



Research
Novel Methodologies in Air Transportation—Review

Airline Disruption Management: A Review of Models and Solution Methods



Yi Su^a, Kexin Xie^a, Hongjian Wang^b, Zhe Liang^{a,*}, Wanpracha Art Chaovalitwongse^c, Panos M. Pardalos^d

^a School of Economics and Management, Tongji University, Shanghai 200092, China

^b Xiamen Airlines, Xiamen 361006, China

^c Department of Industrial Engineering and Institute for Advanced Data Analytics, University of Arkansas, Fayetteville, AR 72701, USA

^d Center for Applied Optimization, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:

Received 19 April 2020

Revised 26 June 2020

Accepted 3 August 2020

Available online 9 February 2021

Keywords:

Disruption management

Aircraft recovery problem

Crew recovery problem

Integrated recovery problem

ABSTRACT

The normal operation of aircraft and flights can be affected by various unpredictable factors, such as severe weather, airport closure, and corrective maintenance, leading to disruption of the planned schedule. When a disruption occurs, the airline operation control center performs various operations to reassign resources (e.g., flights, aircraft, and crews) and redistribute passengers to restore the schedule while minimizing costs. We introduce different sources of disruption and corresponding operations. Then, basic models and recently proposed extensions for aircraft recovery, crew recovery, and integrated recovery are reviewed, with the aim of providing models and methods for different disruption scenarios in the practical implementation of airlines. In addition, we provide suggestions for future research directions in these topics.

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The aviation industry has become a major player in the global economy. As people become increasingly dependent on air travel, the number of scheduled flights worldwide grows every year. According to the International Air Transport Association (IATA) [1], civil aviation passenger demand increased by 4.2% while capacity increased by 3.4% worldwide in 2019. Hence, demand appears to grow faster than capacity, suggesting great development potential for the aviation market. With the development of the aviation industry, airline planning and scheduling problems have attracted much attention, and most airlines benefit from advanced optimization methods. Sophisticated models and effective solutions have been developed for each stage of planning, as reported in detailed overviews by Eltoukhy et al. [2] and Zhou et al. [3].

Under various circumstances, such as aircraft mechanical problems or severe weather conditions, flights cannot be operated as planned. Data provided by the Bureau of Transportation Statistics

(BTS) show that approximately 21% of flights in the United States during 2019 experienced more than 15 min of arrival delay. Similarly, the average flight on-time rate in China was 81.43% according to the Civil Aviation Administration of China (CAAC)'s *Statistical bulletin of civil aviation industry development in 2019* [4]. Likewise, the Punctuality League report by the OAG [5] shows that only three airlines achieve a greater than 90% on-time rate worldwide. Flight delays and cancellations have become important factors that affect passengers' airline preferences. Consequently, if irregular operation cannot be properly recovered, the economic and social benefits for airlines may be undermined. Therefore, disruption management has become a major problem in airline operation management.

There are many differences between airline planning and airline recovery. Planning focuses on optimization, whereas recovery targets a feasible yet possibly suboptimal solution that can be obtained in real time. Furthermore, recovery may be more uncertain than planning, depending on the degree and type of disruption. Unlike flight planning, which can be obtained several months before operation, recovery solutions should be obtained and implemented as quickly as possible after a disruption. In this paper, we focus on the characteristics of recovery and review common models and methods. Details of the models provide the

* Corresponding author.

E-mail address: liangzhe@tongji.edu.cn (Z. Liang).

characteristics and applications of recovery options that help airlines to choose appropriate methods to meet their specific requirements.

The remainder of the paper is organized as follows. Section 2 presents an introduction of disruption management, including sources of disruption and possible recovery operations. Basic models and solution methods for aircraft recovery are provided in Section 3. Section 4 presents the general model and some extensions of crew recovery. Section 5 presents integrated recovery considering multiple resources, including passengers. Section 6 summarizes our work and provides directions for future work.

2. Disruption management

Many studies have explored comprehensive recovery methods to handle disruptions efficiently and effectively. Clarke [6] reviews the practices in airline operations control centers during irregular operations. Filar et al. [7] examine disturbance handling in airports. Kohl et al. [8] provide a detailed overview of many aspects of airline disruption management and report on large-scale research on airline disruption management. Clausen et al. [9] review techniques of aviation recovery and introduce models of schedule planning. In Chapter 9 of their book, Belobaba et al. [10] present schedule recovery and robust planning to mitigate the impact of irregular operation. Barnhart and Smith [11] describe basic models of different resource recoveries and airline disruption management tools (Chapter 6.3). Changes to the mathematical models for aviation recovery problems and other related methods are also mentioned by Floudas and Pardalos [12].

Below, we briefly explain sources of disruption and disruption propagation to the airline flight network. We then describe common airline recovery operations.

2.1. Sources of disruption

We classify the sources of flight disruption into two categories:

(1) Airline resource disruption (e.g., aircraft, crew). This type of disruption is caused by factors such as additional maintenance due to aircraft mechanical failures and fuel shortage, or when crew members are absent due to illness or personal emergency.

(2) External environmental disruption (e.g., weather, air traffic control). Air travel is weather-sensitive. Even minor weather conditions may reduce the airport departure and arrival rates, causing flight delays. Under severe weather conditions, coercive measures such as airport closures and air traffic control are adopted to ensure the safety of passengers and airline assets.

As various consecutive flights are arranged for each aircraft and crew, and certain connection rules should be satisfied between

flights, disruptions are likely to affect subsequent flights, producing a down-line impact. Fig. 1 illustrates an example with two aircraft and four flights. It takes at least 30 min for aircraft to transit, and 45 min for crew members to do so.

Suppose that airport A is closed from 07:00 to 08:30 due to a thunderstorm. Therefore, flight 1 cannot depart until 8:30, rendering aircraft 2 and crew 1 unable to complete the required transference (Fig. 2). Flights 2 and 4 are both affected.

In addition, the capacity of airports and airspaces should be considered during disruptions, because airports and airspaces accept aircraft of various airlines, but the allocated resources (e.g., boarding gates, runways) are limited at a given time. Overall, even a very minor disruption can cause significant losses due to disagreements in resource allocation among airlines. Therefore, it is important to adopt recovery methods to prevent or mitigate the down-line impact.

2.2. Recovery operations

Flight recovery mainly comprises the following operations, which we explain considering the example in Section 2.1.

(1) Delaying flights. The departure time of affected flights and related flights may be delayed. In Fig. 3, the departure of flights 2 and 4 is delayed by at least 15 min to satisfy the minimum transit times of the aircraft and crew members.

(2) Cancelling flights. During recovery, if the allocated resources to carry out a flight are not feasible, or if the flight can take place but the delays would exceed a limit, the flight is cancelled. As flight cancellation incurs high costs, this operation is usually the last recovery option for airlines.

(3) Swapping resources (rerouting). When aircraft or crew members are not prepared for the next flight, other aircraft or crew available in the same airport can substitute for the original ones to carry out the flight. The recovered aircraft or crew is then reallocated to other flights when available. For example, when a disruption occurs, aircraft 1 and 2 can be swapped in airport B for flight 2 to be operated using aircraft 1 without delay, as illustrated in Fig. 4.

(4) Using reserved resources (aircraft and crew). Reserved resources are available in airports and do not perform any flight tasks.

(5) Deadheading and ferrying. Deadheading means that the crew is transported to another airport as passengers, whereas ferrying means that an aircraft is assigned to an unscheduled flight without passengers. Given the high costs incurred by these operations, they are rarely adopted.

(6) Speed controlling. Various studies have recently addressed speed controlling as a recovery operation that modifies the flight time to reduce the impact of a disruption and its corresponding delay.

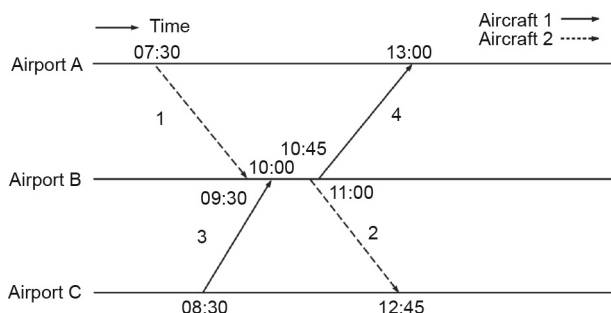


Fig. 1. An illustration of a down-line impact on a network of four flights scheduled with two aircraft and two crews. Crew 1 is assigned to flights 1 and 4, while crew 2 is assigned to flights 2 and 3.

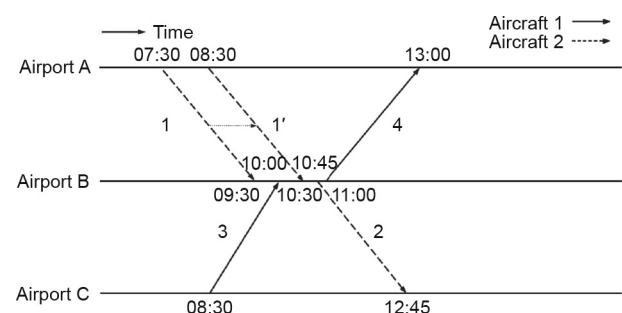


Fig. 2. An illustration of a down-line impact on a network when flight 1 cannot depart on time. The delayed flight 1 is represented by flight 1'. Flights 4 and 2 are affected due to the lateness of crew 1 and aircraft 2 which are assigned to flight 1.

