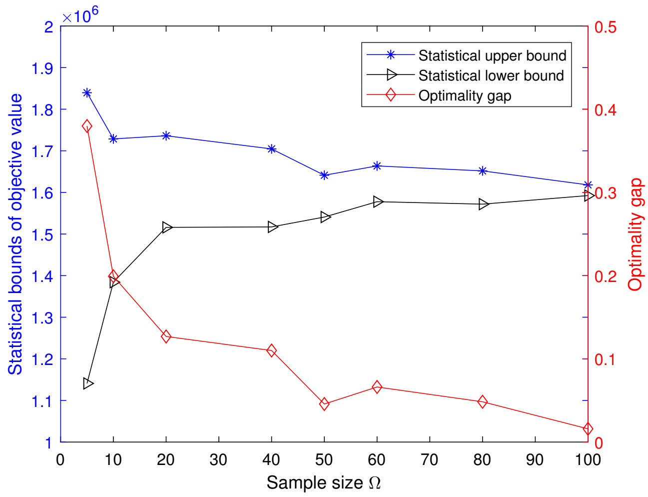

Figure 4. The illustration of SAA algprithm performance with various sample sizes.

Table 4. Results by the SAA and the string-based heuristic algorithm.

	SAA Algorithm		String-Based Heuristic		
Ω	**Obj ($\times 10^6$)**	**Time (s)**	**Obj ($\times 10^6$)**	**Time (s)**	**Gap (%)**
5	1,839,616.95	0.86	1,792,602.17	0.79	2.56
10	1,728,597.22	2.37	1,690,696.77	2.21	2.19
20	1,736,227.64	6.05	1,640,814.05	3.87	5.50
40	1,704,595.29	22.72	1,632,981.88	12.51	4.20
50	1,641,263.24	35.47	1,576,778.13	15.16	3.93
60	1,663,659.90	87.44	1,550,432.61	19.03	6.81
80	1,651,733.70	124.43	1,524,294.09	21.18	7.72
100	1,618,120.93	327.32	1,505,504.50	46.89	6.96
200	1,589,559.99	2909.68	1,450,192.18	353.19	8.77
250	-	-	1,424,505.36	1365.51	-
average		390.70		63.33	5.37 *

The average computational time of string-based heuristic is calculated except in the scenario size equal to 250. '-': there is no feasible solution. '*': calculate the average gap except for the '-' cases.

It can be found from Table 4 that when the sample size Ω is small, the calculation time of SAA is slightly smaller than that of the string-based heuristic. However, with the increase in sample size, the calculation time of the former is much longer than the latter. In addition, as the size of the problem increases, for example, when $\Omega \geq 200$, SAA cannot achieve a feasible solution in the limited time of 3600 s. In contrast, the computational time of the string-based heuristic increases relatively in line with the increase in the number of scenarios. These results show that the computational speed of our string-based heuristic is about six times faster than SAA. In terms of solution quality, we can find that SAA's objective is higher than that of our string-based heuristic, and the average gap between them is 5.37%. The reason may be that SAA can maximise the passenger demand of each flight leg, while the string-based heuristic is to maximise the passenger demand of each flight-leg string, so the string-based heuristic is inferior to SAA in terms of demand capture. However, the string-based heuristic ensures that an aircraft can fly more than one flight leg in one day and improves the daily utilization rate of the aircraft.

5.3. Sensitivity Analysis

For airline companies, different risk preferences may produce different expected profits. Therefore, we discuss the influence of different risk levels ρ and confidence level α on the decision making in SAA. Table 5 shows the computational results of the weight parameter ρ representing the risk level in the interval [0,1] with fixed $M = 5$ and the sample size $\Omega = 50$. As shown in Table 5, both the lower and upper bounds of tje objective value increase with the increase in risk level ρ. This phenomenon implies that, with the increase in risk level, the uncertainty also increases in the objective value obtained. If the decision maker wants to adopt a more risky attitude, the larger computational budget should be given.

Table 5. Impact of different risk level ρ on objective value ($\alpha = 0.95$) for SAA.

