Operational Risk in Airline Crew Scheduling: Do Features of Flight Delays Matter?

ABSTRACT

Our work is motivated by the increasing demand in the aviation sector and simultaneously aggravated poor punctuality. The airlines play an important role to improve their service level and mitigate the profit risks incurred due to their poor resources planning. However, to identify and mitigate operational risks faced by the airlines is complicated, as they are coming from both internal and external factors. Due to the realistic nature, we explore the flying time characteristics, and further model the consecutive interdependent departure-arrival times. It is a key feature included in this study that has not been studied in literature. We characterize the flying time of each flight by the heteroscedastic regression model. The analytical closed-form for the recursive relationship of the expected departure and arrival times of connective flight legs is then carried out. Accordingly, we propose a novel data-driven bi-criteria mathematical model in which the interdependent structures of the departure and arrival times of the consecutive flights is incorporated into the robust optimization. A column generation-based algorithm is developed to solve the proposed model. We found that, for more than 23% of the flights explored, the expected flying times are significantly influenced by its actual departure times. The real-data based computational examples identify that our proposed model sufficiently improves the reliability of the crew pairings decisions by reducing the total deviated time from the schedules with a slight increase of the total basic crew operations cost. Some managerial implications for robust crew pairing and determination of robustness level are discussed as well.

Subject Areas: Data-driven, Crew pairing, Robust optimization, and Operational risks.

INTRODUCTION

The aviation sector has achieved sharp development during the last two decades and contributes significantly to GDP, employment and social benefit. As reported by IATA in July 2018, Air passenger volume has kept growing since 2010 after the Great Depression. The industry-wide revenue passenger kilometers (RPK) increased by 6.1% year-on-year in May 2018. However, the steady increase in passengers demands each year aggravates the poor on-time performance of air transport. It is reported by Li and Song (2016), in the U.S., almost 29% of delays occurred due to the air carrier, 36% were due to the aircraft arriving late, and 31% were due to the National Aviation System. Only 4% of delays were due to heavy weather conditions. Although the airlines have put a tremendous amount of money on their resources planning for achieving efficient operations, how to reduce the disruption risks in terms of the economic losses is still a challenging task to them.

Among the airline operations problems, the crew pairing problems hold an important position. In 2017, the average crew costs per block minute reached $22.67, which ranked first among total direct operating costs per block minute (Airlines for America 2017). A crew pairing is a round-trip of a sequence of flights. An inefficient connection between flights incurs unnecessary long waiting time for the crew or large flight delays for the successive flights on schedule. In the existing literature, many studies focused on recovery control over crew pairing. Some severe disruptions can only be tackled by recovery options. However, most of the frictional disruptions can be reduced by proactive planning which may further decrease the utilization of recovery options. The existing research on proactive planning modeled the problems by discrete delays (Ahmadbeygi et al. 2010; Dunbar et al. 2014). Some studies adopted the worst case of the delay situation as inputs without using stochastic information (Ehrgott and Ryan 2002, Lu and Gzara 2015). In addition, they usually assumed independency between the departure and arrival delays. However, such assumption and simplification may violate the real situations.

For the busiest international airports around the world, the departure time of the day is a key factor which influences the airports' on-time performance. Reported by the Bureau of Transportation Statistics[1], the best time to fly is between 6 and 7 a.m. Flights with scheduled departures in that timeframe are arrived just 8.6 minutes late, on average. However, for every hour later extra minutes of delays are expected for arrival time. The delay times may hit a peak between 6 and 7 p.m., and they remain at above 20 minutes. In a recent study, Chung et al. (2017) identified that the arrival time of the flight is affected by its departure time. On the other hand, shared by a manager of a major airline, some pilots may control the flight speed within some ranges to compensate for the long-propagated delay in practice. Obviously, the above discussion indicates that the flying time is not an independent random variable, but the one affected by its departure time, which further determines the arrival time. Therefore, in our study we relax the assumption and model the flying time in a more realistic way, which makes both theoretical and practical contributions. Our study identifies that the realistic consideration of the departure-arrival times interdependency can help match well between flights and crews, which reduces the time deviation from the schedules with a slight increase in basic crew cost. Because it makes avoid underestimate on the cascading risks and overestimate on the flight delays due to unnecessarily long buffer times.