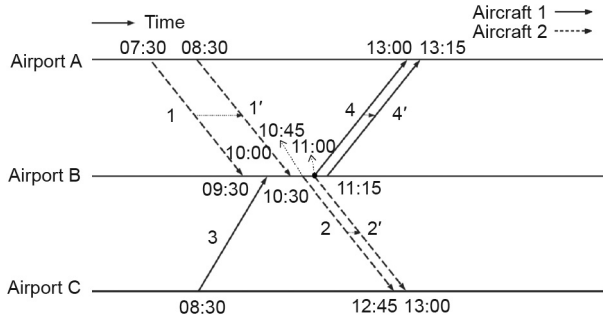


Fig. 3. Example of delaying flight 4 and flight 2 to mitigate the total delay.

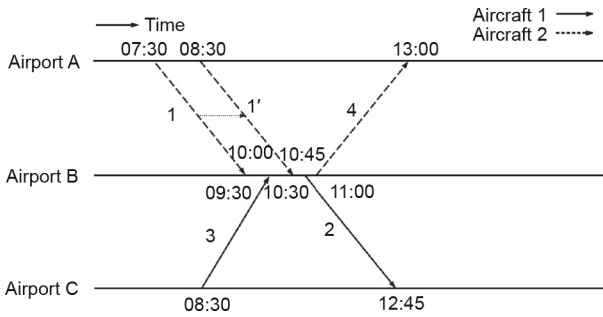


Fig. 4. Example of swapping aircrafts 1 and 2 to mitigate the total delay.

(7) Passenger reallocation. If itineraries are disrupted, passengers can be reallocated to itineraries with the same origin and destination.

Multiple recovery operations can be adopted simultaneously, according to airline preferences and capabilities. Given the complexity of adopting operations, the recovery problem is commonly separated into a sequence of subproblems that are solved in order. Usually, aircraft recovery is solved first, followed by crew recovery and passenger recovery. In the next three sections, we detail specific models and solutions for aircraft recovery, crew recovery, and integrated recovery considering passengers.

3. Aircraft recovery

As both crew rescheduling and passenger resettlement depend on the aircraft arrangement, effective reassignment of aircraft is essential for disruption recovery. Compared with that of planning, the time horizon of recovery is relatively short, varying from hours to days. Aircraft recovery aims to reschedule aircraft routes affected by disruptions at minimum cost while ensuring that flights after the recovery period will not be affected by disruptions. In addition, aircraft should be located at specific stations at the end of the recovery period to carry out the subsequent planned flights.

Aircraft recovery is typically modeled as a network problem. Like many network routing problems, the adopted models are usually arc-based or path-based, as detailed along with various extensions in the following subsections.

3.1. Arc-based model for aircraft recovery

Arc-based models allow handling of disruptions such as airport closure and air traffic control caused by severe weather, security issues, military operations, and other factors. The model is usually built in a time-space network, as proposed by Hane et al. [13], and is widely used for fleet assignment problems. The network has

three types of nodes and three types of arcs, as detailed below. An example of the time-space network is shown in Fig. 5.

3.1.1. Nodes

The supply node is the first node that indicates the beginning of the recovery period at each station when a disruption occurs. The demand node is the last node that indicates the end of the recovery period at each station. It indicates, for example, aircraft positioned in the designated airports to proceed with the flight schedule. An intermediate node is a node with time-station information representing the departure or arrival of a specific flight.

3.1.2. Arcs

A flight arc is an arc representing a flight with its scheduled departure/arrival times and stations. A ground arc is an arc representing an aircraft staying on the ground. A copied flight arc is an arc representing the delay options (e.g., 10, 30, or 60 min) for original flights. Every delay option refers to a flight arc, in which the departure (arrival) time is the scheduled departure (arrival) time plus the delay for the same departure (arrival) station.

A general arc-based model can be summarized from Refs. [14–21]. The aircraft recovery problem can be formulated on a time-space network, like the one illustrated in Fig. 5; this method was first introduced by Thengvall et al. [18].

Assuming a single-fleet problem, the general mathematical model is described below.

Model 1: Basic arc-based model for aircraft recovery

$$\min \sum_{f \in F} \sum_{t \in I(f)} c_f^t x_f^t + \sum_{f \in F} c_f z_f \quad (1)$$

$$\text{s.t.} \quad \sum_{(f,t) \in F_{n^-}} x_f^t + y_{n^-} - \sum_{(f,t) \in F_{n^+}} x_f^t - y_{n^+} = 0, \quad \forall n \in N \quad (2)$$

$$\sum_{(f,t) \in F_{b^+}} x_f^t + y_{b^+} = \text{Num}_b, \quad \forall b \in B \quad (3)$$

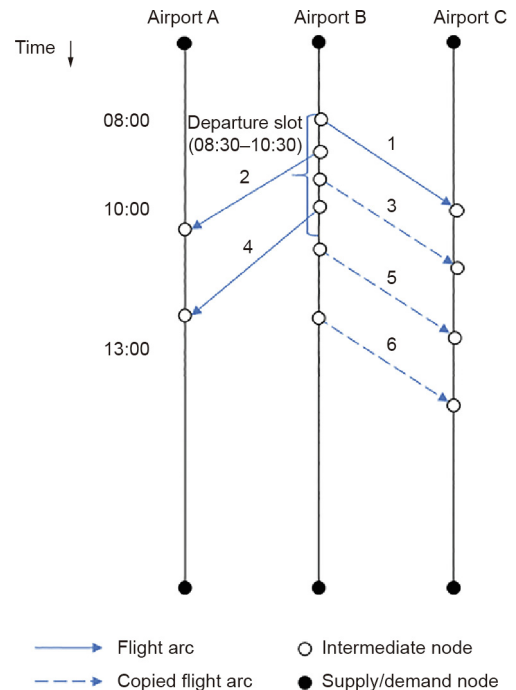


Fig. 5. Example of slot capacity for aircraft recovery.

$$\sum_{(f,t) \in F_e^-} x_f^t + y_{e^-} = \text{Num}_e, \quad \forall e \in E \quad (4)$$

$$\sum_{t \in I(f)} x_f^t + z_f = 1, \quad \forall f \in F \quad (5)$$

$$\sum_{(f,t) \in F_s} x_f^t \leq A_s, \quad s \in S \quad (6)$$

$$x_f^t \in \{0, 1\}, \quad \forall f \in F, \quad \forall t \in I(f) \quad (7)$$

$$z_f \in \{0, 1\}, \quad \forall f \in F \quad (8)$$

$$y_{n^+}, y_{n^-} > 0, \quad \forall n \in N \cup B \cup E \quad (9)$$

where F is the set of flight, among which F_{n^+} and F_{n^-} denotes the set of flight inbound to and outbound from node n , respectively; N is the set of intermediate nodes; F_{b^+} is the set of flight inbound to node b ; B is the set of supply nodes comprising the first node at each station and representing the starting point of the recovery period; F_{e^-} is the set of flight outbound from node e ; E is the set of demand nodes comprising the last nodes at each station and representing the end point of the recovery period; F_s is the set of flights involved in slot s (F_s contains the departure flights if slot s is a departure slot and all the flights if slot s restricts both departure and arrival); S is the set of disrupted slots, which may be caused by weather conditions, security issues, military operations, and other factors; $I(f)$ is the set including the original flight arc f and corresponding copied flight arcs, and $I(F) = \bigcup_f I(f)$ denotes the set of all the possible flight arcs; c_f^t is the assignment cost of the t th copy of flight f (the delay cost is included if the t th copy is not an original flight, i.e., $t > 0$); c_f is the cancellation cost of flight f ; A_s is the restricted number of aircraft allowed in slot s ; Num_b is the supplied aircraft at supply node b ; Num_e is the required aircraft for demand node e ; x_f^t is the decision variable with a value of 1 if the t th copy of flight f is chosen and a value of 0 otherwise; y_{n^+} is the number of aircraft on the ground after node n ; y_{n^-} is the number of aircraft on the ground before node n ; y_{b^+} is the number of aircraft on the ground after node b ; y_{e^-} is the number of aircraft on the ground before node e ; and z_f is the decision variable with a value of 1 if flight f is cancelled and a value of 0 otherwise.

The objective function in Eq. (1) aims to minimize the total assignment cost, delay cost, and cancellation cost. In addition, aircraft flow balance is maintained by the constraints in Eqs. (2)–(4). Specifically, the constraints in Eqs. (3) and (4) indicate that the aircraft are supplied to operate a sequence of flights and reach the demand nodes at the end of the recovery period. The flight coverage constraints in Eq. (5) ensure that each flight is either cancelled or operated according to its scheduled/delay option. The constraints in Eq. (6) require the number of departure and arrival flights involved in a slot to remain below the slot capacity.

Fig. 5 illustrates the reduction of slot capacity, where air traffic control takes place at airport B from 08:30 to 10:30, in which the departure slot capacity decreases from three to two, and flight arcs 1–4 are involved. Copied flight arcs 5 and 6 are set after the slot to represent the delayed flights due to the disruption.

Under severe weather conditions, such as hurricanes or typhoons, aircraft should be moved into hangars or be fixed by a ground lock. Otherwise, the aircraft should be relocated to other airports to ensure their safety. Therefore, the number of aircraft on the ground with respect to the hangar capacity or ground locks should be guaranteed by the corresponding constraints. The constraints in Eq. (10) limit the number of ground arcs involved in the corresponding slots.

$$y_{n^+} \leq A_g, \quad n \in N_g, \quad g \in G \quad (10)$$

where G is the set of ground arcs along with capacity; A_g is the restricted number of ground arcs; N_g denotes the set of some special intermediate nodes; and the number of aircraft on the ground after these nodes is limited.

Air traffic control can be used for a specific airport and for any set of flights of interest. Therefore, the set of flights involved in Eq. (10) can be extended to any candidate set designated by a controller.

Thengvall et al. [19] extend the single-fleet model for handling multiple fleets. The main differences between fleets are the configuration, which is mainly reflected in the number of seats and maximum fly distance, and the crew requirements. Swapping between fleets requires more stringent conditions than swapping within one fleet, as a deviation in the preassigned capacity may lead to substantial profit loss.