Risk Level (ρ)	LB	UB	Gap (%)	95% Confidence Interval
0	1,369,021.16	1,377,434.98	0.61	[−2.57, 3.79]
0.1	1,383,874.70	1,411,807.76	1.98	[−2.88, 6.84]
0.2	1,413,732.27	1,469,021.49	3.76	[−1.68, 9.21]
0.3	1,504,980.43	1,532,751.88	1.81	[−7.02, 10.64]
0.4	1,520,653.48	1,565,837.49	2.89	[−2.30, 8.07]
0.5	1,566,319.57	1,641,263.24	4.57	[−1.66, 10.80]
0.6	1,638,038.23	1,695,416.14	3.38	[−2.52, 9.28]
0.7	1,652,787.03	1,674,949.02	1.32	[−6.13, 8.78]
0.8	1,697,087.96	1,761,609.05	3.66	[−2.56, 9.88]
0.9	1,785,408.35	1,884,734.35	5.27	[−4.32, 14.86]
1	1,889,730.75	1,979,785.04	4.55	[−5.02,14.12]

Table 6 shows the computational results with different confidence levels α. It can be observed that the upper and lower bounds of the objective value decrease with the increase in α. A larger confidence level α means fewer scenarios are regarded as worse scenarios, which must record a worse CVaR value (i.e., a higher CVaR in our minimization model). The gap decreases first and then increases. This means an appropriate confidence level may save computational budget, e.g., $\alpha = 0.75$.

Table 6. Impact of different confidence level α on objective value ($\rho = 0.5$) for SAA.

α	LB	UB	Gap (%)	95% Confidence Interval
0.65	1,818,377.19	1,882,036.82	3.38	[2.39, 4.37]
0.70	1,807,080.04	1,867,932.98	3.26	[1.94, 4.57]
0.75	1,793,029.32	1,830,669.71	2.06	[0.19, 3.92]
0.80	1,732,724.52	1,803,608.19	3.93	[1.45, 6.41]
0.85	1,708,538.75	1,761,950.01	3.03	[−0.65, 6.72]
0.90	1,672,544.65	1,758,560.76	4.89	[0.41, 9.38]
0.95	1,566,319.57	1,641,263.24	4.57	[−1.66, 10.80]
0.99	1,472,215.07	1,606,160.13	8.34	[−2.42, 19.10]

Considering that the degree of fluctuations in passenger demand will vary with the seasons and important events of the year, we change the standard deviation of passenger demand to reflect the effect on the objective value. Table 7 shows that both in SAA and the string-based heuristic, the objective value increases with the growth in the standard deviation of the demand fluctuation. It can be found that when the standard deviation of demand fluctuation is less than 100%, the results of the string-based heuristic are better than that of SAA, indicating that in the case of small demand fluctuations, fleet assignment based on flight-leg strings can result in more profits. When the demand fluctuations become large, the objective value of our string-based heuristic is slightly lower than that of SAA.

Table 7. Impact of the standard deviation of demand for SAA and string-based heuristic.

std.dev_demand (%)	SAA	String-Based Heuristic	Gap (%)
20	1,117,808.60	1,306,114.31	−16.85
50	1,198,492.66	1,271,366.31	−6.08
100	1,641,263.24	1,576,778.13	3.93
150	2,370,769.47	2,154,198.83	9.14
200	2,773,488.68	2,791,479.11	4.40

Gap is calculated by $\frac{\text{SAA} - \text{string-based heuristic}}{\text{SAA}} \times 100\%$.

Figures 5–7 show the impact of changes in demand, fuel prices, and risk levels on the quantity of different aircraft types required by the SAA method. Note that the quantity of different aircraft types required was calculated as the average of five sampling experiment results and rounded up.