Although the topic regarding profit risk is popular in the supply chain management literature (Guo et al. 2017), it is seldom discussed and explored in the airline scheduling

[1] For more details, please refer to the official website of Bureau of Transportation Statistics https://www.transtats.bts.gov/ONTIME/.

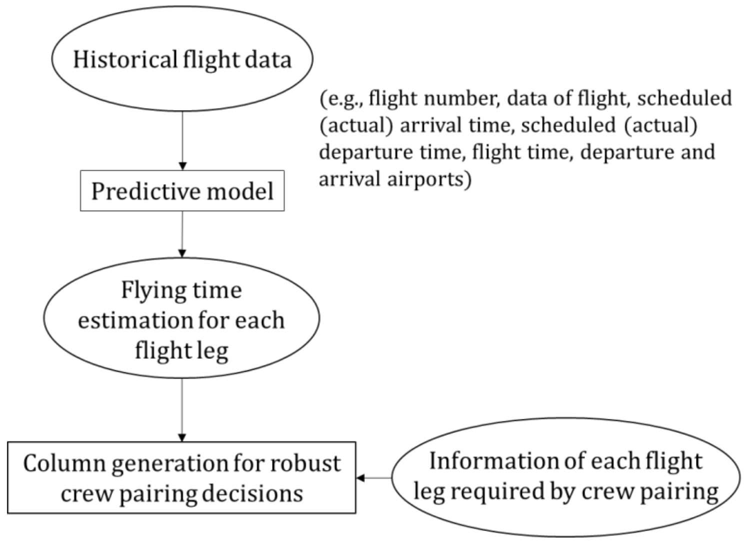
literature, especially for the risk incurred due to the decision making on the operational level. Motivated by the high uncertainties and profit risks faced by the airlines as well as the aforementioned observations, we propose a novel data-driven bi-criteria robust optimization method. Fig. 1 presents the main workflow of the proposed optimization model. The key questions here are as follows: (i) How does the departure time of the flight leg affect its flying time and further determine the arrival time? (ii) How to model the interdependency between the departure and arrival times in the robust optimization? (iii) How does the interdependency of the departure and arrival times of each flight leg affect the robustness of the crew pairing and mitigate the operational risks?

Based on the historical data which is from one of the global best-service airlines, the correlation analysis is conducted, which shows that, for more than 23% flight legs, the flying times are significantly influenced by the actual departure times from two dimensions, i.e., mean and variance. To make the problem tractable, we use the heteroscedastic regression model to predict the structure of the flying times. To reflect the real situation, a recursive relationship of the expected departure and arrival times of each flight leg is proposed and the corresponding analytical results are then obtained.

By applying the obtained explicit expressions of the expected departure and arrival times, a new bi-criteria optimization model for the robust crew pairing problem is proposed to make a trade-off between cost and robustness. A novel penalty function is proposed which works as a measure of robustness. It aims to reduce the chance of propagated and primary delays as well as the chance of flight cancelation or the utilization of reserve crew due to the infeasibility of the crew plan. The computational examples demonstrate the cost-efficiency of the proposed model and verify the impacts of the interdependency between the departure and arrival times on the robustness of the crew pairing. The results suggest that an appropriate level of the robustness consideration may help the airlines attain the best cost-effective operations decisions. However, extreme risk-aversion for the flight delays may do more harm than good to the airlines. Moreover, the cost ratio of the basic crew cost to the compensation expenses due to flight delays may affect the impact of the robustness level on the optimal decisions. When the cost ratio is low, a higher robustness level can be set for a risk-averse decision maker.

In the next section, we review some recent related studies. It is followed by the development of the data-driven robust optimization for the crew pairing problem, including the details of the model description, the predicted model of the flying time, the modeling of the consecutive departure and arrival times, and the bi-criteria robust optimization model. The proposed column generation approach for solving the bi-criteria robust optimization model is developed then. Accordingly, the computational study is performed and discussed regarding the impacts of the departure-arrival times interdependency as well as the robustness level on the operations decisions. In the end, we conclude the study with a further discussion of managerial implications, a summary of the contributions and limitations of this study, and the possibilities for future research. All the proofs are included in the appendix.

Figure 1: Workflow of the proposed data-driven robust optimization.



LITERATURE REVIEW

This paper mainly relates to two streams of literature, which includes robust crew pairing problems and big data analytics in operations management.