The multi-fleet model adopts a set of time-space networks, one for per sub-fleet [19]. As in the single-fleet model, a flight can be operated by all aircraft in the scheduled sub-fleet. Furthermore, some flights can be operated by multiple sub-fleets during recovery. As shown in Fig. 6, sub-fleet 2 can operate the flights originally scheduled to sub-fleet 1; hence, flights 2 and 3 belong to the network of sub-fleet 2. This model is solved using the IBM CPLEX Optimizer. For the instance with 12 fleets containing 1–6 sub-fleets, 1434 flights during the 24 h recovery period, and 10 h of airport closure, the model provides a near-optimal solution after 1838.4 s of computation. In this near-optimal solution, 81 flights are operated by sub-fleets different from the original assignments. This indicates that swapping between fleets provides more flexibility and improves the solution quality.

Aircraft routes cannot be directly obtained from the basic arc-based model, because the flow in the time-space network does not distinguish a specific aircraft. Therefore, the arc-based model cannot handle disruptions such as the unexpected maintenance of a specific aircraft. To solve this problem, Vink et al. [21] extend the arc-based model to incorporate specific aircraft. A set of parallel networks is established, where each network represents a specific aircraft. This set of networks is bundled by the flight coverage constraints. The model is extended by including extra aircraft index for each variable in Model 1.

Constructing the network for each aircraft and solving the corresponding integer linear program can be very time-consuming. To achieve real-time performance, Vink et al. [21] develop a selection algorithm from the proposal by Vos et al. [20] that comprises three stages. The number of selected aircraft involved in each stage is limited in order to speed up the solving process. If no solution is found by the selected aircraft, the set of candidate aircraft is

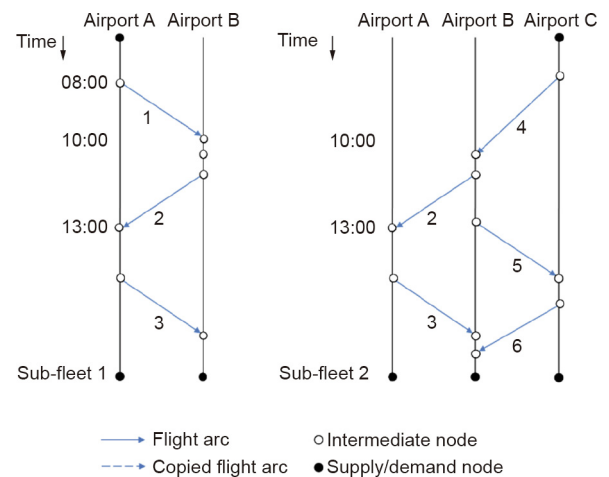


Fig. 6. Example of a time-space network for a multi-fleet model.

expanded and moved to the next stage. It takes 22 s on average to solve an instance with 100 aircraft, 600 daily flights, and a 16 h recovery period by the selection algorithm, while it takes 10 min for the integer linear program considering the entire set. The selection algorithm finds the global optimal solution in seven of ten scenarios. On average, the gap between the selection algorithm solution and the global optimal solution is 6%.

3.2. General path-based model for aircraft recovery

The disruption concerning individual aircraft can also be handled using a path-based model. Compared with the arc-based model, the path-based model assigns aircraft to a route that includes detailed information such as flight delay and aircraft swap. Argüello et al. [22], Rosenberger et al. [23], Eggenberg et al. [24], Wu et al. [25], and Liang et al. [26] have developed models based on paths (routes). The path-based model for aircraft recovery can also handle disruptions such as airport closure and air traffic control by adding side constraints to the model. In addition, situations such as unplanned maintenance of an aircraft can be handled by this model.

The path-based model enables a feasibility check (i.e., guaranteed minimum turnaround times) when generating routes. Constraints and regulations can be implicitly included during route generation and eliminated from the model. The path-based model is described below.

Model 2: Basic path-based model for aircraft recovery

$$\min \sum_{a \in A} \sum_{r \in R_a} c_r^a x_r^a + \sum_{f \in F} c_f z_f \quad (11)$$

$$\text{s.t. } \sum_{r \in R_a} x_r^a = 1, \quad \forall a \in A \quad (12)$$

$$\sum_{a \in A} \sum_{r \in R_a} \delta_f^r x_r^a + z_f = 1, \quad \forall f \in F \quad (13)$$

$$x_r^a \in \{0, 1\}, \quad \forall a \in A, \quad \forall r \in R_a \quad (14)$$

$$z_f \in \{0, 1\}, \quad \forall f \in F \quad (15)$$

where A is the set of aircraft; R_a is the set of routes that can be assigned to aircraft a ; c_r^a is the cost of assigning aircraft a to route r ; δ_f^r is the parameter with a value of 1 if route r contains flight f and a value of 0 otherwise; and x_r^a is the decision variable with a value of 1 if aircraft a is assigned to route r and a value of 0 otherwise.

The objective function in Eq. (11) aims to minimize the assignment cost and cancellation cost. The resource utilization constraints in Eq. (12) imply that each aircraft can operate one route. The origin and destination of routes r correspond to the stations where recovery starts and ends. The flight coverage constraints in Eq. (13) indicate that each flight is covered by only one aircraft, and the flight should be cancelled if no aircraft is available.

Table 1 summarizes the characteristics of the arc- and path-based models. Compared with the arc-based model, the path-based model has fewer constraints but a greater number of possible routes. As cancellation and delay are the most common recovery operations, we should analyze their formulations. Cancellation is represented by cancellation variables in both the arc- and path-based models. However, delay has a variety of representations, as detailed in Section 3.3.

Table 1

Characteristics of the arc- and path-based models.

Decision variable	Arc-based model (Leg)	Path-based model (Path)
Number of variables	$O(A I(F))$	A lot
Number of constraints	$O(A F)$	$O(F)$
Formulation of delay	Decision variables of copied flight arcs	Embedded in route during route generation
Formulation of cancellation	Decision variable	Decision variable

$|A|$: number of elements in the set of aircraft; $|F|$: number of elements in the set of flights; $|I(F)|$: number of elements in the set of all the possible flights; $O()$: space complexity, a measure of the amount of working storage an algorithm needs.

3.3. Recovery with delay

Flight delay is an important recovery operation. A short delay for recovery can sometimes be absorbed by the buffer time between flights, allowing flight connections to be maintained. Thus, selecting a delay for a minor disruption is effective.

Using copied flights to represent delay, as described by Model 1, is widely adopted in aircraft recovery. The delay options are given by a discrete set of predetermined delay times. Thengvall et al. [18,19], Eggenberg et al. [24], Vos et al. [20], and Vink et al. [21] adopt copied flights as delay options in their formulations. However, discrete delay presents some disadvantages. ① Delay may be overestimated. For example, the actual delay of a flight may be 12 min, but the only available option is 20 min, so an unnecessary delay cost is incurred. ② To prevent overestimation, a smaller time interval can be set, but doing so increases the problem size.

The departure and arrival times of flights can also be decision variables in models. Aktürk et al. [27] propose aircraft rescheduling considering swapping, delaying, and adjusting cruise speed, in which the departure delay and cruise time are decision variables, as detailed below.

Fig. 7 illustrates this model. Case I shows the original schedule of flight f and subsequent flight $N(f)$. Then, a delay combining cruise time decision is added to the original schedule in case II. The relationship of flight f and flight $N(f)$ is easily obtained, as follows:

$$r_f + l_f + t_f - t_f^0 + TA_f \leq d_{N(f)} + l_{N(f)} \quad (16)$$

where $N(f)$ is the next flight after operating flight f in the original schedule; r_f is the scheduled arrival time of flight f ; t_f^0 is the original

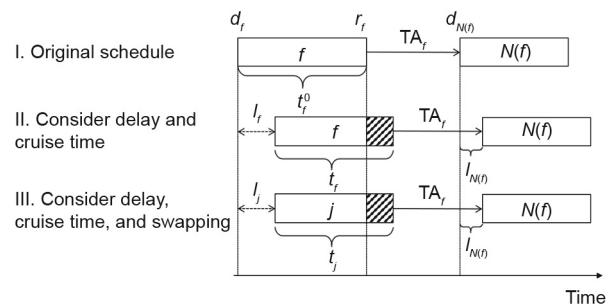


Fig. 7. Illustration of a relationship of swapping flights. I. Original schedule, in which flight $N(f)$ follows flight f . II. Schedule considering delay and cruise time reduction of flight f . III. Schedule considering delay, cruise time reduction, and swapping, in which flight f is swapped with flight j . d_f : scheduled departure time of flight f ; r_f : scheduled arrival time of flight f ; TA_f : turnaround time of flight f ; l_f : departure delay time for flight f ; t_f : departure delay time for flight f ; t_f^0 : original cruise time of aircraft operating flight f ; t_f : cruise time of aircraft operating flight f ; t_j : cruise time of aircraft operating flight j ; $N(f)$: next flight after operating flight f in the original schedule; $d_{N(f)}$: the scheduled departure time of flight $N(f)$; $l_{N(f)}$: the departure delay time for flight $N(f)$.

cruise time of aircraft operating flight f ; TA_f is the turnaround time of flight f ; l_f is the departure delay time for flight f ; t_f is the cruise time of aircraft operating flight f ; $d_{N(f)}$ is the scheduled departure time of flight $N(f)$; and $l_{N(f)}$ is the departure delay time for flight $N(f)$.

