International Journal of Industrial Engineering Computations 15 (2024) 171–188

Contents lists available at GrowingScience

International Journal of Industrial Engineering Computations

homepage: www.GrowingScience.com/ijiec

Research on optimization of flight crew scheduling considering pilot fatigue

Hui Lin[a], Chao Guo[a,b*], Jianxin You[a] and Ming Xia[a]

[a]*School of Economics and Management, Tongji University, Shanghai, 200092, China*
[b]*Civil Aircraft Flight Test Center, Commercial Aircraft Corporation of China, Ltd., Shanghai, 201323, China*

C H R O N I C L E

Article history:
Received August 5 2023
Received in Revised Format
September 6 2023
Accepted October 23 2023
Available online
October 23 2023

Keywords:
Crew Scheduling
Pilot Fatigue
Alertness
Optimization
Mixed Integer Programming
Column Generation

A B S T R A C T

Safety is a top concern for the civil aviation industry, and the risk of safety accidents will increase due to pilot fatigue. To ensure the safety of civil aviation, this paper proposes a method to solve the crew scheduling problem considering pilot fatigue. In order to reflect individual differences and fatigue levels of pilots, an improved three-stage alertness calculation model is first proposed based on subjective and objective perspectives to represent pilots' alertness levels and fatigue working duration quantitatively. Then, for the crew scheduling problem considering pilot fatigue, a mixed integer programming model is constructed to simultaneously achieve the optimization objectives of reducing the overall scheduling cost and crew fatigue working duration. Next, since the actual crew scheduling problem is large-scale, a solution algorithm based on a column generation framework is developed to improve the quality and efficiency of solving the large-scale crew scheduling problem. Furthermore, in the case study, we collected actual data from an airline company to validate the effectiveness of our proposed method. Finally, through multiple experimental comparisons and analyses, to balance the two optimization objectives mentioned above, it is more reasonable to handle pilot fatigue working duration with soft constraints. Sensitivity analysis reveals the variation rules of the crew cost and fatigue, providing some valuable managerial insights for the crew scheduling problem considering pilot fatigue.

1. Introduction

Flight safety is a crucial safeguard for the continuous development of the civil aviation industry. Pilot fatigue poses an increased risk of safety incidents, making pilot alertness a vital prerequisite for ensuring safe operations and fulfilling their responsibilities (Arsintescu et al., 2022; Junya & Ruishan, 2022). International flight crews often face the challenge of working during periods of low circadian rhythm, contributing to a prevalent state of fatigue (Dijk & Czeisler, 1995). Fatigue not only heightens the likelihood of accidents but also undermines pilots' work efficiency (Oken et al., 2006). Compared to shorter flights, international flights have a more significant impact on pilot alertness, leading to sleep disturbances (Reis et al., 2016). Moreover, take-off and landing times, as well as workload, further influences pilots' fatigue levels (Honn et al., 2016; Bennett, 2019). The crew scheduling involves two sub-problems, which are crew task pairing and crew assignment (Deng et al., 2023). Traditional crew scheduling models primarily focus on minimizing economic costs. However, crew scheduling cannot solely consider economic costs because safety is of paramount concern in the civil aviation industry. China Civil Aviation Regulations Part 121 Revision 5 (CCAR-121R5) further enhances fatigue management for flight crews by specifying their flight and rest times, ensuring they do not work continuously for more than five days. It indicates the regulatory agencies' increased emphasis on crew fatigue and risk management. Nevertheless, some airlines have proposed higher crew scheduling requirements to further alleviate crew fatigue, such as reducing the number of early morning and late-night shifts and improving the regularity and stability of scheduling plans (Yildiz et al., 2017).

---

* Corresponding author
E-mail: gcdzq@163.com (C. Guo)
ISSN 1923-2934 (Online) - ISSN 1923-2926 (Print)
2024 Growing Science Ltd.
doi: 10.5267/j.ijiec.2023.10.005

Based on the analysis and discussion of the existing research on pilot fatigue, it is currently only relying on regulations and company requirements to uniformly manage factors that are prone to fatigue of crew members. A margin that is too wide can easily lead to a significant decline in crew resource utilization, while a margin that is too narrow can reduce the level of crew fatigue. This paper intends to pre-measure the fatigue levels of pilots at each time point in the flight or duty planning cycle and control their fatigue levels more accurately by combining their individual needs and differences in physical fitness. Through a unified calculation method, the fatigue level of each crew member in the planning period is balanced from a quantitative level, to rationally arrange flight duty tasks in the planning stage of crew resource scheduling. It relieves the overall fatigue levels of the crew while reducing the total labor cost of the planning crew, which can further ensure flight safety.

The main contributions of this study are as follows. Firstly, considering the influence of subjective and objective factors on pilot fatigue, this study constructs an improved three-stage model of pilot alertness to quantify fatigue levels of pilots with different characteristics. Secondly, this study expands the crew scheduling problem. In other words, in the crew scheduling problem involving the number of duty days of flight loops and crew members' preference characteristics, the consideration of pilot fatigue is added to reduce the labor cost and total fatigue time of the crew. A mixed integer programming model is constructed to solve the extended crew scheduling problem. Finally, the crew scheduling problem is large-scale, so the large-scale crew scheduling problem is divided into the master problem and sub-problems. A solution algorithm based on column generation is designed to solve the mixed integer programming model for the crew scheduling problem mentioned above.

This study is organized as follows. In Section 2, the literature is reviewed, and research gaps are presented. Section 3 describes the main professional terms involved in the crew scheduling problem and the solution method for this problem. In Section 4, an improved three-stage model of pilot alertness and a mixed integer programming model with the objectives of reducing the overall scheduling cost and crew fatigue levels are constructed for solving the crew scheduling problem considering pilot fatigue. Section 5 develops a solution algorithm based on the framework of the column generation algorithm for solving the mixed integer programming model mentioned in Section 4. Section 6 presents numerical experiments and the related analysis. The conclusions of this paper are summarized in Section 7.

2. Literature review

2.1 Research status of integration problems related to crew scheduling

Crew pairing and crew assignment, as two parts of the crew scheduling plan, are closely related. Scholars' solution ideas can be divided into two categories: one is to add consideration to the characteristics of the crew assignment problem for the subsequent solutions in the crew pairing problem. The other is the model of integrating crew assignment and pairing problems, directly solving the crew scheduling plan with flights as input. Some scholars have studied the first category of solution approaches. Yan and Chang (2002) considered factors such as multiple crew bases, mixed aircraft types, and multiple cabin assignments when solving the flight crew pairing problem. The above considerations made the solution of the crew pairing scheme better match the subsequent personnel assignment problem. To improve the coupling of the two-stage solution steps, Quesnel et al. (2017) added constraints on the total working time limit of each personnel base in the crew pairing problem, and then proposed four branch pricing heuristic algorithms to solve the problem. Quesnel et al. (2020) added language constraints to the solution of the crew pairing problem to improve the usability of the crew pairing scheme, considering that the requirement of language qualification of crew members on certain flights should be met when personnel assignment is carried out. Saddoune et al. (2011, 2012) established a spatiotemporal network model to represent the synthesis problem and solved it by the column generation algorithm. Then, they proposed a variety of dynamic constraint aggregation methods to overcome the degradation problem of solving large cases. Zeighami and Soumis (2019) and Zeighami et al. (2020) proposed an extension of the crew pairing problem, which includes problems that need to be solved in the subsequent assignment problem without time conflicts with the occupation of pilot training, vacation, and other fields. They constructed mathematical models for solving integrated crew pairing and personalized crew assignment problems, which were solved using a combination of Benders decomposition and column generation algorithms. Zeghal and Minoux (2006) developed an integer model that integrates the crew pairing and crew assignment problems, and proposed a heuristic algorithm based on local search trees to solve the model. During this process, a rounding strategy was implemented to shorten the solution time. Souai and Teghem (2009) developed three heuristic algorithms to improve constraint rules in genetic algorithms. Saemi et al. (2022) designed a meta-heuristic algorithm based on ant colony optimization to solve the integrated crew pairing and assignment problem. Based on this, it can be concluded that the proposed algorithm performs well in dealing with the above integrated problem.

So far, some scholars have also studied how to integrate the crew pairing and assignment models based on aircraft recovery problems. At the same time, all the constraints of these two stages have obtained the optimal scheduling plan. Stojković et al. (1998) focused on the crew scheduling problem with a time span of one month and proposed that the flight schedule could be divided into a planning stage and an actual operation stage. The optimization goal in the planning stage was to minimize the total cost. However, when there were deviations in the actual operation stage, a new task cycle needed to be generated to replace the currently unfeasible task cycle arranged in the planning stage. Medard and Sawhney (2007) proposed an integer programming model for the problem of integrated scheduling during aircraft recovery, which considered the qualification of each pilot and the limit of available working time, to ensure the match between the generated scheduling plan and the pilots themselves. Chen and Chou (2017) formulated the aircraft recovery problem as a combinatorial optimization problem containing multiple objectives and constraints. They improved the genetic algorithm to obtain the Pareto solution to the original problem. Evler et al. (2022) proposed a comprehensive recovery model that incorporates aircraft turnover and its related recovery potential into existing comprehensive aircraft, crew and passenger recovery solutions. Teymouri et al. (2023) proposed a novel two-stage robust scenario optimization model in order to solve the problem of comprehensively considering crew scheduling and aircraft maintenance routes. They verified the performance of the model through a numerical example. Considering the complexity of the practical situation, the adaptive large neighborhood search algorithm was used to solve the practical problem effectively.