Figure 5 represents the quantity of different aircraft types required when the standard deviation of demand fluctuation increases from 20% to 200% with the standard deviation of fuel-price fluctuation set at 20% and risk level ρ set as 0.5. We can see that when the standard deviation of demand fluctuation is 20%, the demand of B737-Max8 is the largest, which is 16. When the standard deviation of demand fluctuation increases from 50% to 200%, the demand for B787-8, B737-Max8 and B737-700HH decreases, while the demand for B787-9 and A330-300 increases, and there is no significant change in demand for A330-200. The results indicate that (1) under the known aircraft capacity A330-300 > B787-9 > A330-200 > B787-8 > B737-Max8 > B737-700HH, when the standard deviation of demand fluctuation is small (that is, the demand is relatively stable), the small- and medium-sized aircraft types (i.e., B787-8, B737-Max8 and B737-700HH) are more popular; (2) when the demand fluctuates greatly, the medium and large aircraft types (i.e., B787-9 and A330-300) are more popular, indicating that when the demand is unstable but the fuel price is relatively stable, the use of medium and large aircraft types to meet the demand as much as possible can obtain more profits; (3) the capacity of the A330-200 is not much different from that of the B787-9, but the fuel consumption per kilometer of the A330-200 is relatively large, so the A330-200 has never been dominant under different demand fluctuations.

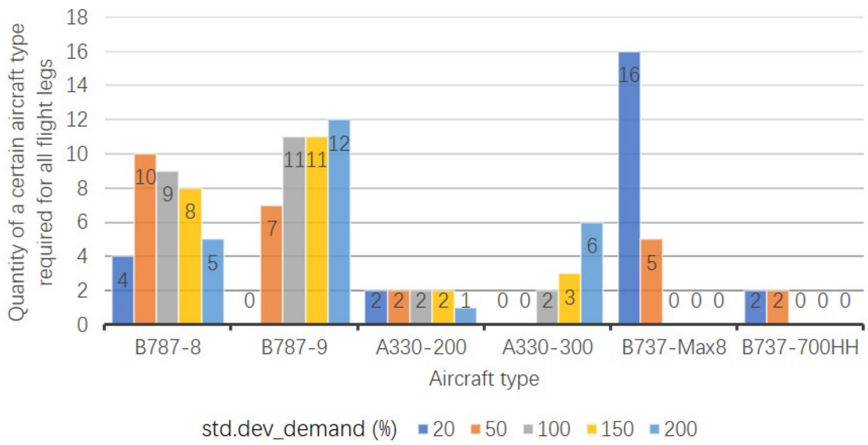

Figure 5. Impact of demand for required quantity of different aircraft types.

Figure 6 represents the required quantity of different aircraft types when the standard deviation of the fuel-price fluctuation increases from 20% to 100% with the standard deviation of the demand fluctuation set at 50% and risk level ρ set as 0.5. According to Figure 6, we can find that (1) as the standard deviation of fuel-price fluctuation increases, the demand for small aircraft types (i.e., B737-Max8 and B737-700HH) increases because the fuel consumption per kilometer of these two aircraft types is the smallest. However, in comparison, B737-Max8 has more advantages, because the unit fuel consumptions of B737-Max8 and B737-700HH are both similar, but the former has a larger capacity and can better meet the demand fluctuation of 50%; (2) with the increase in fuel-price fluctuations, the demand for medium and large aircraft (i.e., B787-8 and B787-9) decreases, but they still have a great advantage in the required quantity; (3) the aircraft types A300-200 and A330-300 have no obvious demand advantage under the fluctuation in fuel price. This is because A330 aircraft is fuel inefficient and when fuel prices fluctuate, fuel costs are too high, so they are not used, which seems to be consistent with reality.