Eq. (16) indicates that if flights f and j are not swapped, the actual departure time of $N(f)$ depends on the actual arrival time of flight f . Otherwise, flight $N(f)$ follows flight j after swapping, as shown in case III of Fig. 7. The actual departure time of flight $N(f)$ is given by $r_j + l_j + t_j - t_j^0 + TA_j$. Therefore, Eq. (17) indicates that flight $N(f)$ follows either flight f or flight j . The model proposed by Aktürk et al. [27] establishes a conic quadratic mixed-integer programming and can be solved using CPLEX. It takes 196 s on average to solve an instance with 207 flights and 60 aircraft with a one day recovery period.

$$(r_f + l_f + t_f - t_f^0 + TA_f)1 \left(- \sum_{j \in S(f)} x_{f,j} \right) + (r_j + l_j + t_j - t_j^0 + TA_j) \sum_{j \in S(f)} x_{f,j} \leq d_{N(f)} + l_{N(f)} \quad (17)$$

where $S(f)$ is the set of flights that can be swapped with flight f and $x_{f,j}$ is the decision variable with a value of 1 if flights f and j are swapped and a value of 0 otherwise.

Liang et al. [26] propose an adaptive delay method for airport traffic control, in which the number of departures or arrivals of specific airports are limited in some slots. This can be achieved by introducing the constraints in Eq. (18) to the path-based model.

$$\sum_{a \in A} \sum_{r \in R_a} \phi_s^r x_r^a \leq A_s, \quad \forall s \in S \quad (18)$$

where ϕ_s^r is the number of time slots s used by route r .

This problem is solved using column generation with a multi-label shortest-path algorithm for the corresponding pricing subproblem. Two labels, referring to cost and delay, are considered in the shortest-path problem, and the dominated paths are removed during searching. To ensure subproblem optimality, the affected flights are duplicated at the beginning and end of each slot. The largest evaluated instance under disruptions of air traffic control over 2 h and unexpected maintenance over 8 h with 638 flights and 44 aircraft is optimized within 356 s.

3.4. Recovery with maintenance

Maintenance is commonly treated as a fixed activity during aircraft recovery. Eggenberg et al. [24] treat maintenance as a resource-consuming and renewal process. Column generation with

the resource-constrained elementary shortest path is used to find the routes with maintenance feasibility in each subproblem. A real-life instance of 16 aircraft and 242 flights over 7 d with a flight-to-aircraft ratio of 18.4 is solved within 3603 s.

Liang et al. [26] determine maintenance requirements using a path-based model on a connection network. Besides the consuming resource concept used by Eggenberg et al. [24], scheduled maintenance operations can be swapped. Swapping can only take place if aircraft are of the same fleet type and the resource limitation is satisfied. Maintenance swapping improves the flexibility of recovery, increasing the probability of finding better recovery solutions. Again, column generation with the resource-constrained shortest path is used. Three additional labels—namely, total fly time, number of takeoffs and landings, and elapsed time—are considered in the multilabel shortest-path algorithm. Compared with fixed maintenance, the cost of recovery with flexible maintenance decreases by 61.47%.

An arc-based model considering maintenance is proposed by Vink et al. [21]. Two types of maintenance operations are considered. One is planned maintenance, which is difficult to change to another time or station (e.g., when the maximum elapsed time between two maintenance operations is nearly reached). This maintenance must be performed as scheduled on aircraft. The other type of maintenance provides various time and space options, and at least one option is selected to meet the maintenance requirements in a route.

Table 2 [14–31] summarizes the features and solving methods of different studies on aircraft recovery approaches.

4. Crew recovery

Before presenting crew recovery, we briefly introduce crew scheduling to provide a background on some necessary concepts. Crew scheduling consists of arranging a set of flight tasks within a given period (e.g., 15 or 30 d) for the pilot, so that each flight in the schedule can be operated by one or more crews. A crew can only operate the fleet if properly qualified. Therefore, flights are grouped by fleet, and crew scheduling is solved separately for each fleet. In addition, the crew schedule should obey regulations from governing agencies and labor unions to ensure security and operability.

Compared with crew scheduling, crew recovery involves fewer airports and shorter schedule times, notably reducing the scale of the problem. To fully leverage the small scale of the problem, the flights not involved in the recovery period are usually treated as fixed activities. Hence, crew recovery aims to find a solution with

Table 2
Features and solving methods of aircraft recovery approaches.

Authors	Year	Model	Delay	Multi-fleet	Maintenance	Solving method
Jarrah et al. [28]	1993	Arc-based	Decided by ready time	No	No	Network flow algorithm
Yan et al. [14,15]	1996, 1997	Arc-based	—	No	No	Lagrangian relaxation + sub-gradient
Argüello et al. [22]	1997	Path-based	Constructed in paths	Yes	No	GRASP
Thengvall et al. [18]	2000	Arc-based	Copied flights	No	No	Solver
Bard et al. [17]	2001	Arc-based	Decided by ready time	No	No	Solver
Thengvall et al. [19,29]	2001, 2003	Arc-based	Copied flights	Yes	No	Lagrangian relaxation + sub-gradient
Rosenberger et al. [23]	2003	Path-based	Constructed in paths	No	Yes	Aircraft selection + solver
Andersson and Värbrand [16]	2004	Arc-based	Decided by ready time	No	No	Dantzig–Wolfe decomposition
Andersson [30]	2006	—	Constructed in paths	Yes	No	Tabu search and simulated annealing
Liu et al. [31]	2008	—	Decided by ready time	Yes	No	Multi-objective genetic algorithm
Eggenberg et al. [24]	2010	Path-based	Copied flights	No	Yes	Column generation + multilabel shortest path
Aktürk et al. [27]	2014	—	Decision variables	Yes	No	Solver
Vos et al. [20]	2015	Arc-based	Copied flights	No	Yes	Aircraft selection + solver
Wu et al. [25]	2017	Path-based	Constructed in paths	No	No	Distributed fixed-point computation
Liang et al. [26]	2018	Path-based	Adaptive delay	No	Yes	Column generation + multilabel shortest path
Vink et al. [21]	2020	Arc-based	Copied flights	Yes	Yes	Aircraft selection + solver

GRASP: greedy randomized adaptive search procedure.

a minimum cost for reassigning the available crews to affected flights, while crews start and end at given fixed activities [32].

Similar to crew scheduling, a crew recovery problem is usually constructed on a single-fleet network. Labor regulations applicable to planning should be complied during recovery. Moreover, preassigned activities, such as vacation and training, should be reflected in the roster of the crew after recovery.

Teodorović and Stojković [33] are the first to develop a mechanism to solve the crew recovery problem during daily operation. They adopt two methods to construct a crew duty—namely, a first-in-first-out scheme and dynamic programming—to minimize the ground time of a crew. However, the optimality of the problem is neglected. Subsequent studies have applied optimization to solve the crew recovery problem. Below, we present a general model for crew recovery as well as an extension and the corresponding method.

4.1. General model for crew recovery

Wei et al. [32], Stojković et al. [34], Lettovský et al. [35], Yu et al. [36], and Guo et al. [37] use a set partitioning model to solve the crew recovery problem. Below, we detail the formulation proposed by Lettovský et al. [35].

Model 3: Basic model for crew recovery

$$\min \sum_{k \in K} \sum_{p \in P} c_p^k x_p^k + \sum_{f \in F} c_f z_f + \sum_{f \in F} c_f^D s_f + \sum_{k \in K} q_k v_k \quad (19)$$

$$\text{s.t.} \sum_{k \in K} \sum_{p \in P} \beta_f^p x_p^k + z_f - s_f = 1, \forall f \in F \quad (20)$$

$$\sum_{p \in P} x_p^k + v_k = 1, \forall k \in K \quad (21)$$

$$x_p^k \in \{0, 1\}, k \in K, p \in P \quad (22)$$

$$s_f \in \text{Integer}, z_f \in \{0, 1\}, f \in F \quad (23)$$

$$v_k \in \{0, 1\}, k \in K \quad (24)$$

where K is the set of crews; P is the set of pairings; c_p^k is the cost of assigning crew k to pairing p ; c_f^D is the deadhead cost of flight f ; q_k is the cost of idle crew k ; β_f^p is the parameter with a value of 1 if pairing p contains flight f and a value of 0 otherwise; x_p^k is the decision variable with a value of 1 if crew k is assigned to pairing p and a value of 0 otherwise; v_k is the decision variable with a value of 1 if crew k has no assignment and a value of 0 otherwise; and s_f is the number of crews with deadhead for flight f .

This model considers coverage constraints and assignment constraints. The objective function in Eq. (19) aims to minimize the pairing cost, cancellation cost, deadhead cost, and cost of idle crews. Eq. (20) indicates that a flight can be covered more than once, and the surplus coverage identifies the number of flight f used for deadheading. If flight f cannot be covered, the cancellation variable z_f has a value of 1 and incurs the corresponding cancellation cost. Eq. (21) indicates that one crew can operate at most one pairing.

Wei et al. [32] solve the above formulation using a heuristic algorithm, in which a set of pairings is generated using a shortest-path algorithm on a time-space network. This model aims to find a feasible schedule with minimum deviation from the original plan by setting different costs for flight arcs. An instance of 6 airports, 51 flights, and 18 pairings over two days is evaluated. The computation time varies from 1 to 6 s.

In addition, the structure of crews should be considered for operational solutions. Crew members are usually divided into

three ranks: captain, first officer, and second officer. To obtain a realistic solution, Medard and Sawhney [38] consider a multi-rank extension by modifying Eq. (20) into Eq. (25).

$$\sum_{k \in K} \sum_{p \in P} \beta_f^p x_{p,q}^k \geq n_{f,q}, \forall f \in F, \forall q \in Q_f \quad (25)$$

where Q is the set of ranks, with Q_f denoting the ranks required for flight f ; $n_{f,q}$ is the minimum number of crew members required for rank q of flight f ; and $x_{p,q}^k$ is the decision variable with a value of 1 if crew k for rank q is assigned to pairing p and a value of 0 otherwise.