Robust crew pairing problems

As mentioned earlier, the stochastic variability in airline operations brings about large disruption cost and widespread negative impact on different aspects, including passengers, shippers, and airlines. So, more and more researchers study the robustness in airline operations and make a trade-off between cost and robustness. Although crew pairing is the last stage of the major airline operations (i.e., flight schedule design, fleet assignment, aircraft routing, and crew pairing), it occupies a tremendous expenditure of the airlines, which is just second after fuel cost. Therefore, robust crew pairing problems have drawn increasing attention from both academics and practitioners. In different studies, the robustness may have a specific definition, it mainly refers to the ability to remain the plans feasible or flexible under the variability or disruptions. The related studies can be mainly divided into three categories.

Recovery control

The first category aims to provide recovery solutions when the disruptions happen, such as crew swaps, reserve crews, to remain the original schedules feasible. Shebalov and Klabjan (2006) produced robust crew schedules by maximizing the number of move-up crews with the constraint to control that the crew cost does not increase too much above the one under the traditional cost-oriented crew plan. Another robust version of the airline crew pairing problem was proposed by Muter et al. (2013), in which case the robustness refers to the ability to accommodate extra flights at the time of operation by disrupting the original plans as minimally as possible. The robustness is attained by incorporating the planned crew pairing with several predefined recovery solutions. Bayliss et al. (2017) proposed a scenario based mixed integer programming for an airline reserve crew scheduling problem under crew absence uncertainty and delay. The objective was to minimize the disruption level under a total set of input scenarios, which was measured by a defined delay cancellation ratio. Recently, based on the previous study, Bayliss et al. (2019) proposed a new probabilistic model for the reserve crew scheduling, in which the crew absence is assumed to follow a binominal distribution. The objective of the optimization model is to minimize the total flight cancelation and reserve induced delays. Recovery control is usually based on the planned schedules and without incorporation with data analytics on flight information. In this study, instead of considering severe disruption, i.e., flight cancelation, we propose a novel data-driven proactive plan in which a new measure of robustness is proposed. It aims to minimize both the departure and arrival delays as well as the over-time of each duty and pairing so as to reduce the chances of flight cancelation and the use of reserve crew induced by unproper crew scheduling.

Integrated planning

The second category aims to improve the robustness by integrated planning. These literature focused on the study of the interdependency between the aircraft schedules and crew schedules and its impact on the propagated delays. Mercier et al. (2005) proposed an enhanced model incorporating robustness to handle the linking constraints. That was to impose a penalty whenever a crew changes aircraft on a restricted connection. Here, the robustness was improved through minimization of the number of connections between the flight legs that are not flown by the same aircraft. Based on this study, Weide et al. (2010) proposed a non-robustness measure of an integrated solution, which is a sum of sit-time dependent penalties overall restricted aircraft changes. They extended their integrated approach to both technical crew case and flight attendants case. The propagated delay is directly affected by its parent's flight's delay, which equals the parent's flight's delay minus the slack time between the connected flights. Both aircraft-induced delay and crew-induced delay may contribute to it. By the analysis of the dependency of the aircraft routing and crew scheduling, Dunbar et al. (2012) accurately estimated the propagated delays with stochastic information and minimized the cost of propagated delay in the integrated framework. An extensive study was done by Dunbar et al. (2014) with further consideration of re-timing for departure times. In the above-mentioned studies, they assumed the primary delays were independent of propagated delays. They were estimated according to the historical data, and the calculation of the propagated delay was based on the estimated values of primary delays induced by aircrafts and crews, respectively. Mohamed Ben Ahmed et al. (2018) proposed an integrated robust model for aircraft routing and crew pairing, in which a penalty would be induced for the cases when the critical aircraft and crew pairing connection occurs to avoid short buffer times. In reality, the cause of delay cannot be easily divided into aircraft-induced and crew-induced. Many other factors, such as airport congestion, air traffic control, etc., induces delays as well. On the other hand, the pilots may speed up as a way to complement its late departure time, which directly affects the final arrival time of the flight. Therefore, the primary delay is a random variable dependent on its propagated delay, which is simply assumed to be independent variables in the existing literature. Therefore, in our problem, we do not separate the delays due to different sources but focus on the study of the interdependency between the primary delay and its propagated delay.