Other scholars have also considered integrating problems from different stages of crew scheduling and aviation operation management to solve them simultaneously. Cordeau et al. (2001) and Mercie et al. (2005) adopted the Benders decomposition algorithm to solve the integration problem of aircraft maintenance routes and crew pairing. Sandhu and Klabjan (2007) and Gao et al. (2009) studied the correlation between aircraft assignment and crew pairing problems. They considered improving the aircraft purity of each airport during the aircraft assignment, so as to provide more possibilities for detachable flights to improve the robustness of the plan. Dunbar et al. (2014) integrated the aircraft maintenance routing problem and the crew pairing problem considering flexible time. Mohamed et al. (2016) constructed an integrated mixed integer programming model for the integrated aircraft routing and crew pairing problem. They solved the model by using integer linear programming, the Dantzig Wolfe decomposition method and the Benders decomposition method respectively. In addition, a genetic algorithm was employed to generate the feasible path and flight loop in the sub-problem solution. Ahmed et al. (2018) presented an integrated polynomial-sized mixed integer nonlinear programming model, aiming to generate an economical and robust scheduling scheme under the premise of matching aircraft and crew scheduling rules. Parmentier and Meunier (2020) designed a cut generation method to solve the integration problem. Ahmed et al. (2022) constructed a large-scale but compact mixed integer programming model to address an integrated scheduling problem considering fleet allocation, flight routing, and crew pairing. They presented a mathematical heuristic approach that includes a decomposition method and a neighborhood search algorithm to solve the model efficiently. To enhance airlines' operational sustainability and promote the circular economy, Wen et al. (2023) proposed a novel crew pairing problem, which aims to minimize the total basic operating costs, total fuel consumption, greenhouse gas emissions, as well as the robustness cost of generating crew pairings. A solution algorithm based on column generation technology was developed.

2.2 Research status of crew fatigue and flight safety

Scholars in the past have conducted extensive research on various factors causing pilot fatigue, the harm of pilot fatigue and the measures to reduce the degree of fatigue. Gregory et al. (2010) analyzed the sleep status and fatigue levels of pilots based on the investigation results of aviation medicine and proposed that the phenomenon of pilot fatigue is universal. Most pilots showed sleep inertia in the investigation, and fatigue would affect their own flight performance. Powell et al. (2007) focused on the fatigue factors of pilots who perform short-distance flights. They mainly discussed how the total duration of a day, including all flight segments and duty hours, the number of flight segments, take-off airports, and specific working time affect fatigue levels of pilots. Gander et al. (2013) aimed to compare and analyze fatigue levels of pilots performing ultra-long range and long-range flight missions, and the research results showed that the increased sleep time of pilots during ultra-log range flights could effectively alleviate the accumulation of fatigue. Gander et al. (2015) found in their study that the fatigue degree of pilots is highly influenced by sleep duration and circadian rhythm. For example, adequate sleep before take-off can reduce fatigue during flight and improve reaction speed. Lee and Kim (2018) further conducted a systematic analysis and research on the factors that contribute to the risk of flight crew fatigue, verifying that pilot fatigue was influenced by seven independent variables: flight direction, crew scheduling, partnership, flight environment, job allocation, racial differences and hotel environment. Seah et al. (2021) used scales, recorders, and other tools to measure the relationship between sleep, fatigue and daily performance, and developed and tested a flight operation workflow that marked fatigue prone to reduce risk. Sabaner et al. (2022) conducted a cross-sectional survey to assess crew fatigue and sleep problems during the early stages of the COVID-19 pandemic. The study indicated that anxiety, fear, and uncertainty during the COVID-19 pandemic would lead to fatigue and sleep disorders in flight crew. Li et al. (2023) analyzed the relationship between sustained attention and fatigue of flight crew during exempt and non-exempt flights in the COVID-19 pandemic phase. They found that pilot fatigue reached its highest in the early morning. Increasing the number of flight crew members on exempt flights, allowing for more in-flight rest shifts, and taking extra breaks on non-exempt flights may help reduce pilot fatigue and maintain pilots' alertness.

Steiner et al. (2012) also mentioned that the current practices of mitigating fatigue risks by restricting pilot flight time through aviation regulations have proven to be insufficient, highlighting the need for more precise control over crew fatigue levels in aviation operations management. Additionally, they introduced the use of a fatigue risk management system (FRMS) as a continuous monitoring approach to assess fatigue-related risks, ensuring that crew members always maintain a certain level of alertness and enhance flight safety. Honn et al. (2016) also argued that in the aviation industry, in addition to the common shift duration and the rest time between consecutive shifts, flight time and the number of flight segments also had an important impact on the alertness of pilots. Meanwhile, Novak et al. (2020) suggested incorporating biological fatigue calculation models into crew scheduling during the planning stage to mitigate fatigue to some extent. They recommended using fixed scheduling cycles to better manage and reduce the fatigue of pilots.

According to the published papers, research gaps in crew scheduling problems are as follows:

- In the research of integrated crew scheduling problems, most considerations have focused on optimizing costs rather than prioritizing flight safety, which cannot purse more reasonable and fair crew scheduling objectives. As a result, this paper proposes a mixed integer programming model that combines reducing the total scheduling cost with the overall fatigue levels of the crew as optimization goals.

- Although scholars have proposed several integrated models and algorithms for crew scheduling problems, there are still areas for improvement to some extent. For instance, there is a lack of comprehensive consideration for the associated characteristics of crew pairing and crew assignment. Furthermore, due to the limitation of the large scale of integration problems, it is difficult to solve large-scale problems within an acceptable time frame. Based on the above shortcomings, this paper considers the number of duty days in a flight loop and the preference characteristics of crew members. A crew scheduling solution algorithm is designed in this paper based on the column generation algorithm framework.
- Regarding crew fatigue issues, the current literature primarily focuses on qualitative management. Existing research mainly controls the characteristics that can lead to increased fatigue levels of crew members during crew pairings and assignments, aiming to avoid or achieve a balance. However, for multiple types of fatigue characteristics such as night shifts, second take-offs and landings, positioning and so on, it is difficult to accurately assess their respective impacts on fatigue from a qualitative perspective. Therefore, this paper measures the fatigue levels of pilots with different characteristics from both subjective and objective perspectives.

3. Problem description

We introduce the main professional terms involved in the crew scheduling problem.

- Flight segment: the airplane flies from one airport to another airport (without making a stop at a third airport) on its flight.
- Deadheading: the flight of a crew member as a passenger in an aircraft or ground vehicle for the purpose of performing the assigned flight task at the request of the certificate holder, except for his or her transportation to and from suitable local accommodation. Deadheading belongs to duty, and deadheading time cannot be used as rest time.
- Duty: all the tasks performed by the crew members in accordance with the requirements of the certificate holder, such as flight duty, deadheading, backup, training and so on.
- Flight duty period: the time period starting from when crew members report to the designated location for a flight assignment (including flying, deadheading or ferrying), until the engines are shut down after the final flight with no intention for further aircraft movement.
- Connection: the time between two adjacent flight segments within the same duty day.
- Rest period: the continuous time period from when crew members arrive at a suitable accommodation place until they leave the suitable accommodation place for the next mission. During this period, the certificate holder shall not arrange any work or disturb the crew members. The time spent on duty and using transportation to travel between suitable accommodation locations and duty stations for assigned flight tasks shall not be counted as the rest period.
- Placeholder: ground training, duty and other activities already arranged by the crew prior to the planning scheduling.
- Base: the airport where crew members are employed, and each crew member has their own assigned base.
- Duty day: a working day for crew members between two consecutive overnight rests, consisting of a sequential and organized series of multiple flight missions, deadheading tasks, and other duty tasks in terms of time and space. The interval between two consecutive tasks is the crew members' layover rest. The duty day specifies the work content for crew members in a day.
- Pairing: a series combination of consecutive duty days that does not exceed a certain number, starting from a certain personnel base and ultimately returning to the same base.
- Personal schedule: the arrangement of a specified pilot's work and rest schedule within a certain period of time (such as two weeks, one month, etc.), consisting of a series of flight loops, simulator training, duty shifts, and off days.
- Assignment: the allocation of specific flights or crew pairings to pilots and cabin crew.