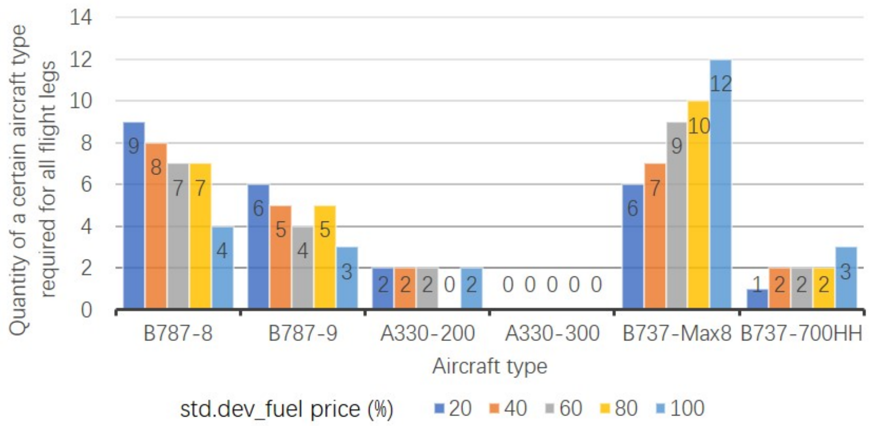

Figure 6. Impact of fuel price for required quantity of different aircraft types.

Figure 7 represents the required quantity of different aircraft types when the risk level ρ increases from 0 (risk neutral) to 1 with the standard deviation of demand fluctuation set at 50% and the standard deviation of the fuel-price fluctuation set at 40%. It can be found that with the increase in the risk level, the demand for B737-Max8 increases. Meanwhile, the demand for B787-8 and B787-9 decreases, and the demand for A330-200, A330-300 and B737-700HH does not change significantly. This shows that when demand and fuel prices fluctuate significantly, as risk aversion increases, the aircraft with smaller capacities and higher fuel efficiency is more popular because they are more flexible for controlling costs. It is worth noting that, due to a certain degree of demand fluctuation (i.e., 50%) in this experiment, although the B737-700HH has high fuel efficiency, the capacity is too small to be dominant under demand fluctuations.

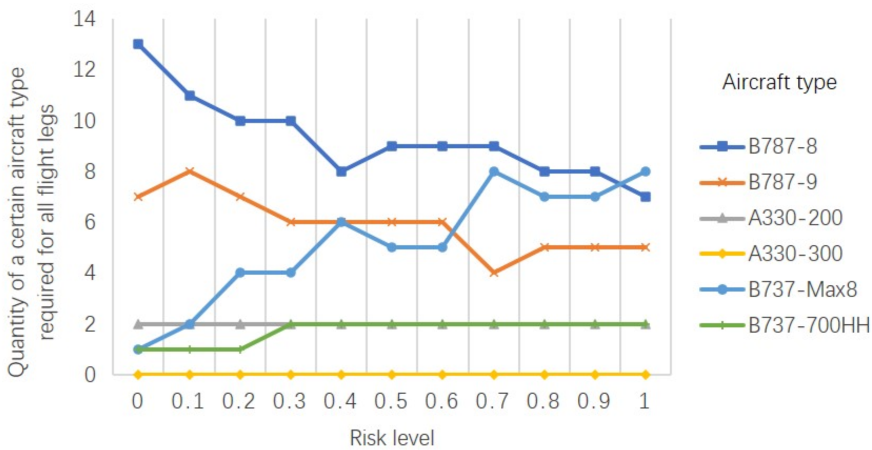

Figure 7. Impact of risk level for required quantity of different aircraft types.

In sum, the following managerial insights can be drawn:

1. A larger computational budget should be given if the decision maker wants to adopt a more risky attitude.
2. In the case of small demand fluctuations, fleet assignment based on flight-leg strings can generate more profits. When the demand fluctuates significantly, the airline can give priority to the aircraft with moderate fuel efficiency and capacity, such as B787-8 and B787-9.
3. When the fuel price fluctuates greatly, the demand for B737-Max8 with high fuel efficiency rises markedly. However, with almost the same fuel efficiency as B737-Max8, the advantage of B737-700HH is not significant, because its small capacity is not a good choice for 50% fluctuation in demand. In addition, the medium and large aircraft such as B787-8 and B787-9 are good choices when the fuel prices fluctuate very little.
4. As the degree of risk aversion increases, among the six aircraft types studied in this paper, the required proportion of B737-Max8 increases.
5. In the context of uncertainty, large capacity and fuel-inefficient aircraft such as the A330-200 and A330-300 do not have advantages, while more fuel-efficient aircraft such as B737-Max8 are clearly more popular.