Proactive planning

The third category mainly focuses on the study of proactive robust planning. Most of them took the measure of robustness by penalizing the delay related factors. Ehrgott and Ryan (2002) developed a bicriteria optimization framework to generate Pareto optimal schedules. The robustness was measured by a penalty function which was composed of the difference between buffer and worst-case delays. Schaefer et al. (2005) implemented a procedure for finding an appropriate solution to the problem of minimizing expected crew cost for crew scheduling with operations disruptions. The flight time disruption and ground delay were estimated by different probability distributions, respectively, including Gamma, Beta, Erlang. Schaefer and Nemhauser (2006) introduced a method for determining schedule perturbations that improved on-time performance without increasing crew costs. Yen and Birge (2006) developed a scenario-based stochastic model of the crew scheduling problem. They focused on the study of the interactions among crew pairings and its impact on the flight delays. The random variables, the flight delays in the stochastic model, was represented by a large number of scenarios which are generated by using truncated gamma or lognormal distribution to match disruption data from airlines. Ahmadbeygi et al. (2010) controlled the robustness of the crew scheduling by redistributing existing slack to reduce delay propagation. Minor modifications were made to the flight schedule (scheduled departure time) while leaving the original fleeting and crew scheduling decisions unchanged. Lu and Gzara (2015) presented a robust crew pairing formulation which was to optimize the problem against the worst instances. The robustness of the solutions was controlled by a user-defined parameter referred to as a protection level that limits the number of delays. Data uncertainty was modeled by intervals which are set by the planner. Quesnel et al. (2017) extended the crew pairing problem with additional constraints on the limitation of the total working time at each crew base to improve the solution robustness of further crew assignment step. The estimation of propagated delay induced by crew scheduling was studied by Wei and Vaze (2018). The authors developed a robust process to generate crew itineraries that were similar to real-world airline crew itineraries, based on an accurate and stable estimation of crew-propagated delays and disruptions. Four different features, including aircraft change, pushback, crew legality, and crew swaps were proposed to investigate their impacts on the crew-propagated delay and disruption. Our study is related to the previous study done by Chung et al. (2017). The authors proposed a robust crew pairing and schedule optimization model for cascading delay risk. Flight arrival delay was estimated by cascade neural network based on the historical data. The results demonstrated that the prediction accuracy on flight arrival delay can be increased when relating it with the flight departure delay.

Based on the previous results, we further analyze the interdependency between the flying time and its departure time in both practical and theoretical way. In that way, the flying time of a flight leg is represented as a heteroscedastic regression model which depends on the state of the random variable, i.e., its departure time. Different from the previous study, the main focus of this study is, through the analysis on the interdependency of the departure delay with propagated delay (coming from the delay of previous flight), to verify its impact on the robustness of the crew pairing decisions.

Big data analytics in operations management

In recent years, more and more operations management studies integrate big data analytics into their optimal decision-making processes. It has been widely applied in many areas such as revenue management, transportation management, risk analysis, and service operations (Choi et al. 2018; Cohen 2018). By leveraging the potential information hidden in big data, the decision makers seek to identify patterns and trends (Ettl et al. 2019), conduct predictive modeling and analytics (Demir 2014), and obtain superior strategies to facilitate their decision makings (Chan et al. 2016; Sun et al. 2019). Different big data analytics techniques are explored accordingly, such as statistics, machine learning and data mining. Each technique has its own strengths and weaknesses, which determines its suitable application fields. Among the statistical methods, regression models are commonly used in practice. For instance, Demir (2014) applied different data analytics techniques to forecast the patients at risk of unplanned readmissions. The results show that regression method had superior performance regarding predictive ability, compared to the data-driven methods. Ang et al. (2015) proposed a Lasso regression model called Q-Lasso method for wait time prediction, which provides insights for healthcare management. Sun et al. (2019) proposed a partially profiled LASSO regression model to analyze the promotion effects on the retailer sales and profits, which provides implications on promotion strategies. Different from the existent studies, our study proposes a heteroscedastic regression model to capture the complex structure of the flying time of each flight leg according to historical data.