For the crew scheduling problem studied in this paper, our proposed method consists of two parts: model construction and algorithm solution. In the first part, an improved three-stage model of pilot alertness is first employed to measure alertness and fatigue working duration of pilots. Then, a mixed integer programming model is constructed to describe the crew scheduling problem considering pilot fatigue. In the second part, a solution algorithm based on a column generation framework is developed for solving the proposed mixed integer programming model.

4. Model

In this section, we refer to the existing relevant literature, which examines the effects of different variables on pilot fatigue (Powell et al., 2007; Gander et al., 2015). From an objective perspective, this study primarily considers two aspects: on one hand, the environment of take-off and landing airports, flight type, climate, geographical factors and time period determine the objective flight difficulty of each flight segment. On the other hand, when crew members stay overnight at various rest stations, the rest environment at different rest stations would also bring different benefits or impacts to the recovery of pilot fatigue. Subjectively, in addition to qualification documents and pre-arranged occupying activities, individual differences in sleep patterns, adaptability to the environment, and geographical preferences also have varying effects on the actual fatigue experience and cumulative fatigue tolerance of pilots.

From both objective and subjective perspectives, this research selects factors that have different effects on the actual fatigue feeling and cumulative fatigue endurance of pilots in the planning stage and improves the calculation model of alertness, which quantitatively calculates the fluctuation process of alertness and fatigue levels of different pilots in the implementation of different work plans. Subsequently, considering the pilot fatigue, a mixed integer programming model is constructed to deal with the crew scheduling problem.

4.1 Proposed improved three-stage model of pilot alertness

In the basic alertness calculation model proposed by Åkerstedt and Folkard (1997), the internal environmental stability and the circadian rhythm were mainly considered. Internal environmental stability represents the change process in which a person's alertness decreases gradually while maintaining wakefulness, and then recovers after a period of sleep, denoted by *S*. Biorhythm describes the influence of the biological clock on changes in human alertness, which can be divided into 24-hour and 12-hour biorhythms, denoted by *C* and *U* respectively. The alertness level is calculated by *S*+*C*+*U*. Fig. 1 shows how someone's alertness changes over a 24-hour period from 8 a.m. on day one to 8 a.m. on day two without constraints. Interval [*b, d*] represents the sleep stage, interval [*b, c*] represents deep sleep stage, interval [*c, d*] represents normal sleep stage, and interval [*d, a*] represents wakefulness stage. If a person is unable to fall asleep at point *b* due to a job, the alertness will continue to decrease.

Michael et al. (2014) linearly transformed the level of alertness expressed by Eq. (1) into a measurement standard for sleepiness, known as the Karolinska Sleepiness Scale (KSS), to complete the quantitative measurement of the fatigue level.

$$f = a_f + b_f(S + C + U) \tag{1}$$

where $a_f = 10.6, b_f = -0.6$.

However, the values of each parameter in the three-stage calculation model of basic alertness are fixed. It is assumed that the work intensity remains constant and the recovery rate of fatigue during rest is not influenced by any other factors. In addition, since no specific aircrew is involved in the stage of crew pairing, if the individuals who perceive fatigue are entirely identical, which differs from the aircrew assignment problem studied in this paper. Therefore, this paper improves the three-stage calculation model of alertness, so that the calculation results are more closely related to fatigue levels of pilots with different characteristics. The specific explanations of the adjusted calculation formulas for each stage are as follows.

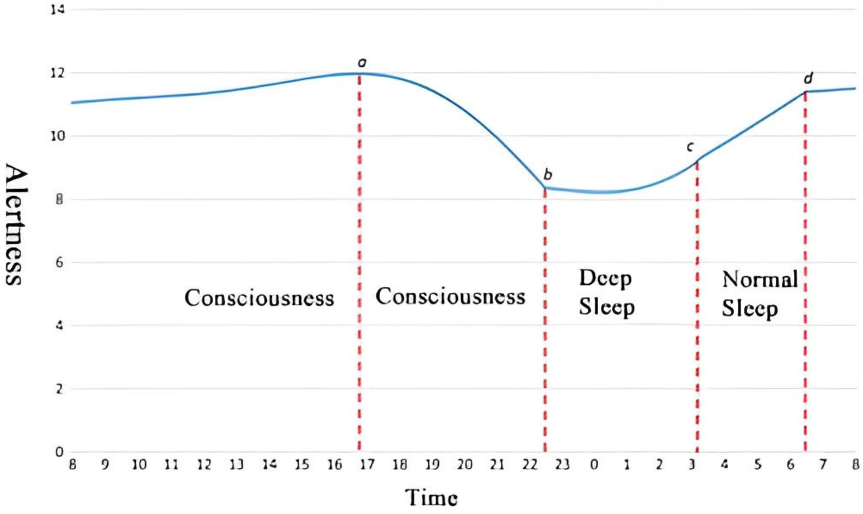

Fig. 1. The process of changes in someone's alertness level over a 24-hour period

4.1.1 Stability of internal environment

In the first part of the formula, the change in the value of *S* reflects the maintenance of homeostasis within the human body. In the waking state, as the value of *S* decreases, it reflects an increase in fatigue and a decrease in alertness. Conversely, in the sleep state, fatigue is restored and then vigilance is increased, leading to a corresponding increase in the value of *S*.

The degree of fatigue during the work phase depends not only on the duration of continuous work, but also on the intensity of the work. Within each duty period, pilots' work can be roughly divided into two parts: the actual flight process of operating flights and the ground duty part, which includes check-in upon arrival, layovers between consecutive flights, and check-out after completion of duty. This study presents that ground duty belongs to the normal work intensity category, while the work intensity of the flight part is affected by various factors, such as the environment of the relevant take-off and landing airports, flight type, climate, geographical factors, and the time. Taking these factors into consideration, the objective flight difficulty coefficient of each flight segment can be evaluated directly. On the one hand, it reflects the difficulty of the work and the requirement for pilots’ concentration. On the other hand, it also corresponds to the speed at which pilots consume alertness and accumulate fatigue.

The recursion formulas for the value of S during flight and duty are shown in Eq. (2) and Eq. (3) respectively. Compared to ground duty, it is considered that the fatigue accumulation speed is faster during the flight. Therefore, the difficulty coefficients of the flight segments are all greater than 1. The greater the difficulty coefficient values, the faster the value of S decreases under the same flight time.

Based on the above situation, the work of pilots can be roughly divided into two categories: flight and ground duty. This article divides each working time segment based on the take-off and landing moments of each flight. For any waking time segment, the value of S will be decreased by a certain amount relative to its initial value at the beginning of that time segment.

$$S(S_w, \Delta Tleg) = la + (S_w - la)e^{(d)(\Delta Tleg)(\alpha)} \tag{2}$$

$$S(S_w, \Delta Twork) = la + (S_w - la)e^{(d)(\Delta Twork)} \tag{3}$$

where d=-0.0353, la = 2.4, ha = 14.3. S_w represents the initial value at the beginning of the flight or duty mission. α represents the flight difficulty factor of the flight segment. $\Delta Tleg$ represents the flight duration of the segment. $\Delta Twork$ represents the operating duration of the selected ground duty segment.

During the rest period, the human body primarily restores its self-regulatory state through sleep. According to the value of S at the initial moment of sleep, which represents the level of fatigue, the human body will enter either the deep sleep state with a high recovery rate or the light sleep state with a gentle recovery rate. The corresponding calculation formulas are shown as Eq. (6) and Eq. (7) respectively. When the value of S is no less than 12.2 at the initial moment, the human body directly enters a shallow sleep state for recovery, and Eq. (7) is used for the recursive calculation. Conversely, when the value of S is less than 12.2 at the initial moment of the human body falling asleep, it will enter the deep sleep state and perform a recursive calculation by applying Eq. (6). According to Eq. (4), the time required to recover to the value of S equal to 12.2 can be obtained clearly. If the total rest time exceeds the critical time, the human body will transform from deep sleep to light sleep at the critical point. With the rest of the time spent in the light sleep state, Eq. (7) is still used for the recursive calculation.

$$bt = \frac{bl - S_s}{(bl - ha)g} \tag{4}$$

$$S'(S_s, \Delta Tasleep) = \begin{cases} S_1'(S_s, \Delta Tasleep), if\,(S_s < bl) \\ S_2'(\Delta Tasleep, bt), otherwise \end{cases} \tag{5}$$

$$S_1'(S_s, \Delta Tasleep) = S_s + g(bl - ha)\Delta Tasleep \tag{6}$$

$$S_2'(\Delta Tasleep, bt) = ha - (ha - bl)e^{(\Delta Tasleep - bt)g} \tag{7}$$

$$g_1 = g_0 \times restLevel \tag{8}$$

$$g_2 = g_0 \times restLevel \times 0.8 \tag{9}$$

where $g_0 = -0.3813$, $ha = 14.3$. $bl = 12.2$ indicates that the value of S at the junction between deep and light sleep states is 12.2. bt represents the duration of sleep required to recover from its initial value when entering the sleep state S_s to bl. g_0 represents the value of the recovery rate in general situations.