6. Conclusions

In this work, we investigated a risk-averse stochastic airline fleet-assignment problem, in which both passenger demand and fuel price are assumed to be uncertain. Our established model differs from the traditional two-stage stochastic programming models by incorporating the risk-averse CVaR measure. The aim is to maximise the expected total profits under a given risk level. Then, the SAA algorithm is applied to solve our problem. Further, we developed a string-based heuristic algorithm which is more efficient at handling large-scale problems. Numerical experiments were carried out to illustrate the effectiveness of the proposed algorithms.

In future research, we intend to (1) extend the presented programming model by integrating aircraft routing and crew-scheduling problems into the airline management [45,46]; (2) consider the multimodal transportation such as the air–rail intermodality to reduce carbon emissions. These extensions may complicate the problem, and the realistic flight network may contain hundreds or thousands of flights legs. In the environment of deep uncertainty and data scarcity, more robust approaches and efficient heuristics should be developed; (3) whether uncertainty is good or bad should also be investigated [47]; (4) although the paper considers aircraft with different fuel efficiency, their emissions are not taken into account. Therefore, the incorporation of carbon-emission cost is a further research direction [48].

Author Contributions: Conceptualization, M.L.; methodology, Y.D. (Yueyu Ding); data curation, L.S.; formal analysis, R.Z.; investigation, Y.W.; validation Z.Z.; writing—original draft preparation, Y.D. (Yue Dong); writing—review and editing, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (NSFC) under Grants 72021002, 71771048, 71832001, 72071144, 71871159, and 71890973.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous referees for their constructive comments. The authors would also like to thank Feng Chu for her efforts to improve the quality of our work.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Wei, K.; Vaze, V.; Jacquillat, A. Airline timetable development and fleet assignment incorporating passenger choice. *Transp. Sci.* **2020**, *54*, 139–163. [CrossRef]
2. Ahmed, M.B.; Hryhoryeva, M.; Hvattum, L.M.; Haouari, M. A matheuristic for the robust integrated airline fleet assignment, aircraft routing, and crew pairing problem. *Comput. Oper. Res.* **2022**, *137*, 105551. [CrossRef]
3. Li, Y.; Mahmoudzadeh, A.; Wang, X.B. Airlines Seat Pricing with Seat Upgrading. *Multimodal Transp.* **2022**, *1*, 100054. [CrossRef]
4. News, C.H. Five Ways for Governments to Green Airline Bailouts. Available online: https://www.climatechangenews.com/2020/05/05/five-ways-governments-green-airline-bailouts/ (accessed on 5 May 2020).
5. Rowland, B. Is There Such a Thing as a Green Airline? Available online: https://www.oag.com/blog/green-airline (accessed on 4 May 2021).
6. Ma, Q.; Song, H.; Zhu, W. Low-carbon airline fleet assignment: A compromise approach. *J. Air Transp. Manag.* **2018**, *68*, 86–102. [CrossRef]
7. Wu, J.; Zhang, P.w.; Wang, Y.; Shi, J.J. Integrated aviation model and metaheuristic algorithm for hub-and-spoke network design and airline fleet planning. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *164*, 102755. [CrossRef]
8. Sherali, H.D.; Zhu, X. Two-stage fleet assignment model considering stochastic passenger demands. *Oper. Res.* **2008**, *56*, 383–399. [CrossRef]
9. Pilla, V.L.; Rosenberger, J.M.; Chen, V.; Engsuwan, N.; Siddappa, S. A multivariate adaptive regression splines cutting plane approach for solving a two-stage stochastic programming fleet assignment model. *Eur. J. Oper. Res.* **2012**, *216*, 162–171. [CrossRef]