DATA-DRIVEN ROBUST OPTIMIZATION FOR CREW PAIRING PROBLEM

Our objective is to generate cost-robustness trade-off solutions for crew pairing by a new proposed bi-criteria optimization model. The robustness is measured by a new proposed penalty function. Although the planned crew pairing obeys the regulations by FAA in terms of maximum flying time and elapsed time in a duty, the stochastic variability and disruption in air operations often make the crew scheduling illegal due to overtime in reality. The airlines will pay extra for reserve crews or overtime compensation. We, therefore, set the expected time exceeding the maximum elapsed time as a measure of the solution robustness to control the illegal situations. The propagated delay is the main source of overtime. It involves many aspects, such as aircraft routing, crew schedules, primary delays, operational recovery actions, etc. We do not conduct analysis on each of these factors, but to analyze whether interdependency exists between the propagated delay and primary delay. Through the regression analysis, it is identified that there exists a significant relationship between the flying time of the flight leg and its departure time. Thus, a new approach is proposed to model the flying time, arrival time and departure time of each flight leg. The primary delay which depends on flying time in our case is no longer independent from its propagated delay. A detailed description of the problem is displayed in the following section. Afterward, the heteroscedastic regression analysis for modeling the characteristics of the flying time is explored. Finally, the bi-objective mathematical program is then formulated.

Problem description

A crew pairing is a crew trip composed of a set of flights $\{j_1, j_2, \ldots, j_n\} \subset$ F, which starts and ends with the same specified station, called *crew base*. The crew pairings have to cover total flights *F*, and each flight *j* can only be served by one pairing. In the crew pairing, the flight legs are separated by *rest periods.* It usually refers to the overnight rest. The set of flight legs between two rest periods is called a *duty.* Meanwhile, there is a *sit time* between two consecutive legs in a duty, regulated as a minimum connection time for crews, aircrafts as well as passgengers. When a crew is on duty, they fly the consecutive flight legs under a series of regulations and contractual restrictions, for instance, the maximum flying time in a duty, a maximum elapsed time from the base of a pairing, etc. Under the consideration of those constraints, the typical objective for a crew pairing problem is to minimize the planned crew cost. Introduced by Saddoune et al. (2013), the basic cost usually includes the fixed cost for the crew, deadhead cost, connection time cost, rest period cost, and pairing minimum duty guaranteed (PMDG) cost.

Besides, a novel penalty function is proposed as a measure of robustness. It not only reflects the propagated delays, primary delays but also the disruption risk induced by exceeding the maximum elapsed time. Thus, two objective functions are established for both crew cost and potential risks from flight delay caused by the crew pairings. Our purpose is to get trade-off solutions by Pareto analysis. The calculation of the departure time and arrival time of each flight leg is the key factor for the penalty function. Under stochastic variability, they are random variables and linked each other by flying time. In the existing literature, the flying time or the block time was assumed to be independent of the departure time, which however does not hold in reality. In practice, the actual flight departure and arrival time can significantly deviate from the planned one by many factors, such as airport congestion, weather conditions, time of day, etc. We, therefore, assume it as a heteroscedastic regression model. The elaborate explanations in terms of relationship and calculation of the departure time, arrival time, and flying time of each flight leg are stated in the following section.

The predictive model of flying time

In this section, we identify the characteristics of the flying time of each flight leg by the regression model. Distinct regression models are used based on the characteristics of the data. For instance, Demir (2014) applied logistic regression and regression tree to predict patient-specific probability. As in the previous study (Chung et al. 2017), it is observed that the expected arrival delay of the flight is affected by its departure delay and the delay time is largely absorbed by the flight time. Inspired by this result, we further analyze how the flight time is affected by its departure time for each specific flight leg. Obviously, the flying time is affected by many other factors, such as flight numbers, departure and arrival airports, seasons and weekdays. In our problem, in order to incorporate the model of flying times into the proposed mathematical program and solved by the column generation, the two-variable regression analysis, given the other factors, is conducted between the mean/variance of the flying time and its departure time. The details regarding the modeling of flying time are discussed in the following.