In this article, the rest places are not fixed when pilots need to stay out overnight during the work and each rest station provides a different rest environment for pilots. It is believed that different rest places will bring different gains or influences to the recovery of the pilots’ fatigue degree, which is determined by parameters $restLevel$ in the formula. In addition, due to individual differences among pilots, their adaptation to the climate and diet of many stations will also be different. It is roughly assumed that each pilot can submit a preferred site, such as the hometown or a more familiar city, and so on. When resting at the base and preferred place, the recovery rate is calculated by Eq. (8). When resting at other places, the recovery rate is calculated by Eq. (9).

4.1.2 Biological rhythm

Biological rhythm, also known as the biological clock in life, affects the duration of work or rest at different times of the day and may result in varying levels of alertness. Therefore, additional formulas are needed to correct the alertness level at specific times. This part is composed of 24-hour biological rhythm and 12-hour biological rhythm, represented by C and U respectively in Eq. (10) and Eq. (11). This study also considers that pilots may have different work and rest rules, roughly divided into morning types who habitually wake up early and evening types who prefer staying up late. Different parameter values are used to distinguish between these two types when calculating their biological rhythms.

$$C(A_t,T_t)=C_m+C_a\cos\left[\frac{2\pi}{24}(T_t-p-A_t)\right] \tag{10}$$

$$U(A_t,T_t)=U_m+U_a\cos\left[\frac{2\pi}{12}(T_t-p-A_t-3)\right] \tag{11}$$

where $C_m=0,C_a=2.5$, $U_m=-0.5,U_a=0.5$. If it is a morning person who is accustomed to waking up early, the parameter value $p=16.8$. If it is an evening person who is accustomed to staying up late, the parameter value $p=20.8$. The value of T_t is to convert a specific moment into the corresponding hour value. For example, 16:30 would be represented as 16.5. In addition, pilots often need to cross multiple time zones when performing flight tasks, and the jet lag will also have a certain impact on their fatigue degree, that is, the A_t in the calculation formulas for C and U mentioned above. The recursive formula is shown as Eq. (12). TZ represents the time zone difference between t and $t-1$. For example, when a pilot performs a flight from Beijing to Tokyo, $TZ=1$.

$$A_t(TZ\ ,A_{t-1})=A_{t-1}+\left[1-(daily)^{T_t-T_{t-1}}\right](TZ\ -A_{t-1}) \tag{12}$$

where $daily = 0.5$.

4.1.3 Calculation of alertness level and fatigue working duration at each moment

Since the alertness level affected by self-regulation depends on the state of the crew (awake, deep sleep or normal sleep), and crew fatigue will vary with the different states of crew members. When calculating crew fatigue, it is necessary to determine the current state of the crew members. Therefore, crew fatigue needs to be calculated sequentially within smaller time intervals. According to pilots' work or rest state, this study divides the content of individual scheduling into three categories: flight, duty, and rest. Whenever there is a transition between these categories, calculations need to be made regarding pilots' alertness levels.

Taking the start time of the first task in the personal scheduling timetable arranged within the planning period as the initial moment for alertness propagation, it is assumed that all pilots begin their work in optimal condition at the beginning of the planning period. According to the conclusion of Yildiz et al. (2017), in general, when alertness level $S+C+U \le 8.38$, it indicates that alertness level has reached a critical value, and the human body enters a state of fatigue and needs to rest. When alertness level $S+C+U \ge 11.38$, people will wake up. The rules for setting the initial values of each part in the formula for calculating alertness at the initial moment T_0 are shown in Eqs. (13-16) respectively.

$$A_{t0}=0 \tag{13}$$

$$C_0=C_m+C_a\cos\left[\frac{2\pi}{24}(T_0-p-A_{t0})\right] \tag{14}$$

$$U_0=U_m+U_a\cos\left[\frac{2\pi}{12}(T_0-p-A_{t0}-3)\right] \tag{15}$$

$$S_0=11.38-C_0-U_0 \tag{16}$$

Afterwards, the components of the scheduling timetable are iterated successively. Whenever the pilot switches between flying, duty or rest states, the values for internal environmental stability, 24-hour biorhythm and 12-hour biorhythm at the current moment are calculated using the recursive formulas introduced in Section 4.1.1 and Section 4.1.2. The recursion calculation is performed successively until the scheduling timetable is completed.

Similarly, based on the conclusion that people will wake up when alertness level $S+C+U \ge 11.38$, this study believes that the rest period partially includes but is not entirely equivalent to the recovery time of sleep. That is, when the duration of rest is sufficient for pilots to recover to the optimal state, $S+C+U = 11.38$, it is considered that the S value will not continue to increase during the remaining rest time. Furthermore, at the end of the rest period, $S = 11.38-C-U$, it can be regarded as the rest (non-sleep) state where pilots can choose to enter the sleep state at any time after consumption to recover to their optimal state and start their next work segment.

Additionally, it is worth noting that pilots often unavoidably must work for a certain duration in a fatigued state when performing high-intensity, high-density flight tasks. The duration of working in a fatigued state is also an important factor that cannot be overlooked about flight safety. Therefore, based on the calculated alertness levels of pilots at each moment within the planning period, this study estimates the total duration of their work while fatigued. For simplicity in the calculation, it is approximated that there is a linear trend in alertness levels between consecutive moments. Hence, the calculation method for the total fatigue working duration is shown as Eqs. (17-19).

$$\text{If } Z_{S+C+U}^{t}\le 8.38 \text{ and } Z_{S+C+U}^{t-1}\le 8.38\text{ , } T_{f1}=(T_t-T_{t-1}) \tag{17}$$

$$\text{If } Z_{S+C+U}^{t} \le 8.38 \text{ and } Z_{S+C+U}^{t-1} > 8.38,\ T_{f2} = \frac{8.38 - Z_{S+C+U}^{t}}{Z_{S+C+U}^{t-1} - Z_{S+C+U}^{t}}(T_t - T_{t-1}) \tag{18}$$

$$\text{Total fatigue working duration} = \sum T_f \tag{19}$$

4.2 Construction of a crew scheduling model considering pilot fatigue

Based on the ensemble segmentation model of the personnel assignment problem, this subsection constructs the mixed integer model by considering pilot fatigue. Below, we first define sets, parameters and decision variables involved, and then introduce the relevant meanings of the objective function and constraints of the model.

4.2.1 Definition of notations

To make the constructed model easier to understand, we first define all sets, parameters, and decision variables used in the problem formulation, as shown in Table 1.

Table 1
Definition of notations

Notations	Detailed definitions
Sets	
R	Set of working schedules (indexed with r)
Ω	Set of flight loops (indexed with i)
P	Set of pilots (indexed with p)
R_p	Set of optional working schedules for pilot p
Parameters	
t_r	Fatigue duration for pilots executing working schedule r
c_r	Total cost of individual working on schedule r
π	Target weight adjustment parameters
a_i^r	1, if flight loop i is in the working schedule r , otherwise 0
Decision variables	
x_r	1, if the working schedule r is in the scheme, otherwise 0

4.2.2 Model construction

$$\min \sum_{r\in R} (\pi \cdot t_r + c_r) x_r \tag{20}$$

subject to

$$\sum_{r\in R} a_i^r x_r = 1 \qquad \forall i \in \Omega \tag{21}$$

$$\sum_{r\in R_p} x_r = 1 \qquad \forall p \in P \tag{22}$$

$$x_r \in \{0,1\} \qquad \forall r \in R \tag{23}$$

The objective function (20) is mainly composed of two parts, which respectively represent two optimization objectives: reducing the manpower cost of pilots required for airline operations and minimizing the total fatigue working duration of crew members. To begin with, although the physical fitness of each pilot has been considered when generating a feasible work schedule, pilot fatigue can still have a detrimental impact on the flight safety, since pilots continue to work while fatigued. In addition, during the planning period, the fatigue working duration is considered as a major factor in determining the overall workload intensity for pilots, in addition to the total working hours. Therefore, decreasing the total working hours of all pilots under fatigue can reduce the risk of fatigue-induced errors. Subsequently, hiring and training pilots is a major part of an airline company's labor costs. To improve its economic efficiency, it is necessary to simultaneously reduce these costs, instead of blindly reducing the fatigue working duration. Among them, the parameter π is used to balance the weight of two optimization objectives.

Eq. (21) represents the set covering constraint, which ensures that all flight loops $\forall i \in \Omega$ are included in at least one selected individual schedule obtained from solving the model. The number of such constraints is equal to the total number of flight loops. Eq. (22) refers to personnel assignment constraints. In the final scheduling scheme, each pilot can only be assigned one schedule. The number of such constraints is equal to the total number of pilots. Finally, Eq. (23) represents the value constraint for variables. The model coefficients are presented in Table 2.