In Eq. (25), $x_{p,q}^k$ now reflects the information of rank q . This set of constraints ensures that the minimum number of members with different ranks is satisfied.

4.2. Extensions of crew recovery

Crew recovery can also affect flight delay. Stojković and Soumis [39] consider flight delays when constructing feasible pairings. Moreover, a set of constraints to protect the passenger connection is added to the master problem.

$$\sum_{k \in K} \sum_{p \in P} (b_{j,p}^k - b_{w,p}^k) x_p^k \geq d_{w,j} \forall (w,j) \in W \quad (26)$$

where $b_{j,p}^k$ is the departure time of flight j operated by crew k in pairing p ; $b_{w,p}^k$ is the departure time of flight w operated by crew k in pairing p ; and W is the set of ordered flight pairs (w,j) which are contained in aircraft routes or passenger connections.

If arc (w,j) is in the aircraft route, then $d_{w,j} = bl_w + g_w$, indicating the block time (bl_w) plus the turnaround time (g_w). If there is an important passenger connection on arc (w,j) , then $d_{w,j} = bl_w + c_{w,j}$, indicating the block time plus the passenger connection time ($c_{w,j}$). The flight precedence constraints in Eq. (26) protect the aircraft connections and passenger connections by limiting the period between consecutive flights in a pairing.

Stojković and Soumis [40] develop a multi-rank model by adding synchronization constraints. A set of copies from original flights, defined as tasks, represent the different rank requirements. As these tasks belong to the same flight, Eq. (27) is added to the model of Stojković and Soumis [39].

$$\sum_{k \in K} T_u^k = a_f + T_f, \forall f \in F, \forall u \in N_f \quad (27)$$

where a_f is the earliest departure time of flight f ; T_f is the delay of flight f ; T_u^k is the departure time of task u operated by crew k ; and N_f represents the set of tasks associated with flight f .

Eq. (27) ensures that every task associated with flight f has the same departure time. Column generation is used to solve this multi-rank crew pairing problem and add the synchronization constraints to the master problem. Eq. (27) is rewritten as follows:

$$\sum_{k \in K} \sum_{p \in P} b_{u,p}^k x_p^k = a_f + T_f, \forall f \in F, \forall u \in N_f \quad (28)$$

The model is applied to data from domestic flights in the United States. An instance of 190 flights, consisting of 46 and 20 flights with flexible and fixed departure times, respectively, is evaluated. Each flight requires five crew members, and the number of involved crews is 97. This instance is solved in 1237 s.

Abdelghany et al. [41] use a greedy heuristic to iteratively solve the crew recovery problem. The disrupted flights are grouped into different stages, in which the flights are resource independent, as shown in Fig. 8. Resource independence is defined as those flights not sharing the same crew resources in a stage. Hence, in each iteration, the resource-independent flights are assigned to the available crews using the following formulation.

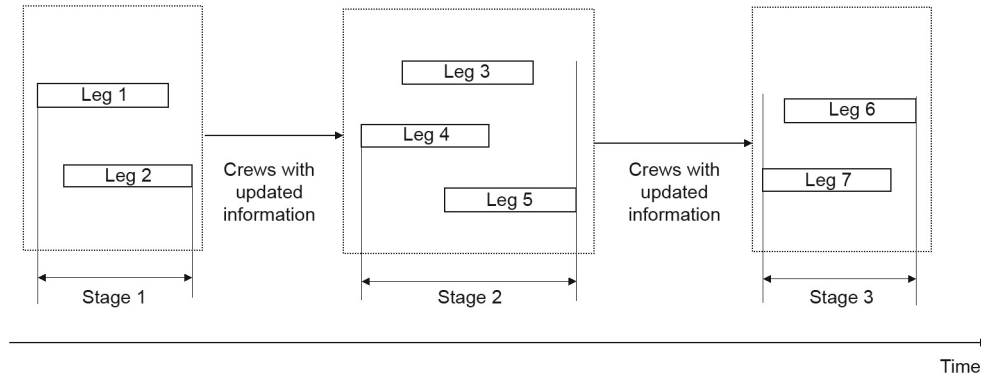


Fig. 8. A rolling framework for crew recovery.

$$\sum_{k \in K} x_{k,f,q} = 1, \forall f \in F, q \in Q_f \quad (29)$$

$$\sum_{f \in F} \sum_{q \in Q_f} x_{k,f,q} \leq 1, \forall k \in K \quad (30)$$

$$dt_f - a_k x_{k,f,q} \geq 0, \forall f \in F, q \in Q_f, k \in K \quad (31)$$

$$dt_f - d_f \geq 0, \forall f \in F \quad (32)$$

$$at_f \leq v_k x_{k,f,q}, \forall f \in F, q \in Q_f, k \in K \quad (33)$$

$$at_f = dt_f + bl_f, \forall f \in F \quad (34)$$

where d_f is the scheduled departure time of flight f ; bl_f is the scheduled block time of flight f ; a_k is the ready time for crew k ; v_k is the specific time in which maximum on-duty time is met for crew k ; $x_{k,f,q}$ is the decision variable with a value of 1 if crew k assigned to flight f for position q ; dt_f is the actual departure time of flight f ; and at_f is the actual arrival time of flight f .

Eqs. (29) and (30) ensure the coverage of flights and that each crew has at most one rank. The constraints in Eqs. (31) and (32) indicate that the actual departure time should not be earlier than the scheduled departure time and the crew-ready time. The constraints in Eq. (33) ensure that the arrival time does not exceed the duty limit. The constraints in Eq. (34) indicate the relationship between departure and arrival times. Data from major airlines in the United States with an 8 h recovery period and 121 crews is solved in 1 min and 51 s. Although this rolling framework cannot retrieve a globally optimal solution, the solution is practical and can be obtained in real time.

To limit the model size and speed up recovery, Lettovský et al. [35] divide the partial pairings in the recovery period into segments. A partial pairing is illustrated in Fig. 9. The flights in duty 2 are split into segments that consist of one or more flight legs. Segments are covered instead of flights, thus reducing the number

of rows in the model. Using segments may reduce both the number of generated pairings and the computation time.

4.3. Solving methods for the crew recovery problem

The general model for crew recovery includes few constraints but a huge number of variables; hence, most recovery approaches are solved using column generation. Column generation is an efficient method to avoid explicitly enumerating all the variables while maintaining the optimality of linear programming relaxation. Under column generation, a path for crew k is generated in the subproblem, and the optimal set of paths for recovery is obtained in the master problem. Although the problems in Refs. [34,35,38,40–42] are solved using column generation, the construction of the networks differs by adopting methods based on elements such as flights, segments, and duties. Table 3 [32–44] summarizes the features and solving methods of different studies on crew recovery.

5. Integrated recovery

In this section, we present characteristics and considerations for the integration of different resources to perform recovery.

5.1. Integrated aircraft and crew recovery

Both aircraft and crews play important roles during recovery, with the former being the scarcest resource for airlines. Complex restrictions and the large scale of the problem hinder the integrated modeling of these two resources.

5.1.1. Basic link

The basic link between aircraft and crew recovery problems is provided by flight cancellation and delay decisions [45–51]. As detailed in Sections 3 and 4, the basic models of aircraft and crew

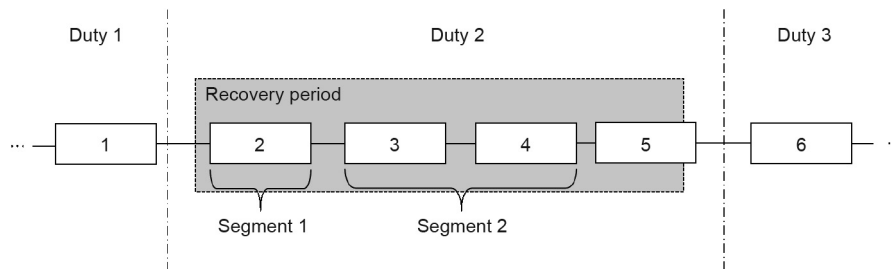


Fig. 9. Illustration of segments. Segment 1 contains one flight (i.e., flight 2), while segment 2 consists of two flights (i.e., flight 3 and flight 4).

Table 3

Features and solving methods of crew recovery approaches.

Authors	Year	Basic element for construction	Solving method
Teodorović and Stojković [33]	1995	Flight	First-in-first-out and dynamic programming
Wei et al. [32]	1997	Flight	Depth-first search + B&B
Stojković et al. [34]	1998	Duty	Column generation + resource-constrained shortest path + B&B
Lettofský et al. [35]	2000	Segment	Column generation + primal-dual simplex + B&B
Stojković and Soumis [39]	2001	Flight	Column generation + resource-constrained shortest path + B&B
Yu et al. [36]	2003	Flight	Depth-first search + B&B
Abdelghany et al. [41]	2004	Flight	Rolling horizon optimization
Guo et al. [37]	2005	Flight	Genetic algorithm + local search
Stojković and Soumis [40]	2005	Flight	Column generation + resource-constrained shortest path + B&B
Nissen and Haase [43]	2006	Duty	Column generation + resource-constrained shortest path + B&B
Medard and Sawhney [38]	2007	Flight	Column generation/depth-first search + solver
Chang [44]	2012	Duty	Genetic algorithm

B&B: branch and bound.

recovery have the common variable z_f , which represents the cancellation of a flight.