Data source and preprocessing

We obtained over two years of flight data (from April 2015 to March 2017) which is provided from one of the major flight data service providers. We select the data of one major global airline, which is headquartered in Hong Kong[2]. A flight network which involves 14 airports

[2] Here we choose to apply the data of one main airline, whose flight network covers 35 countries and regions to

are investigated for data analysis. They are: 1. HKG (Hong Kong) 2. KIX (Osaka) 3. SIN (Singapore) 4. PEN (Penang) 5. OKA (Naha) 6. TPE (Taipei) 7. GUM (Guan) 8. ICN (Seoul) 9. SGN (Ho Chi Minh City) 10. BKK (Bangkok) 11. CMB (Columbo) 12. MAA (Chennai) 13. BOM (Mumbai) 14. DEL (Delhi). Multiple dimensional data are collected, including flight number, date of flight, scheduled arrival time, actual arrival time, scheduled departure time, actual departure time, actual flight time, departure and arrival airports, and aircraft type. To avoid the negative influence incurred due to the messy data, we preprocess the raw data and measure the outliers by the following rules. Since early flight departures will not be a reason causing arrival delays and thus will not cause propagating delays, for any flight departures earlier than the scheduled departure time, it is considered as on-time departures in our model to avoid negative values in the calculations. This practice also consists with the robust planning in which the earlier departure of the flight is not considered (Ionescu et al., 2016, Chung et al., 2017). In addition, since this study is on a regular operational basis, any flight with a departure or arrival delay longer than 3 hours is regarded as an extreme case and is excluded as in Lan et al., 2006 and Ionescu et al., 2016. This is because those extreme cases can be regarded as severe disruptions, which normally cannot be handled by proactive robust planning.

Heteroscedastic regression model

As the scheduled departure time of the flights with the same flight number is varied throughout the year, the flight legs are divided into different groups according to the origin and destination airports as well as its departure time rather than by the flight number. Both departure and arrival times of each flight leg are random variables, which depend on the arrival time of its preceding flight leg and the actual flying time. Figs. 2 and 3 illustrate the frequencies of the actual departure and arrival times of a selected flight leg. The x-axis represents the actual departure/arrival time point. The actual departure and arrival times may have different patterns, but for flight regular uncertainties, they can be approximated by known probability distributions, such as Normal, Gamma, etc. Here, *One-Sample Kolmogorov-Smirnov Test* is applied to test if departure/arrival time fits a known distribution well. For instance, it is verified that the departure time illustrated in Fig.2 is well fitted by Normal distribution $\mathrm{N}(1066, 14^2)$, as the test statistics $p = 0.2 > 0.05$. However, the hypothesis that the arrival time in Fig.3 follows a Normal distribution is rejected as the test statistics $p = 0.00 < 0.05$. This result is also intuitive, as the real situation is much more complex. More than one distribution would be involved. But, in our study, by conducting the Chi-square test, we choose Normal distribution to model the uncertainties of the departure and arrival time of all the flight legs to make the model tractable. Consequently, we define the

conduct our data analysis and computational experiments. It works as a representative to reflect the general situations faced by main airlines.

flight regular uncertainties occurring in the airline operations below.

Definition 1. *Flight regular uncertainties* refer to recurring probabilistic activities or events in airline operations at airport or at air. Historical data are often available to characterize their probability distributions.

Figure 2: An illustration on departure time distribution.

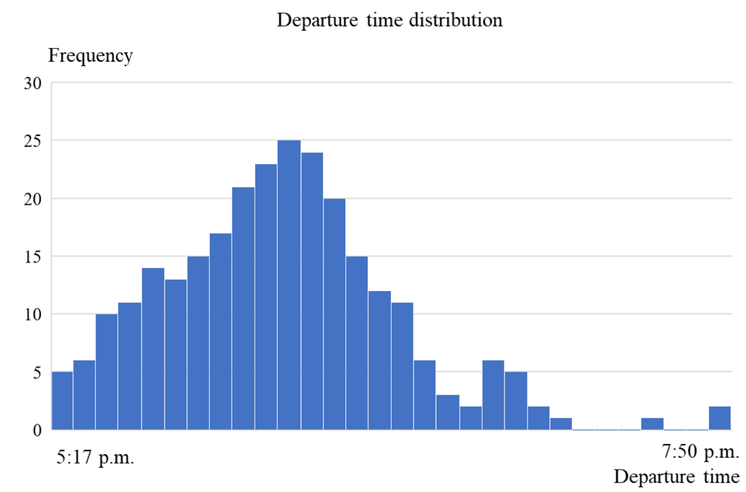
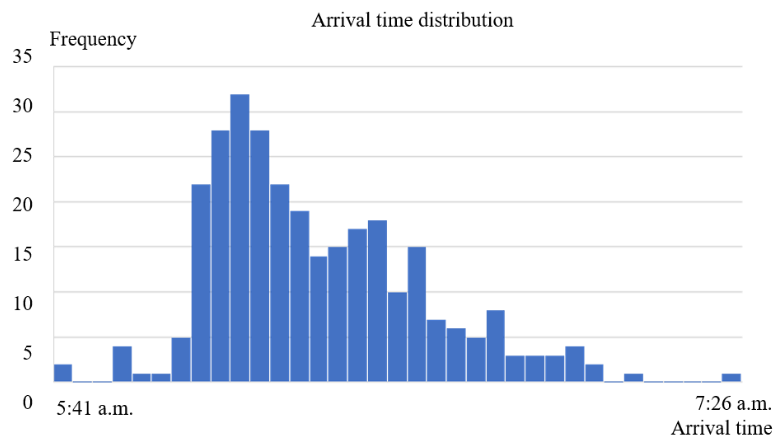