Table 2
Example of model coefficients

Constraints		x_1	x_2	x_3	$\cdots$	x_r	b
Flight loop coverage constraints	Ω_1	1	0	1	$\cdots$	0	1
	Ω_2	0	1	1	$\cdots$	0	1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	1
	Ω_i	1	1	0	$\cdots$	1	1
Pilot assignment constraints	p_1	1	0	0	$\cdots$	0	1
	p_2	0	1	0	$\cdots$	0	1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	1
	p_j	0	0	0	$\cdots$	1	1

5. Model solution based on the column generation algorithm

The number of feasible scheduling plans rises significantly as the number of flight loops increases, and even once the issue reaches a certain scale, it is nearly impossible to list all suitable scheduling timetables. Therefore, although theoretically using the simplex method can guarantee obtaining the solution of integer programming problems after a finite number of iterations, it is not applicable when facing large-scale problems where it is impossible to obtain the set of all variables or select entering and leaving basic variables from a huge number of variables. As a result, for the crew scheduling problem studied in this paper, a solution algorithm based on the column generation framework is developed as follows: construct an initial master problem model, after each round of solving the master problem, solve the scheduling path based on the dual price of each resource to generate sub-problems, and add additional variables with a reduced cost smaller than 0 to the master problem model. Repeat this process iteratively until no better scheduling path can be obtained. At this point, it is considered that the linear programming master problem has reached its optimal solution. Then restore the 0-1 integer constraints on variables and obtain an integer solution.

5.1 Master problem

Remove the integer constraints on variables and relax the master problem into a linear programming problem. To ensure the existence of an initial feasible solution, make the following adjustments to the original problem model: add $|\Omega|$ artificial variables y_i for all flight loops, with coefficients in their column vectors set to 1 only in corresponding task cycle coverage constraints and 0 elsewhere. The coefficient c_i in the objective function represents the penalty cost for unassigned flight loops. Similarly, add $|P|$ artificial variables for all pilots, with coefficients in their column vectors set to 1 only in corresponding pilot coverage constraints. The coefficient c_p in the objective function represents the cost for idle pilots. When optimizing towards minimizing human resource usage, this cost coefficient is negative. When optimizing towards evenly utilizing existing pilots, it is positive. The master problem model is shown as follows:

$$\min \sum_{r\in R}(\pi \cdot t_r + c_r)x_r + c_i y_i + c_p y_p \tag{24}$$

subject to

$$\sum_{r\in R} a_i^r x_r + y_i = 1 \qquad \forall i \in \Omega \tag{25}$$

$$\sum_{r\in R_p} x_r + y_p = 1 \qquad \forall p \in P \tag{26}$$

$$0 \le y_i \le 1 \qquad \forall i \in \Omega \tag{27}$$

$$0 \le y_p \le 1 \qquad \forall p \in P \tag{28}$$

$$0 \le x_r \le 1 \qquad \forall r \in R \tag{29}$$

An example of the master problem after introducing artificial variable relaxation is shown in Table 3.

Table 3
Example of coefficients in the master problem model

Constraints		y_{Ω_1}	$\cdots$	c_{Ω_i}	y_{p_1}	$\cdots$	y_{p_j}	x_1	$\cdots$	x_r	b
		c_{Ω_1}	$\cdots$	c_{Ω_i}	c_{p_1}	$\cdots$	c_{p_j}	$\pi\cdot t_1+c_1$	$\cdots$	$\pi\cdot t_r+c_r$	
Flight loop coverage constraints	Ω_1	1	$\cdots$	0	0	$\cdots$	0	1	$\cdots$	0	1
	Ω_2	0	$\cdots$	0	0	$\cdots$	0	1	$\cdots$	1	1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	1
	Ω_i	0	$\cdots$	1	0	$\cdots$	0	0	$\cdots$	1	1
Pilot assignment constraints	p_1	0	$\cdots$	0	1	$\cdots$	0	1	$\cdots$	0	1
	p_2	0	$\cdots$	0	0	$\cdots$	0	0	$\cdots$	1	1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	1
	p_j	0	$\cdots$	0	0	$\cdots$	1	0	$\cdots$	0	1

5.2 Sub-problem

In the sub-problem, it is essential to identify an entering basic variable that makes the function viable for further optimization. In the column generation algorithm, the criterion for determining whether a variable is more optimal depends on the size of its reduced cost. Specifically, in the model of the minimization problem constructed in this article, when the reduced cost is less than 0, it indicates that using the newly generated work schedule for assignment can reduce the objective function value. At this point, its corresponding variable is added to the model of the master problem. The dual variable corresponding to the flight loop coverage constraint (24) is denoted as η_i, and the dual variable corresponding to the personnel assignment constraint (25) is denoted as *βp*. The formula for calculating the number of sub-problem tests is shown in Eq. (30):

$$(c_r+\pi\cdot t_r)-\sum_{i\in\Omega_r}\eta_i a_{ir}-\beta_p \tag{30}$$

where its own total cost $(c_r+\pi\cdot t_r)$ is the increased part of the value of the objective function, $\sum_{i\in\Omega_r}\eta_i a_{ir}$ represents the sum of the dual price of all flight rings in the scheduling timetable *i*, and together with personnel dual price β_p, they constitute the part that can save the current target value. When Eq. (30) takes a value less than 0, the absolute value is also equal to the actual gain that the individual can bring to the master problem at this time.

The sub-problem aims to generate a series path of flight loops for each pilot that satisfies the condition of having a reduced cost less than 0. Therefore, it is necessary to solve $|p|$ sub-problems in each iteration of the generated column.

Firstly, the flight loop connection network is constructed for each pilot, as shown in Fig. 2.

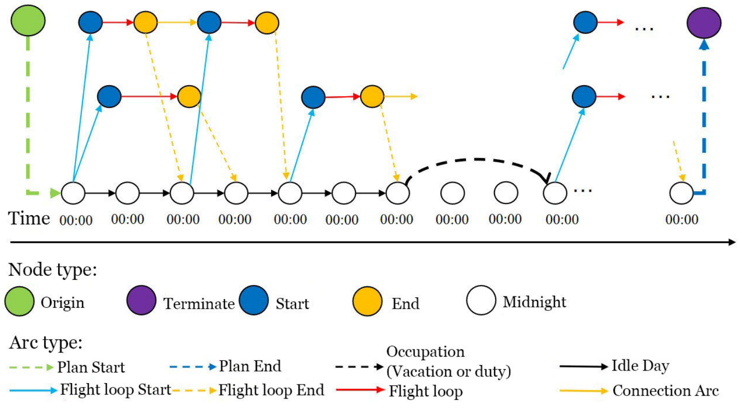

Fig. 2. The connection network of flight loops

The set of all points in the figure is denoted as N^p, indexed by n, including a virtual starting point, a virtual ending point, a series of connection nodes, representing the beginning and end of activities such as flight loops or placeholder tasks, as well as the time node representing midnight. The virtual starting point is connected to the time node corresponding to the start time of the planning period. Similarly, the virtual ending point is only connected to the time node corresponding to the end time of the planning period. There are eight types of arcs in the network diagram, which are arcs indicating plan start, plan end, pre-arranged space placeholder activities of the pilot such as vacation or duty, idle day arcs, arcs indicating flight loop start and end, flight loop arcs and arcs connecting consecutive flight loops or space placeholder tasks. After rearranging the calculation formula of the reduced cost, Eq. (31) can be obtained. Therefore, when constructing the network diagram, the weights of the flight arc and end arc can be defined as $(c_i-\eta_i)$ and $-\beta_p$ respectively. However, based on the calculation method of pilots' alertness level at each time node and total fatigue working duration introduced in Section 4.1.3, the fatigue cost t_r cannot simply be split into the weight of arcs or nodes in the network diagram. Therefore, when this sub-problem is converted into finding the shortest path between the virtual start point and virtual end point, it is also necessary to add the fatigue cost part according to the task composition of the entire path.

$$\pi \cdot t_r + \sum_{i\in\Omega_r}(c_i-\eta_i)a_{ir} - \beta_p \tag{31}$$

The impact of factors such as the flight difficulty, rest environment at overnight stations and pilots' adaptation level, as well as their own work and rest patterns, mentioned in the problem analysis has been reflected through parameter adjustments in the three-stage alertness calculation model described in Section 4.1. Additionally, the pilots with varying physical fitness have different tolerances and durations for fatigue work states. Therefore, this paper considers the cumulative fatigue working duration for each pilot as a class of resources that should not exceed its upper limit value. Thus, this sub-problem can also be solved by analogy to the shortest path problem with resource constraints (SPPRC) problem.