10. Liu, M.; Liang, B.; Zheng, F.; Chu, F. Stochastic airline fleet assignment with risk aversion. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 3081–3090. [CrossRef]
11. Şafak, Ö.; Çavuş, Ö.; Aktürk, M.S. Multi-stage airline scheduling problem with stochastic passenger demand and non-cruise times. *Transp. Res. Part B Methodol.* **2018**, *114*, 39–67. [CrossRef]
12. Talluri, K.T. Swapping applications in a daily airline fleet assignment. *Transp. Sci.* **1996**, *30*, 237–248. [CrossRef]
13. Jiang, H.; Barnhart, C. Robust airline schedule design in a dynamic scheduling environment. *Comput. Oper. Res.* **2013**, *40*, 831–840. [CrossRef]
14. Ryerson, M.S.; Hansen, M. The potential of turboprops for reducing aviation fuel consumption. *Transp. Res. Part D Transp. Environ.* **2010**, *15*, 305–314. [CrossRef]
15. Naumann, M.; Suhl, L.; Friedemann, M. A stochastic programming model for integrated planning of re-fleeting and financial hedging under fuel price and demand uncertainty. *Procedia-Soc. Behav. Sci.* **2012**, *54*, 47–55. [CrossRef]
16. Hsu, C.I.; Eie, W.Y. Airline network design and adjustment in response to fluctuation in jet fuel prices. *Math. Comput. Model.* **2013**, *58*, 1791–1803. [CrossRef]
17. Hane, C.A.; Barnhart, C.; Johnson, E.L.; Marsten, R.E.; Nemhauser, G.L.; Sigismondi, G. The fleet assignment problem: Solving a large-scale integer program. *Math. Program.* **1995**, *70*, 211–232. [CrossRef]
18. Rushmeier, R.A.; Kontogiorgis, S.A. Advances in the optimization of airline fleet assignment. *Transp. Sci.* **1997**, *31*, 159–169. [CrossRef]
19. Barnhart, C.; Kniker, T.S.; Lohatepanont, M. Itinerary-based airline fleet assignment. *Transp. Sci.* **2002**, *36*, 199–217. [CrossRef]
20. Bélanger, N.; Desaulniers, G.; Soumis, F.; Desrosiers, J.; Lavigne, J. Weekly airline fleet assignment with homogeneity. *Transp. Res. Part B Methodol.* **2006**, *40*, 306–318. [CrossRef]
21. Dumas, J.; Aithnard, F.; Soumis, F. Improving the objective function of the fleet assignment problem. *Transp. Res. Part B Methodol.* **2009**, *43*, 466–475. [CrossRef]
22. Pilla, V.L.; Rosenberger, J.M.; Chen, V.C.; Smith, B. A statistical computer experiments approach to airline fleet assignment. *IIE Trans.* **2008**, *40*, 524–537. [CrossRef]
23. Noyan, N. Risk-averse two-stage stochastic programming with an application to disaster management. *Comput. Oper. Res.* **2012**, *39*, 541–559. [CrossRef]
24. Sherali, H.D.; Bae, K.H.; Haouari, M. Integrated airline schedule design and fleet assignment: Polyhedral analysis and benders' decomposition approach. *Informs J. Comput.* **2010**, *22*, 500–513. [CrossRef]
25. Sherali, H.D.; Bae, K.H.; Haouari, M. A benders decomposition approach for an integrated airline schedule design and fleet assignment problem with flight retiming, schedule balance, and demand recapture. *Ann. Oper. Res.* **2013**, *210*, 213–244. [CrossRef]
26. Cadarso, L.; de Celis, R. Integrated airline planning: Robust update of scheduling and fleet balancing under demand uncertainty. *Transp. Res. Part C Emerg. Technol.* **2017**, *81*, 227–245. [CrossRef]
27. Kenan, N.; Jebali, A.; Diabat, A. An integrated flight scheduling and fleet assignment problem under uncertainty. *Comput. Oper. Res.* **2018**, *100*, 333–342. [CrossRef]
28. Dong, J.X.; Lee, C.Y.; Song, D.P. Joint service capacity planning and dynamic container routing in shipping network with uncertain demands. *Transp. Res. Part B Methodol.* **2015**, *78*, 404–421. [CrossRef]
29. Liu, M.; Liu, X.; Chu, F.; Zhu, M.; Zheng, F. Liner ship bunkering and sailing speed planning with uncertain demand. *Comput. Appl. Math.* **2020**, *39*, 1–23. [CrossRef]
30. Wang, T.; Meng, Q.; Wang, S.; Qu, X. A two-stage stochastic nonlinear integer-programming model for slot allocation of a liner container shipping service. *Transp. Res. Part B Methodol.* **2021**, *150*, 143–160. [CrossRef]