Maier [50] uses the connection network for recovery by focusing on aircraft and crews. The network uses flight copies to represent different delays. He integrates a path-based model for aircraft recovery and a general model for crew recovery by adding delay consistency constraints.

$$\sum_{k \in K} \sum_{p \in D_k} b_{fp}^{k,v} x_p^k - s_f^v - \sum_{a \in A} \sum_{r \in R_a} b_{fr}^{a,v} x_r^a = 0, \quad \forall f \in F, \quad \forall v \in U_f \quad (35)$$

where U_f is the set of discrete copies of flight f indexed by copy v ; D_k is the set of duties that can be assigned to crew k ; $b_{fr}^{a,v}$ is the parameter with a value of 1 if flight f on route r contains copy v of flight f for aircraft a and a value of 0 otherwise; $b_{fp}^{k,v}$ is the parameter with a value of 1 if flight f on duty p contains copy v of flight f for crew k and a value of 0 otherwise; and s_f^v is the number of crew deadheading operations on copy v of flight f .

The constraints in Eq. (35) ensure that the delay on each flight is consistent for the crews and aircraft. Moreover, the number of delay consistency constraints changes depending on the input paths and duties.

5.1.2. Aircraft and crew compatibility

A crew member cannot operate all types of fleets. As a pilot requires the ability to fly a specific aircraft when carrying out a flight, it is important to ensure resource compatibility. Abdelghany et al. [46] and Arkan et al. [51] build arc-based models considering these compatibility constraints, which are described as follows in Ref [46]:

$$(x_f^k + x_f^a)(1 - \zeta_{k,a}) \leq 1, \quad \forall f \in F, \quad \forall a \in A, \quad \forall k \in K \quad (36)$$

where $\zeta_{k,a}$ is the parameter with a value of 1 if crew k is eligible to operate aircraft a and a value of 0 otherwise and x_f^o is the decision variable with a value of 1 if resource o (aircraft a or crew k) is assigned to flight f and a value of 0 otherwise.

The constraints in Eq. (36) ensure that only compatible aircraft and crew can be assigned to a flight. In addition to these constraints, the model by Abdelghany et al. [46] includes coverage constraints for aircraft and crews, and other constraints to ensure delay time feasibility. A greedy iterative heuristic is used to solve the integrated recovery problem. It takes about 36 s to obtain a recovery plan for an instance with 522 aircraft, 1360 pilots, 2040 flight attendants, and 1100 daily flights after ten disrupted flights.

Arkan et al. [51] adopt a special flight network that contains four types of nodes—namely, scheduled flight nodes, source nodes, sink nodes, and must-visit nodes—which represent maintenance requirements or scheduled crew rest periods. Then, an arc-based

model is established on this network and the resource compatibility constraints are formulated as Eq. (37).

$$x_{fg}^k \leq \sum_{a \in A_k} x_{fg}^a, \quad \forall (f, g) \in \text{Conn}, \quad \forall k \in K \quad (37)$$

where Conn is the set of arcs between nodes f and g indexed by (f, g) ; A_k is the set of aircraft that can be operated by crew k ; and x_{fg}^o is the decision variable with a value of 1 if resource o covers arc (f, g) and a value of 0 otherwise. When the resource is an aircraft, the variable is denoted by x_{fg}^a . When the resource is a crew, the variable is denoted by x_{fg}^k .

If a crew is assigned to a flight, the crew should be qualified to operate the corresponding aircraft. By using the problem size control algorithm and passenger aggregation, the total running time for a network containing 1254 flights with 402 aircraft is reduced to less than 12 min when the cruise speed-controlling method is considered.

It is notable that, as the arc-based model is adopted, it is very difficult—or even impossible—to formulate the complex legality constraints for pairings or duties of crews.

5.1.3. Connection feasibility

Different resources, such as aircraft, crews, and passengers, have different transit time requirements. Therefore, the control of the flight interval by delay decisions partly determines whether reasonable and high-quality routings and duties can be generated. Various integrated recovery studies [47,48,50] have incorporated connection considerations into the construction of routings and duties to omit additional constraints that guarantee the feasibility in their models.

The connection feasibility constraints are included in the model proposed by Arkan et al. [51]. These constraints can be formulated as Eq. (38). If a resource is assigned to an arc, the period between the arrival of the previous flight and the departure of the next flight should comply with the minimum transit time required for the resources.

$$dt_g \geq at_f + CT_{fg}^o x_{fg}^o - \overline{AT}_f (1 - x_{fg}^o), \quad \forall o \in \text{Res}, \quad \forall (f, g) \in CC^o \quad (38)$$

where CC^o is the set of connection-critical arcs between nodes. The minimum possible connection time of the arc is below the minimum connection time required for resource o . Res is the set of resources including aircraft and crews indexed by o ; CT_{fg}^o is the minimum connection time required for resource o of arc (f, g) ; and \overline{AT}_f is the latest arrival time of flight f .

As solving an integrated model is computationally intractable, Zhang et al. [49] use a two-stage heuristic algorithm for the integrated recovery problem. The model mainly decomposes the integrated recovery problem into two models to solve the aircraft

and crew recovery problems separately. The two stages are linked by the connection feasibility between consecutive flights for aircraft and crews.

When crew members fly the same aircraft consecutively, there is no need to obey the crew transit time requirements. This can be considered as a short connection, and considering this situation increases the solution flexibility. Maher [48,50] takes this into account. The corresponding variables are described in Section 5.1.1, and the situation can be formulated as Eq. (39). When an aircraft is assigned to a route containing a short connection, the pairing that contains the same short connection is covered by a crew.

$$\sum_{k \in K} \sum_{p \in P} \gamma_{fg}^{k,p} x_p^k - \sum_{a \in A} \sum_{r \in R_a} \gamma_{fg}^{a,r} x_r^a \leq 0, \forall (f, g) \in SC \quad (39)$$

where SC is the set of short connections (connection time below minimum transit time for crew but possibly satisfying aircraft transit) indexed by (f, g) ; $\gamma_{fg}^{k,p}$ is the parameter with a value of 1 if connection (f, g) is contained in pairing p for crew k and a value of 0 otherwise; and $\gamma_{fg}^{a,r}$ is the parameter with a value of 1 if connection (f, g) is contained in route r for aircraft a and a value of 0 otherwise.

Nevertheless, it remains difficult to obtain a complete recovery schedule for aircraft and crews.

5.2. Integrated aircraft recovery considering passengers

Passenger recovery is important for airlines, as irregular schedules can adversely affect the itineraries of passengers. In serious cases, passengers may fail to reach their destination, and airlines may have to refund the tickets or arrange flights with other airlines for them. Hidden costs include the airline's loss of credibility, which cannot be easily estimated and which invisibly affects passengers' future travel choices. Therefore, recovering the itineraries of affected passengers quickly and reasonably improves the market competitiveness of airlines.

5.2.1. Integration of itinerary-based passenger recovery

Most studies on integrated recovery considering aircraft and passengers are based on itineraries [45,47,48,52–55]. In these studies, a binary variable is usually adopted to determine whether an itinerary is affected. An itinerary may be disrupted by either flight cancellation or violation of passenger transit times. This variable and some related constraints in Eqs. (40) and (41) are detailed below.

$$\alpha_i \geq z_f, \forall i \in IT, \forall f \in Itin(i) \quad (40)$$

$$\begin{aligned} & \sum_{a \in A} \sum_{r \in R_a} td_{Itin(i,l+1),r}^D \delta_{Itin(i,l+1),r}^r x_r^a - \sum_{a \in A} \sum_{r \in R_a} td_{Itin(i,l),r}^A \delta_{Itin(i,l),r}^r x_r^a \\ & + t_{(Itin(i,l),Itin(i,l+1))} \geq T_{\min C} - M\alpha_i, \forall i \in IT, \\ & \forall l \in 1, \dots, |Itin(i)| - 1 \end{aligned} \quad (41)$$

where $Itin(i)$ is the set of flights in itinerary i ; IT is the set of itineraries; $t_{(f_1, f_2)}$ is the scheduled departure time of flight 2 minus scheduled arrival time of flight 1; $td_{f,r}^D$ is the delay of departure for copy of flight f in route r ; $td_{f,r}^A$ is the delay of arrival for copy of flight f in route r ; $T_{\min C}$ is the minimum transit time for passengers; M is a large number; $Itin(i, l)$ is the l th flight leg in itinerary i ; and α_i is the decision variable with a value of 1 if itinerary i is disrupted and a value of 0 otherwise.

The constraints in Eq. (40) ensure that if a flight is cancelled, its itinerary is disrupted. The constraints in Eq. (41) ensure that if the passenger transit time is not satisfied in an itinerary, the itinerary is disrupted. In the objective function of the model proposed by Arıkan et al. [52], the items related to passengers include delay cost and spill cost. In addition to the disruption of itineraries, some pas-

sengers may “spill” (i.e., be left over) due to capacity shortage when aircraft is swapped. The fuel cost is expressed as a nonlinear function related to cruise time and is included in the objective function. The problem is reformulated into a model with a linear objective function and conic quadratic constraints. Limited by model complexity, passenger reallocation is not part of the formulation in Ref. [52].

Regarding integrated recovery considering passengers, many studies [47,48,51,54–56] have addressed the rearrangement of passengers. The construction of the corresponding constraints uses a judgment variable in Eqs. (42)–(45).

$$\sum_{m \in \Gamma(i)} \rho_i^m + h_i = PN_i, \forall i \in IT \quad (42)$$

$$\sum_{i \in IT} \sum_{m \in \Gamma(i)} \phi_f^m \rho_i^m \leq \sum_a \sum_{r \in R_a} \delta_f^r Cap_a x_r^a, \forall f \in F \quad (43)$$

$$\sum_{m \in IT} \rho_m^i \leq \sum_{m \in IT} PN_m (1 - \alpha_i), \forall i \in IT \quad (44)$$

$$PN_i (1 - \alpha_i) = \rho_i^i, \forall i \in IT \quad (45)$$

where $\Gamma(i)$ is the set of candidate itineraries for target itinerary i (including itinerary i); PN_i is the number of passengers originally allocated to itinerary i ; ϕ_f^m is the parameter with a value of 1 if flight f is in itinerary m and a value of 0 otherwise; Cap_a is the capacity of aircraft a ; h_i is the number of passengers originally allocated to itinerary i to be served by other airlines or refunded; and ρ_i^m is the number of passengers from itinerary i to be assigned to itinerary m .