Although the connection network is a directed acyclic graph (DAG), in addition to considering the total cost of the path, it is also necessary to simultaneously consider the cumulative fatigue working duration of pilots and determine whether it complies with crew scheduling rules during the construction of the path. Therefore, this pricing sub-problem does not satisfy the conditions of the classic DAG shortest path algorithm. For the crew scheduling problem considering pilot fatigue in this study, we have made improvements to the classical DAG algorithm as follows: designing a data structure which stores four items including pre-order node, cumulative fatigue working duration, the size of the reduced cost and the end time of the last long break. In addition to comparing the total cost of the path represented by the size of the reduced cost, the cumulative fatigue working duration must also be lower than the pilot's accepted limit for the path to be deemed legal. Furthermore, the R6 version of the Rules for the Operation of Carriers of Public Air Transport for Large Aircraft requires that pilots acquire at least 48 hours of consecutive rest time within 144 hours before any flight departure time. However, since the input flight loop may not always span 4 days, there may be several consecutive flight loops in a series before the pilot can obtain their necessary long rest at their base. Thus, even if flight loops can be connected, the end time of the last long rest will determine if subsequent flight loops can be added to the existing path. If parts of paths with higher reduced costs but later long rest end times are deleted too early, it may reduce the possibilities of connecting paths and result in failure to obtain the optimal solution.

Firstly, arrange all nodes chronologically and assign initial labels to each of them. The start node is $<NIL, 0, 0, bgnTime_{n_0}>$, while other nodes are labeled as $<NIL, \infty, \infty, bgnTime_n>$. Here, $bgnTime_n$ denotes the start time of node n, indicating that any path is assumed to have just finished a long rest before starting. Then, the label set for each node is updated in a topological order: If the newly generated label does not dominate any existing labels of the node, the new label is added to the set and all the existing labels that are dominated by the new one is removed. The pseudo-code for extending the RELAX function in the original classical DAG algorithm is shown in Algorithm 1. Additionally, Algorithm 2 called DOMINATE is included, which aims to determine whether one label is superior to another.

Algorithm 1. MULTILABEL-RELAX (*i*, *j*, *w*)

```
for each $l \in i.labelList$;// Traverse the list of labels for node i;
  label $a = Connect(l, j)$;// Construct label $a$ after concatenating node j based on the information in label $l$;
  for each $b \in j.labelList$;// Traverse the existing label set of all nodes $j$;
    mark = unknown;// Mark the status of label $a$ as unknown;
    if DOMINATE $(a,b)$;// Label $a$ is superior to label $b$;
      delete $b$ from $j.labelList$;// Delete $b$ from the label list of $j$;
      mark = good;// Mark the status of $a$ as good;
    if DOMINATE $(b,a)$;// Label $b$ is superior to label $a$;
      mark = bad;// Mark the status of $a$ as bad;
```

```
if mark is not bad;// If $a$ is marked not bad;
    insert $a$ into $j.labelList$ ;// Insert $a$ into the label list of $j$ ;
```

An algorithm DOMINATE (a,b) is designed to determine whether label a is superior to label b . The pseudo-code is shown in Algorithm 2.

Algorithm 2. DOMINATE (a,b)

```
if $a.endT \leq b.endT$ and $a.t \leq b.t$
    and $a.\cos t \leq b.\cos t$
    and $!(a.endT = b.endT$ and $a.t = b.t$
    and $a.\cos t = b.\cos t)$ ;// None of the indicators of label $a$ is inferior to label $b$ and they are not exactly equal;
    return true;// Return true, indicating that label $a$ is superior to label $b$ ;
return false;// Return false, indicating that label $a$ is inferior to label $b$ ;
```

After completing the search for all nodes, the end point can obtain a label that records the minimum reduced cost. The optimal scheduling path can be generated by tracing the preceding nodes in succession. If the reduced cost of the shortest path is less than 0, its corresponding variable will be added to the master problem.

When the minimum reduced cost obtained from solving the task network model for all pilots does not meet the requirement of having a reduced cost less than 0, the iterative process stops. Finally, based on the current master problem model, the integer constraints to variables are restored in order to solve the integer programming problem.

6. Experiment results and sensitivity analysis

6.1 Experiment data

In order to verify the feasibility of our proposed method, the real dataset obtained from an airline company is used for numerical experiments. This dataset includes flight information, airport details, aircraft details, personnel information, seating details, and placeholder activity information. The flight information contains the flight ID, flight date, flight number, domestic/international indicator, departure airport, arrival airport, departure time, arrival time, tail number, and flight time. The airport information includes the three-character code, four-character code, Chinese name, domestic/international indicator, and overnight stay indicator. Aircraft information includes the tail number and type of aircraft. Personnel information includes the personnel ID and personnel base, and the technical authorization level. Seat placement is divided into two categories: flight placement and ground positioning. Flight placement includes the same fields as the flight information, along with additional flights from other airline companies and some commonly used airlines. Ground positioning includes the departure time, arrival time, departure airport, and arrival airport. The placeholder activity information is composed of the pre-arranged duties and rests, including the personnel ID, placeholder airport, start time, end time, and whether it belongs to rest. The statistical summary of the dataset can be presented in Table 4.

Table 4
The statistical summary of the dataset

Parameters	Values
Schedule start date	2020/07/01
Planned end date	2020/07/21
Number of flights	1351
Number of airports	48
Number of overnight stops available	14
Number of aircraft	11
Number of personnel	68
Number of flight seats	5589
Number of ground seats	432
Number of placeholder activities	60

6.2 Analysis of experimental results

To compare subsequent plans, the scheduling of test data is initially optimized solely based on minimizing the total cost without considering pilot fatigue. Subsequently, the improved three-stage calculation model for pilot alertness introduced in this paper is employed to recursively calculate the fatigue levels at each state transition point for pilots in this scheduling plan. Lastly, the cumulative fatigue working duration of each pilot and the maximum fatigue level throughout the entire planning period are analyzed, as shown in Table 5.

Table 5
Statistical analysis of fatigue indicators for the basic plan

Indicators	Values
Minimum value of the highest fatigue level	3
Maximum value of the highest fatigue level	7
Number of pilots with a fatigue level greater than or equal to 5	65
Number of pilots with a fatigue level greater than or equal to 7	51
Minimum cumulative fatigue duration (hours)	0
Maximum cumulative fatigue duration (hours)	23.0525
Number of pilots with the cumulative fatigue duration greater than 6 hours	41
Number of pilots with the cumulative fatigue duration greater than 10 hours	27

In real-life scenarios, some pilots who are older or concerned by airlines should not be in a state of fatigue work, while young and physically fit pilots can undertake more tasks that exacerbate fatigue. It is assumed that the upper limit of cumulative fatigue work hours that can be accepted during the planning period can be determined by considering the individual needs of pilots and the assessment results of their physical fitness. For example, the cumulative maximum allowable duration of fatigued work for older pilots or those concerned by airlines could be set at 0, that is, they are not allowed to perform flight tasks while fatigued. A simulated dataset is generated to determine the fatigue limits for each pilot. For the upper limit of the fatigue working duration, we will try to compare it with soft constraints and hard constraints in algorithms. Soft constraints involve assigning additional penalty costs during the solving process for scheduling time exceeding pilots predetermined acceptable limit of fatigue work hours, while hard constraints involve adding corresponding rule judgments during sub-problem search for feasible scheduling paths and deeming paths with excessive fatigue durations as infeasible. The statistical results of various indicators in the final scheduling plans corresponding to these two approaches are shown in Table 6 and Table 7 respectively, with a comparison presented in Table 8. As evidenced by Tables 5-8, when dealing with the limitations on pilots' fatigue working duration in the form of soft constraints, compared to the basic plan, the number of pilots with the cumulative fatigue working duration greater than 6 and 10 hours diminishes to some extent. The maximum cumulative fatigue working duration also decreases from 23.0525 hours in the basic plan to 19.5694 hours, and the number of pilots exceeding the upper limit of fatigue working duration decreases from 8 to 1. Additionally, there is a decrease in the number of pilots with a maximum fatigue level equal to or greater than 7 during their planning period. However, when handling this rule with hard constraints, although it ensures that no pilot exceeds the restriction on cumulative fatigue working duration, it does not optimize overall crew fatigue indicators as significantly as that with soft constraints and incurs the higher crew cost. This is possibly because when hard constraints are imposed, algorithms must allocate flight tasks to other pilots to satisfy some pilots' special requirements, leading to the poor global optimization effect.

In conclusion, the basic plan that only considers the cost of scheduling exceeds the maximum acceptable fatigue working duration for some pilots. The existing scheduling rules do not entirely avoid pilot fatigue by controlling flight, duty, and rest time. Moreover, the implementation of hard constraints will reduce a lot of possibilities for scheduling. Therefore, to balance both the cost of flight crew scheduling and the control of fatigue, it is more reasonable to handle them with soft constraints.