31. Kliewer, N.; Mellouli, T.; Suhl, L. A time–space network based exact optimization model for multi-depot bus scheduling. *Eur. J. Oper. Res.* **2006**, *175*, 1616–1627. [CrossRef]
32. Zhang, D.; Yu, C.; Desai, J.; Lau, H.; Srivathsan, S. A time-space network flow approach to dynamic repositioning in bicycle sharing systems. *Transp. Res. Part B Methodol.* **2017**, *103*, 188–207. [CrossRef]
33. Şafak, Ö.; Gürel, S.; Aktürk, M.S. Integrated aircraft-path assignment and robust schedule design with cruise speed control. *Comput. Oper. Res.* **2017**, *84*, 127–145. [CrossRef]
34. Hainan Airlines. Fleet Information. Available online: https://www.hainanairlines.com/HUPortal/dyn/portal/DisplayPage?LANGUAGE=US&COUNTRY_SITE=US&SITE=CBHZCBHZ&PAGE=FLET (accessed on 20 May 2020).
35. Aero Corner. Boeing 737 MAX 8—Price, Specs, Photo Gallery, History. Available online: https://aerocorner.com/aircraft/boeing-737-max-8/ (accessed on 5 June 2020).
36. SAURABH. Fuel Consumption of Popular Aircraft. Available online: https://alliknowaviation.com/2019/12/14/fuel-consumption-aircraft/ (accessed on 14 December 2019).
37. Rockafellar, R.T.; Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2000**, *2*, 21–42. [CrossRef]
38. Ahmed, S.; Shapiro, A.; Shapiro, E. The sample average approximation method for stochastic programs with integer recourse. *SIAM J. Optim.* **2002**, *12*, 479–502.
39. Pagnoncelli, B.K.; Ahmed, S.; Shapiro, A. Sample average approximation method for chance constrained programming: Theory and applications. *J. Optim. Theory Appl.* **2009**, *142*, 399–416. [CrossRef]
40. Long, Y.; Lee, L.H.; Chew, E.P. The sample average approximation method for empty container repositioning with uncertainties. *Eur. J. Oper. Res.* **2012**, *222*, 65–75. [CrossRef]
41. Habibi, M.K.K.; Battaïa, O.; Cung, V.D.; Dolgui, A.; Tiwari, M.K. Sample average approximation for multi-vehicle collection—Disassembly problem under uncertainty. *Int. J. Prod. Res.* **2019**, *57*, 2409–2428. [CrossRef]
42. Bertsimas, D.; Gupta, V.; Kallus, N. Robust sample average approximation. *Math. Program.* **2018**, *171*, 217–282. [CrossRef]
43. Bazargan, M. *Airline Operations and Scheduling*; Ashgate Publishing Limited: Aldershot, UK, 2004.
44. Listes, O.; Dekker, R. A scenario aggregation–based approach for determining a robust airline fleet composition for dynamic capacity allocation. *Transp. Sci.* **2005**, *39*, 367–382. [CrossRef]
45. Salazar-González, J.J. Approaches to solve the fleet-assignment, aircraft-routing, crew-pairing and crew-rostering problems of a regional carrier. *Omega* **2014**, *43*, 71–82. [CrossRef]
46. Kenan, N.; Jebali, A.; Diabat, A. The integrated aircraft routing problem with optional flights and delay considerations. *Transp. Res. Part E Logist. Transp. Rev.* **2018**, *118*, 355–375. [CrossRef]
47. Wang, W.; Wu, Y. Is uncertainty always bad for the performance of transportation systems? *Commun. Transp. Res.* **2021**, *1*, 100021. [CrossRef]
48. Brueckner, J.K.; Abreu, C. Airline fuel usage and carbon emissions: Determining factors. *J. Air Transp. Manag.* **2017**, *62*, 10–17. [CrossRef]