The constraints in Eq. (42) guarantee that all passengers on any itinerary are either successfully transported to their destination or refunded. The constraints in Eq. (43) ensure that the number of passengers cannot exceed the capacity of the aircraft assigned to the flight. The constraints in Eq. (44) guarantee that no passenger is allocated to a disrupted itinerary. The constraints in Eq. (45) ensure that passengers follow the original itinerary without any change under no disruption.

Limited by the scale of the problem and the allowable computation time, although various models have considered passenger reallocation, approximations are usually adopted without considering reallocation constraints but by penalizing the number of disrupted itineraries in the objective function. Marla et al. [53] use a mechanism called flight planning that enables flight speed changes in long-haul flights. Unlike the method in Ref. [52], Marla et al. [53] discretize the dynamic selection of flying time instead of considering the cruise time as a variable. An accurate model and an approximate model have been proposed. As obtaining the solution to the accurate model is very time-consuming, the approximate model is used for case analysis. Similarly, Bratu and Barnhart [45] combine aircraft, crew, and passenger recovery by proposing two models called the disrupted passenger metric model (DPM) and the passenger delay metric model (PDM). The latter can more accurately calculate the cost of passenger delay by allowing passenger rescheduling. In an instance involving the reallocation of more than 80 000 passengers, the average calculation time of the PDM is approximately 25 times that of the DPM.

5.2.2. Integration of flight-based passenger recovery

Besides itinerary-based passenger recovery, other methods are available for passenger recovery. The method proposed by Arıkan et al. [51] treats all resources, including aircraft, crews, and passengers, as the same entities and maintains the connection feasibility and flow balance of each entity in a flight network. However, this method increases the network scale, and entity aggregation should be used to reduce the number of entities without

compromising optimality. Hu et al. [56] construct a model based on a time-band network and propose a flight-based passenger-transiting mechanism, in which flights with the same departure and arrival airports are arranged for passengers whose itinerary is disrupted. Maher [48] builds a flight-based passenger recovery model based on the point-to-point network, which is considered to be non-multi-itinerary. The formulation of the model extends the airline recovery formulation in Ref. [50] to include passenger recovery.

5.2.3. Heuristic-based methods

Various heuristic algorithms exhibit high performance for simultaneously reassigning aircraft and passengers in case of disruptions.

Bisaillon et al. [57] propose a large neighborhood search heuristic algorithm that is divided into construction, repair, and improvement phases. The first two phases generate a feasible solution. The third phase implements large changes to improve the solution based on the solution from the previous phases, while preserving feasibility. The construction phases are then executed again iteratively. This algorithm provides high-quality solutions and can handle large-scale problems in real time. Sinclair et al. [58] include additional steps for each phase of the algorithm in Ref. [13] to improve performance. Experimental results show that the destroy-and-create step, which is added to the third phase, has the greatest improvement on the quality of the solution.

Sinclair et al. [59] also propose a column generation post-optimization heuristic algorithm. An integrated linear programming model based on a time-space network is established. The algorithm in Ref. [58] is used to obtain the initial solution and to construct constraints of the restricted master problem (RMP). Then, in order to obtain suitable passenger rearrangement options, the corresponding column generation subproblem is solved.

Jozefowicz et al. [60] propose the new connection flights heuristic method, which is a three-stage heuristic algorithm based on shortest-path calculation. The first stage integrates various types of disruptions into existing schedules. In the second stage, passengers with a disrupted itinerary are assigned to candidate feasible itineraries by solving the corresponding shortest-path problem. In the third stage, a new sub-rotation is inserted into the existing aircraft rotation to allocate additional passengers.

Hu et al. [61] use a greedy randomized adaptive search procedure (GRASP) algorithm for the integrated recovery of aircraft and passengers after airline operation disruption. Suitable passenger reassignment can be obtained based on new aircraft routing in

each iteration by applying the GRASP algorithm. The iterations proceed until one of the stopping criteria is met.

5.3. Integrated airline recovery problem

Thus far, few studies have fully addressed integrated recovery [45,47,48,51], as detailed throughout this section. The model proposed by Bratu and Barnhart [45] is one of the first attempts to solve fully integrated problems, but only considers reserved crews. To obtain real-time solutions for integrated recovery, Arıkan et al. [51] use two preprocessing methods to considerably reduce the number of constraints and variables, and an algorithm to limit the scope of recovery.

Petersen et al. [47] integrate five subproblems for recovery of the schedule, aircraft, crew, itinerary, and passengers. They propose a Benders decomposition framework to solve the integration problem. An example with 800 daily flights and two fleets is solved by means of the integrated method and sequential solution, respectively. Although the integrated method shows higher performance, it requires further improvements to solve larger-scale problems within 30 min.

Maher [48] uses the column-and-row generation method. For integrated recovery, the column generation subproblems are solved for the duty and aircraft routing variables using a shortest-path algorithm. A passenger rearrangement scheme is formulated as a knapsack problem in column generation. The rows, which are added to the model iteratively, represent passenger reallocation constraints and delay consistency on each flight between aircraft and crew. A point-to-point schedule with 262 flights and a hub-and-spoke schedule comprising 441 flights are solved in 427 and 400 s, respectively.

Some characteristics of existing integrated recovery approaches are summarized in Table 4 [45–61].

6. Conclusions and future work

In this review, we presented basic models and extensions to disruption management for aircraft and crew recovery, as well as integrated recovery considering passengers. For aircraft recovery, we reviewed the basic arc-based and path-based models and discussed the situations that these models adapt to. We also summarized the formulations related to delay and maintenance. In terms of crew recovery, we reviewed the basic model and the extensions related to crew ranks and delay. For integrated recovery, the key constraints linking different resources and related algorithms were

Table 4
Characteristics of integrated recovery approaches.

Authors	Year	Type	Model for aircraft recovery	Solving method
Bratu and Barnhart [45]	2006	ACP	AB	—
Abdelghany et al. [46]	2008	AC	AB	Rolling horizon framework
Jafari and Zegordi [54]	2010	AP	AB	Rolling horizon framework
Jafari and Zegordi [55]	2011	AP	PB	—
Bisaillon et al. [57]	2011	AP	—	Large neighborhood search heuristic
Petersen et al. [47]	2012	ACP	PB	Benders decomposition and column generation
Jozefowicz et al. [60]	2013	AP	—	Heuristic approach based on shortest-path problems
Sinclair et al. [58]	2014	AP	—	Large neighborhood search heuristic
Hu et al. [56]	2015	AP	AB	—
Maher [48]	2015	ACP	PB	Column-row generation
Zhang et al. [49]	2015	AC	AB	Two-stage heuristic algorithm
Arıkan et al. [52]	2016	AP	AB	Conic quadratic reformulation
Hu et al. [61]	2016	AP	—	GRASP
Maher [50]	2016	AC	PB	Column-row generation
Sinclair et al. [59]	2016	AP	—	Column generation post-optimization heuristic
Arıkan et al. [51]	2017	ACP	AB	Conic quadratic reformulation
Marla et al. [53]	2017	AP	AB	—

ACP: aircraft, crew, and passengers; AP: aircraft and passengers; AC: aircraft and crew; AB: arc-based; PB: path-based.

introduced. After reviewing the development of disruption management, we provided the following suggestions for future research. In recovery using one stage, more realistic factors should be considered in order for the solution to be more practical and useful. For example, aircraft swapping between different fleets, as a flexible yet more complicated recovery method, can be modeled. In addition, solving the fully integrated problem still remains challenging due to its complexity. Future research should focus on integrating several stages and devising optimization methods to solve the integrated models. In integrated recovery, preprocessing to reduce the size of input data can be adopted in order to reduce the complexity. Decomposition may be used as a preprocessing approach. The loss due to disruption mainly comes from passengers. Therefore, passenger-centered recovery is a promising research topic. For example, passenger preferences should be included for recovery in order to improve service. Likewise, evaluating passenger information through data mining may lead to more accurate optimization.

Acknowledgement

This study is supported by the National Natural Science Foundation of China (71825001 and 71890973).

Compliance with ethics guidelines

Yi Su, Kexin Xie, Hongjian Wang, Zhe Liang, Wanpracha Art Chaovalitwongse, and Panos M. Pardalos declare that they have no conflict of interest or financial conflicts to disclose.