Table 6
Statistical analysis of fatigue indicators with soft constraints on fatigue working duration

Indicators	Values
Minimum value of the highest fatigue level	3
Maximum value of the highest fatigue level	7
Number of pilots with a fatigue level greater than or equal to 5	59
Number of pilots with a fatigue level greater than or equal to 7	44
Minimum cumulative fatigue duration (hours)	0
Maximum cumulative fatigue duration (hours)	19.5694
Number of pilots with the cumulative fatigue duration greater than 6 hours	38
Number of pilots with the cumulative fatigue duration greater than 10 hours	25

Table 7
Statistical analysis of fatigue indicators with hard constraints on fatigue working duration

Indicators	Values
Minimum value of the highest fatigue level	3
Maximum value of the highest fatigue level	7
Number of pilots with a fatigue level greater than or equal to 5	58
Number of pilots with a fatigue level greater than or equal to 7	47
Minimum cumulative fatigue duration (hours)	0
Maximum cumulative fatigue duration (hours)	19.1372
Number of pilots with the cumulative fatigue duration greater than 6 hours	35
Number of pilots with the cumulative fatigue duration greater than 10 hours	26

Table 8
Comparison of indicators

Indicators	Basic plan	Soft constraints on fatigue working duration	Hard constraints on fatigue working duration
Crew cost	142973	147272.4	147820.4
Sum of the highest fatigue levels	448	423	425
Total fatigue working duration	530.915	522.255	525.636
Number of pilots exceeding the maximum fatigue working duration	8	1	0
Flight coverage rate	98.51%	98.15%	98.15%
Number of flight days	579	578	578
Average flight duration on flight days	4.62	4.62	4.62
Number of duty days	609	609	610
Average flight duration on duty days	4.39	4.39	4.38

6.3 Sensitivity analysis experiment

In order to control the level of pilot fatigue in the overall scheduling plan, we will test the model constructed in Section 4.2 and evaluate the performance of algorithms designed in Section 5.2. To further analyze the impact of fatigue relative to the crew cost on the final scheduling plan, different values for parameter π are set for sensitivity analysis through multiple iterations and experiments. The statistical results of each group experiment include total crew costs, crew fatigue levels, flight coverage rate, number of flight days, average daily flight hours on flight days, number of duty days and average daily flight hours on duty days are presented in Table 9. Among them, the crew cost and crew fatigue levels displayed represent percentage differences compared to a control group experiment ($\pi=0$) that only considers minimizing total costs.

Table 9
Experimental results of different parameters π using cumulative fatigue hours as an evaluation indicator

π	Crew cost (%)	Crew fatigue (%)	Flight coverage rate (%)	Flight days	Average flight hours per flight day	Duty days	Average flight hours per duty day
0	+0.00	+0.00	98.51	579	4.62	609	4.39
10	+3.50	-0.73	98.08	578	4.62	606	4.41
50	+0.62	-0.85	98.44	578	4.62	609	4.38
100	+2.33	-0.71	98.22	578	4.62	609	4.39
150	+2.85	-0.48	98.15	579	4.63	606	4.42
200	+0.70	-0.09	98.44	579	4.62	610	4.38
250	+2.72	-2.77	98.01	573	4.62	604	4.39
300	+2.31	-3.05	98.08	574	4.62	606	4.38
350	+4.66	-3.27	97.80	572	4.63	603	4.39
400	+5.76	-3.00	97.65	574	4.63	602	4.41
450	+12.22	-7.89	96.87	562	4.67	590	4.45
500	+18.52	-7.24	96.09	554	4.66	583	4.43
Minimum	+0.62	-7.89	96.09	-	4.62	-	4.38
Maximum	+18.52	-0.09	98.44	-	4.67	-	4.45
Average	+5.11	-2.74	97.80	-	4.63	-	4.40

By observing Table 9, it can be seen that when the parameter value is small (e.g. 10-200), the total cumulative fatigue working duration for pilots reduces slightly, but there is no clear trend in the increase of crew costs and decrease in crew fatigue. The changes in the average daily flight hours on flight days and average daily flight hours on duty days corresponding to the unit utilization rate remain within the expected range. However, as the parameter value continues to increase, in order to minimize penalty costs caused by crew fatigue, the model and algorithms sacrifice flight coverage rate during the optimization process so as to allow pilots to complete flight and duty tasks without fatigue. Therefore, this paper believes that by properly controlling parameter π , the overall fatigue levels of the aircrew can be limited within an acceptable range, without excessively increasing the aircrew cost.

6.4 Managerial insights

The Civil Aviation Administration's March 2021 R7 version of the regulations for the Certification of Carrier Operation Qualification of Large Aircraft Public Air Transport places a strong emphasis on the establishment of a fatigue risk management system for each airline company. The importance attached to crew fatigue is increasing. Nonetheless, the restrictions of the original scheduling rules can limit the space available to manage fatigue, leading to a decrease in the utilization rate of crew resources or a lack of crew resources to guarantee the existing production adequately.

In the experiment mentioned above, hard constraints on the fatigue working duration can be seen to handle scheduling manually. Typically, the scheduler first determines the scheduling plan for pilots with special needs, such as poor physical condition, medical attention, or additional flight time restrictions, and selects those that meet all restrictions from a given set of task cycles to construct work schedules for these pilots. Subsequently, assignments between the remaining pilots and the remaining task cycles are completed. This further division of personnel and task cycle subsets will further reduce the global optimality of the results. Additionally, the fixed task cycles mean that adjustments cannot be made at smaller units during scheduling, leading to a poorer overall quality of scheduling results to meet fatigue requirements.

Therefore, from the perspective of balancing production operations and crew fatigue management, our proposed method integrates the quantitative measurements of crew fatigue with crew scheduling considerations. The proposed method can effectively improve the state of crew fatigue by considering each crew member as an individual with various attributes and assigning different flight tasks according to those attributes. During the planning stage, levels of crew fatigue should be accurately monitored to reduce the likelihood of fatigue-related issues. At the same time, it ensures control over operating costs while reducing fatigue levels without sacrificing too many cost items.

7. Conclusions

From the perspectives of cost reduction and aviation safety improvement, this paper studies the optimization of flight crew scheduling considering pilot fatigue. On the one hand, by fully utilizing the available time of crew members and improving their utilization rate, it is possible to reduce the number of required pilots to some extent, lower labor costs, and enhance efficiency. On the other hand, if flight tasks are not arranged reasonably, the accumulated fatigue from long-term high-intensity work can pose a certain threat to aviation safety. Based on the literature review in Section 2, it is understood that existing aviation management regulations' limitations on flight duration and duty hours are insufficient for controlling pilot fatigue. In this paper, we quantitatively consider the impact of assigning various flight and duty tasks on each pilot's level of fatigue during the planning stage. On this basis, we propose a method for solving the crew scheduling problem considering pilot fatigue.

Firstly, the three-stage calculation model for basic alertness is improved to quantitatively calculate the alertness level of pilots at various time nodes. By reflecting the level of fatigue through alertness, the fatigue work duration of pilots with different characteristics during work is monitored to avoid pilots being assigned to work schedules that exceed their physical fitness to withstand.

Then, a mixed integer programming model is established to represent the crew scheduling problem considering pilot fatigue. Indicator variables are set in the model to determine whether the work schedule is assigned. Under the constraints of covering all task cycles and each pilot being able to execute at most one work schedule, the optimization objective is to minimize the total fatigue working duration for all pilots and the manpower cost incurred by pilots. Considering that this problem falls into a large-scale category in the practical situation, it is nearly impossible to directly enumerate all feasible work schedules for its resolution. Therefore, during the solution process, the solution algorithm based on column generation decomposes the original problem into a master problem involving only partial variables and several sub-problems for finding the shortest path.

Finally, the feasibility of our proposed method is verified through the numerical example. The performance of the proposed solution in this paper is compared with that of the assignment plan without considering pilot fatigue, further demonstrating the effectiveness of the proposed solution. Through multiple experimental comparisons and analyses, it is more reasonable to handle pilot fatigue working duration with soft constraints. Sensitivity analysis is conducted, revealing the variation rules of the crew cost and fatigue. Based on the research work in this paper, the following conclusion can be drawn: in solving the crew assignment problem, precise control over individual differences in pilots' fatigue levels during their work period is achievable, which has profound significance for society, airlines and pilots themselves.

The method provided in this article also has certain limitations. For augmented crew, the number of flight crew members scheduled exceeds the minimum requirement for operating the aircraft type. This allows for the substitution of a flight crew member with another qualified crew member, while the replaced crew member can have rest onboard during the flight. The recovery of alertness during the onboard rest period is closely related to the onboard rest facilities and rest patterns, which has not been considered in the proposed method. We will make improvements in future research to ensure that the improved method is applicable to a wider range of scenarios for the crew scheduling problem considering pilot fatigue.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 72171170), the Fundamental Research Funds of the Central Universities (No. 22120210535, 2023-6-YB-02).

References

Ahmed, M. B., Hryhoryeva, M., Hvattum, L. M., & Haouari, M. (2022). A matheuristic for the robust integrated airline fleet assignment, aircraft routing, and crew pairing problem. *Computers & Operations Research*, *137*, 105551.

Ahmed, M. B., Mansour, F. Z., & Haouari, M. (2018). Robust integrated maintenance aircraft routing and crew pairing. *Journal of Air Transport Management*, *73*, 15-31.

Åkerstedt, T., & Folkard, S. (1997). The three-process model of alertness and its extension to performance, sleep latency, and sleep length. *Chronobiology International, 14*(2), 115-123.

Arsintescu, L., Pradhan, S., Chachad, R. G., Gregory, K. B., Mulligan, J. B., & Flynn-Evans, E. E. (2021). Early starts and late finishes both reduce alertness and performance among short-haul airline pilots. *Journal of Sleep Research*, *31*(3), e13521.

Bennett, S. A. (2019). Pilot workload and fatigue on four intra-European routes: a 12-month mixed-methods evaluation. *Journal of Risk Research*, *22*(8), 983-1003.

Chen, C.H., & Chou, J. H. (2017). Multiobjective optimization of airline crew roster recovery problems under disruption conditions. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *47*(1), 133-144.

Cordeau, J. F., Stojkovic, G., Soumis, F., & Desrosiers, J. (2001). Benders decomposition for simultaneous aircraft routing and crew scheduling. *Transportation Science*, *35*(4), 375-388.

Deng, B., Ding, R., Li, J., Huang, J., Tang, K., & Li, W. (2023). Hybrid multi-objective metaheuristic algorithms for solving airline crew rostering problem with qualification and language. *Mathematical Biosciences and Engineering*, *20*(1), 1460-1487.

Dijk, D. J., & Czeisler, C. A. (1995). Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. *Journal of Neuroscience*, *15*(5), 3526-3538.

Dunbar, M., Froyland, G., & Wu, C. -L. (2014). An integrated scenario-based approach for robust aircraft routing, crew pairing and re-timing. *Computers & Operations Research*, *45*, 68-86.

Evler, J., Lindner, M., Fricke, H., & Schultz, M. (2022). Integration of turnaround and aircraft recovery to mitigate delay propagation in airline networks. *Computers and Operations Research*, *138*, 105602.

Gander, P. H., Mulrine, H. M., Van den Berg, M. J., Smith, A. A. T., Signal, T. L., Wu, L. J., & Belenky, G. (2015). Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. *Journal of Sleep Research*, *24*(1), 110-119.

Gander, P. H., Signal, T. L., Van den Berg, M. J., Mulrine, H. M., Jay, S. M., & Mangie, J. (2013). In-flight sleep, pilot fatigue and Psychomotor Vigilance Task performance on ultra-long range versus long range flights. *Journal of Sleep Research*, *22*(6), 697-706.

Gao, C., Johnson, E., & Smith, B. (2009). Integrated airline fleet and crew robust planning. *Transportation Science*, *43*(1), 2-16.

Gregory, K. B., Winn, W., Johnson, K., & Rosekind, M. R. (2010). Pilot fatigue survey: exploring fatigue factors in air medical operations. *Air Medical Journal*, *29*(6):309-319.

Honn, K. A., Satterfield, B. C., McCauley, P., Caldwell, J. L., & Dongen, H. P. A. (2016). Fatiguing effect of multiple take-offs and landings in regional airline operations. *Accident Analysis & Prevention*, *86*, 199-208.

Junya, S., & Ruishan, S. (2022). Forecasting crew fatigue risk on international flights under different policies in China during the COVID-19 outbreak. *Frontiers in Public Health*, *10*, 996664.

Lee, S., & Kim, J. K. (2018). Factors contributing to the risk of airline pilot fatigue. *Journal of Air Transport Management*, 67, 197-207.

Li, J. Q., Zhang, X. N., & Zhou, Y. R. (2023). The sustained attention characteristics of flight crews on exempt and non-exempt flights. *Chronobiology International*, *40*(7), 926-937.

Medard, C. P., & Sawhney, N. (2007). Airline crew scheduling from planning to operations. *European Journal of Operational Research*, *183*(3), 1013-1027.

Mercie, A., Cordeau, J. F., & Soumis, F. (2005). A computational study of benders decomposition for the integrated aircraft routing and crew scheduling problem. *Computers & Operations Research*, *32*(6), 1451-1476.

Michael, I., Wessel, V. L., Tomas, K., Christer, U., Stephen, H., Göran, K., David, K. & Torbjörn, Å. (2014). Validating and extending the three process model of alertness in airline operations. *Plos One*, *9*(10), e108679.

Mohamed, N. F., Zainuddin, Z. M., Salhi, S., Mohamed, N.H., & Mohamed, N. A. (2016). Development of hybrid algorithm for integrated aircraft routing problem and crew pairing problem. *Indian Journal of Science and Technology, 9*(48), 108696.

Novak, A., Badanik, B., Brezonakova, A. & Lusiak, T. (2020). Implications of crew rostering on airline operations. *Transportation Research Procedia*, *44*, 2-7.

Oken, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*, *117*(9), 1885-1901.

Parmentier, A., & Meunier, F. (2020). Aircraft routing and crew pairing: updated algorithms at Air France. *Omega*, 93, 102073.

Powell, D. M., Spencer, M. B., Holland, D., Broadbent, E., & Petrie, K. J. (2007). Pilot fatigue in short-haul operations: Effects of number of sectors, duty length, and time of day. *Aviation, Space, and Environmental Medicine*, 2007, *78*(7), 698-701.

Quesnel, F., Desaulniers, G., & Soumis, F. (2017). A new heuristic branching scheme for the crew pairing problem with base constraints. *Computers & Operations Research*, *80*, 159-172.

Quesnel, F., Desaulniers, G. & Soumis, F. (2020). A branch-and-price heuristic for the crew pairing problem with language constraints. *European Journal of Operational Research*, *283*(3), 1040-1054.

Reis, C., Mestre, C., Canhao, H., Gradwell, D., & Paiva T. (2016). Sleep and fatigue differences in the two most common types of commercial flight operations. *Aerospace Medicine and Human Performance*, *87*(9), 811-815.

Sabaner, E., Kolbakir, F., & Ercan, E. (2022). Evaluation of fatigue and sleep problems in cabin crews during the early COVID-19 pandemic period. *Travel Medicine and Infectious Disease*, *50*, 102430.

Saddoune, M., Desaulniers, G., Elhallaoui, I., & Soumis, F. (2011). Integrated airline crew scheduling: A bi-dynamic constraint aggregation method using neighborhoods. *European Journal of Operational Research*, *212*, 445-454.

Saddoune, M., Desaulniers, G., Elhallaoui, I., & Soumis, F. (2012). Integrated airline crew pairing and crew assignment by dynamic constraint aggregation. *Transportation Science*, *46*(1), 39-55.

Saemi, S., Komijan, A. R., Tavakkoli-Moghaddam, R., & Fallah, M. (2022). Solving an integrated mathematical model for crew pairing and rostering problems by an ant colony optimisation algorithm. *European Journal of Industrial Engineering*, *16*(2), 215-240.

Sandhu, R., & Klabjan, D. (2007). Integrated airline fleeting and crew-pairing decisions. *Operations Research*, *55*(3), 439-456.

Seah, B. Z. Q., Gan, W. H., Wong, S. H., Lim, M. A., Goh, P. H., Singh, J., & Koh, D.S.Q. (2021). Proposed data-driven approach for occupational risk management of aircrew fatigue. *Safety and Health at Work*, *12*(4), 462-470.

Souai, N., & Teghem, J. (2009). Genetic algorithm based approach for the integrated airline crew-pairing and rostering problem. *European Journal of Operational Research*, *199*(3), 674-683.

Steiner, S., Fakleš, D., & Gradišar, T. (2012). Problems of crew fatigue management in airline operations. *International Conference on Traffic and Transport Engineering*, 29-30 November, Belgrade, Serbia, 617-623.

Stojković, M., Soumis, F., & Desrosiers, J. (1998). The operational airline crew scheduling problem. *Transportation Science*, *32*(3), 232-245.

Teymouri, A., Sahebi, H., & Pishvaee, M. S. (2023). Airline operational crew-aircraft planning considering revenue management: A robust optimization model under disruption. *International Journal of Industrial Engineering Computations*, *14*, 381-402.

Wen, X., Chung, S. H., Ma, H. L., & Khan, W. A. (2023). Airline crew scheduling with sustainability enhancement by data analysis under circular economy. *Annals of Operations Research*, https://doi.org/10.1007/s10479-023-05312-7.

Yan, S., & Chang, J.-C. (2002). Airline cockpit crew scheduling. *European Journal of Operational Research*, *136*, 501-511.

Yildiz, B. C., Gzara, F., & Elhedhli, S. (2017). Airline crew pairing with fatigue: Modeling and analysis. *Transportation Research Part C: Emerging Technologies*, *74*, 99-112.

Zeghal, F. M., & Minoux, M. (2006). Modeling and solving a Crew Assignment Problem in air transportation. *European Journal of Operational Research*, *175*, 187-209.

Zeighami, V., & Soumis, F. (2019). Combining Benders' decomposition and column generation for integrated crew pairing and personalized crew assignment problems. *Transportation Science*, *53*(5), 1479-1499.

Zeighami, V., Saddoune, M., & Soumis, F. (2020). Alternating Lagrangian decomposition for integrated airline crew scheduling problem. *European Journal of Operational Research*, *287*(1), 211-224.