References

- [1] The International Air Transport Association. Infographic: economics air passenger demand [Internet]. Montreal: The International Air Transport Association; 2019. Available from: https://www.iata.org/en/iata-repository/publications/economic-reports/Infographic_Economics_Air_Passenger_Demand_2019/.
- [2] Eltokhy AEE, Chan FTS, Chung SH. Airline schedule planning: a review and future directions. *Ind Manage Data Syst* 2017;117(6):1201–43.
- [3] Zhou L, Liang Z, Chou CA, Chaovalitwongse WA. Airline planning and scheduling: models and solution methodologies. *Front Eng Manag* 2020;7:1–26.
- [4] Civil Aviation Administration of China. Statistical bulletin of civil aviation industry development in 2019. Beijing: Civil Aviation Administration of China; 2020. Chinese.
- [5] OAG. Punctuality League [Internet]. Luton: OAG; 2020. Available from: <https://www.oag.com/punctuality-league-2020-report?submissionGuid=8ba7ac6c-475b-4821-9456-4c8203499733>.
- [6] Clarke MDD. Irregular airline operations: a review of the state-of-the-practice in airline operations control centers. *J Air Transp Manage* 1998;4(2):67–76.
- [7] Filar JA, Manyem P, White K. How airlines and airports recover from schedule perturbations: a survey. *Ann Oper Res* 2001;108(1–4):315–33.
- [8] Kohl N, Larsen A, Larsen J, Ross A, Tiourine S. Airline disruption management—perspectives, experiences and outlook. *J Air Transp Manage* 2007;13(3):149–62.
- [9] Clausen J, Larsen A, Larsen J, Rezanova NJ. Disruption management in the airline industry—concepts, models and methods. *Comput Oper Res* 2010;37(5):809–21.
- [10] Belobaba P, Odoni A, Barnhart C, editors. The global airline industry. Hoboken: John Wiley and Sons, Ltd.; 2009.
- [11] Barnhart C, Smith B, editors. Quantitative problem solving methods in the airline industry: a modeling methodology handbook. Berlin: Springer; 2012.
- [12] Floudas CA, Pardalos PM, editors. Encyclopedia of optimization. 2nd ed. Berlin: Springer; 2009.
- [13] Hane CA, Barnhart C, Johnson EL, Marsten RE, Nemhauser GL, Sigismondi G. The fleet assignment problem: solving a large-scale integer program. *Math Program* 1995;70(1–3):211–32.
- [14] Yan S, Yang DH. A decision support framework for handling schedule perturbation. *Transp Res Pt B Methodol* 1996;30(6):405–19.
- [15] Yan S, Lin CG. Airline scheduling for the temporary closure of airports. *Transport Sci* 1997;31(1):72–82.
- [16] Andersson T, Värbrand P. The flight perturbation problem. *Transp Plann Technol* 2004;27(2):91–117.
- [17] Bard JF, Yu G, Argüello MF. Optimizing aircraft routings in response to groundings and delays. *IEE Trans* 2001;33(10):931–47.
- [18] Thengvall BG, Bard JF, Yu G. Balancing user preferences for aircraft schedule recovery during irregular operations. *IEE Trans* 2000;32(3):181–93.
- [19] Thengvall BG, Yu G, Bard JF. Multiple fleet aircraft schedule recovery following hub closures. *Transp Res Pt A Policy Pract* 2001;35(4):289–308.
- [20] Vos HWM, Santos BF, Omondi T. Aircraft schedule recovery problem—a dynamic modeling framework for daily operations. *Transp Res Procedia* 2015;10:931–40.
- [21] Vink J, Santos BF, Verhagen WJC, Medeiros I, Filho R. Dynamic aircraft recovery problem—an operational decision support framework. *Comput Oper Res* 2020;117:104892.
- [22] Argüello MF, Bard JF, Yu G. A grasp for aircraft routing in response to groundings and delays. *J Comb Optim* 1997;1(3):211–28.
- [23] Rosenberger JM, Johnson EL, Nemhauser GL. Rerouting aircraft for airline recovery. *Transport Sci* 2003;37(4):408–21.
- [24] Eggenberg N, Salani M, Bierlaire M. Constraint-specific recovery network for solving airline recovery problems. *Comput Oper Res* 2010;37(6):1014–26.
- [25] Wu Z, Li B, Dang C, Hu F, Zhu Q, Fu B. Solving long haul airline disruption problem caused by groundings using a distributed fixed-point computational approach to integer programming. *Neurocomputing* 2017;269:232–55.
- [26] Liang Z, Xiao F, Qian X, Zhou L, Jin X, Lu X, et al. A column generation-based heuristic for aircraft recovery problem with airport capacity constraints and maintenance flexibility. *Transp Res Pt B Methodol* 2018;113:70–90.
- [27] Aktürk MS, Atamtürk A, Gürel S. Aircraft rescheduling with cruise speed control. *Oper Res* 2014;62(4):829–45.
- [28] Jarrah AIZ, Yu G, Krishnamurthy N, Rakshit A. A decision support framework for airline flight cancellations and delays. *Transport Sci* 1993;27(3):266–80.
- [29] Thengvall BG, Bard JF, Yu G. A bundle algorithm approach for the aircraft schedule recovery problem during hub closures. *Transport Sci* 2003;37(4):392–407.
- [30] Andersson T. Solving the flight perturbation problem with meta heuristics. *J Heuristics* 2006;12(1–2):37–53.
- [31] Liu TK, Jeng CR, Chang YH. Disruption management of an inequality-based multi-fleet airline schedule by a multi-objective genetic algorithm. *Transp Plann Technol* 2008;31(6):613–39.
- [32] Wei G, Yu G, Song M. Optimization model and algorithm for crew management during airline irregular operations. *J Comb Optim* 1997;1(3):305–21.
- [33] Teodorović D, Stojković G. Model to reduce airline schedule disturbances. *J Transp Eng* 1995;121(4):324–31.
- [34] Stojković M, Soumis F, Desrosiers J. The operational airline crew scheduling problem. *Transport Sci* 1998;32(3):232–45.
- [35] Lettovský L, Johnson EL, Nemhauser GL. Airline crew recovery. *Transport Sci* 2000;34(4):337–48.
- [36] Yu G, Argüello M, Song G, McCowan SM, White A. A new era for crew recovery at continental airlines. *Interfaces* 2003;33(1):5–22.
- [37] Guo Y, Suhl L, Thiel MP. Solving the airline crew recovery problem by a genetic algorithm with local improvement. *Oper Res* 2005;5(2):241–59.
- [38] Medard CP, Sawhney N. Airline crew scheduling from planning to operations. *Eur J Oper Res* 2007;183(3):1013–27.
- [39] Stojković M, Soumis F. An optimization model for the simultaneous operational flight and pilot scheduling problem. *Manage Sci* 2001;47(9):1290–305.
- [40] Stojković M, Soumis F. The operational flight and multi-crew scheduling problem. *Yugosl J Oper Res* 2005;15(1):25–48.
- [41] Abdelghany A, Ekkollu G, Narasimhan R, Abdelghany K. A proactive crew recovery decision support tool for commercial airlines during irregular operations. *Ann Oper Res* 2004;127(1–4):309–31.
- [42] Desrochers M, Soumis F. A generalized permanent labeling algorithm for the shortest path problem with time windows. *Inf Syst Res* 1988;26(3):191–212.
- [43] Nissen R, Haase K. Duty-period-based network model for crew rescheduling in European airlines. *J Sched* 2006;9(3):255–78.
- [44] Chang SC. A duty based approach in solving the aircrew recovery problem. *J Air Transp Manage* 2012;19:16–20.
- [45] Bratu S, Barnhart C. Flight operations recovery: new approaches considering passenger recovery. *J Sched* 2006;9(3):279–98.
- [46] Abdelghany KF, Abdelghany AF, Ekkollu G. An integrated decision support tool for airlines schedule recovery during irregular operations. *Eur J Oper Res* 2008;185(2):825–48.
- [47] Petersen JD, Sölveling G, Clarke JP, Johnson EL, Shebalov S. An optimization approach to airline integrated recovery. *Transport Sci* 2012;46(4):482–500.
- [48] Maher SJ. A novel passenger recovery approach for the integrated airline recovery problem. *Comput Oper Res* 2015;57:123–37.
- [49] Zhang D, Henry Lau HYK, Yu C. A two stage heuristic algorithm for the integrated aircraft and crew schedule recovery problems. *Comput Ind Eng* 2015;87:436–53.
- [50] Maher SJ. Solving the integrated airline recovery problem using column-and-row generation. *Transport Sci* 2016;50(1):216–39.
- [51] Arıkan U, Gürel S, Aktürk MS. Flight network-based approach for integrated airline recovery with cruise speed control. *Transport Sci* 2017;51(4):1259–87.
- [52] Arıkan U, Gürel S, Aktürk MS. Integrated aircraft and passenger recovery with cruise time controllability. *Ann Oper Res* 2016;236(2):295–317.

- [53] Marla L, Vaaben B, Barnhart C. Integrated disruption management and flight planning to trade off delays and fuel burn. *Transport Sci* 2017;51(1):88–111.
- [54] Jafari N, Zegordi SH. The airline perturbation problem: considering disrupted passengers. *Transp Plann Technol* 2010;33(2):203–20.
- [55] Jafari N, Zegordi SH. Simultaneous recovery model for aircraft and passengers. *J Franklin Inst* 2011;348(7):1638–55.
- [56] Hu Y, Xu B, Bard JF, Chi H, Gao M. Optimization of multi-fleet aircraft routing considering passenger transiting under airline disruption. *Comput Ind Eng* 2015;80:132–44.
- [57] Bisailon S, Cordeau JF, Laporte G, Pasin F. A large neighbourhood search heuristic for the aircraft and passenger recovery problem. *4OR-Q J Oper Res* 2011;9(2):139–57.
- [58] Sinclair K, Cordeau JF, Laporte G. Improvements to a large neighborhood search heuristic for an integrated aircraft and passenger recovery problem. *Eur J Oper Res* 2014;233(1):234–45.
- [59] Sinclair K, Cordeau JF, Laporte G. A column generation post-optimization heuristic for the integrated aircraft and passenger recovery problem. *Comput Oper Res* 2016;65:42–52.
- [60] Jozefowiez N, Mancel C, Mora-Camino F. A heuristic approach based on shortest path problems for integrated flight, aircraft, and passenger rescheduling under disruptions. *J Oper Res Soc* 2013;64(3):384–95.
- [61] Hu Y, Song Y, Zhao K, Xu B. Integrated recovery of aircraft and passengers after airline operation disruption based on a GRASP algorithm. *Transport Res E-Log* 2016;87:97–112.