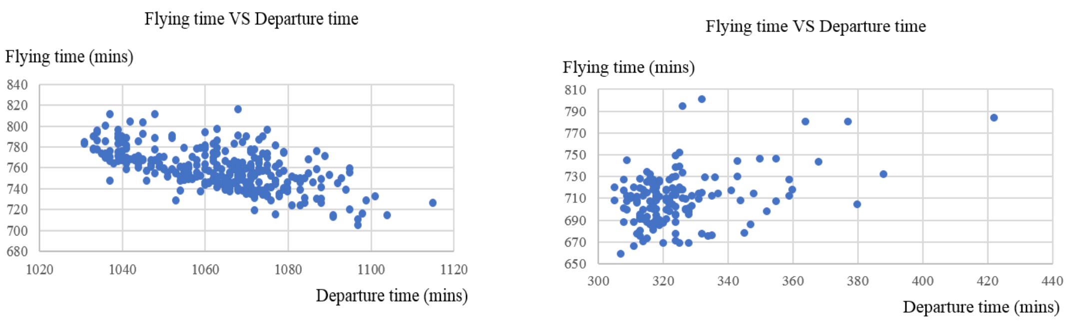
Figure 3: An illustration on arrival time distribution.



This definition is commonly used in the transport literature for the cases when shipping uncertainty is considered (Li et al. 2015). Figs. 4a and 4b display the correlation between the departure time and its flying time of a given flight leg. Fig. 4a shows the historical data of the flight departing from Hong Kong at 5 p.m. to London Heathrow at 6:20 a.m. There is a strong negative correlation between its departure time and flying time. This reflects the real practice that the aircraft would speed up to chase the planned arrival time, so the flying time is decreased when the departure time increases. On the other hand, Fig. 4b presents a strong positive correlation between the departure time and its flying time of the flight departing from

Hong Kong at 5 a.m. to Los Angeles at 5:25 p.m. This corresponds to the situations when the aircraft have to circle overhead and await permission to land when encountering the peak hour at the destination airport. On the other hand, it is observed that the volatility of the flying time varies along with the departure time. Motivated by the phenomenon, the systematic correlations analysis is conducted. The raw dataset is divided into 122 datasets corresponding to the specific departure time slots among 14 different airports from April 2015 to March 2017. The statistics results verify that, for over 23% of the total flight legs explored, their expected flying times are significantly correlated to the actual departure times (i.e., the statistical measure of p-value is less than the significance level of 0.05). In addition, for 21% of the total flight legs, their flying time variances are significantly correlated to their actual departure times (i.e., the statistical measure of p-value is less than the significance level of 0.05). It is shown that the correlation between the flying time and its departure time is related to the time slot when it gets departure from the airport, which does not exist for all the flight legs. This phenomenon is intuitive. As for the non-peak hours for the airport, the departure and arrival time would be on-time. However, for the flight legs departed at early in the afternoon, the flying time would be shorter to avoid facing airport congestion. While for those flight legs departed at late in the afternoon, the flying time would be longer due to the long queue to get landed.

Figure 4: Illustrations on the correlation between departure time and flying time.



a. Negative correlation.

b. Positive correlation.

Consequently, in the case of heteroscedasticity, a parametric shape of the flying time for this conditional volatility is considered as follows. The rationality of modeling the flying time as a heteroscedastic regression model can be mainly stated from two aspects in our study. Firstly, according to the description and correlation analysis on the historical data, it is found that both mean and variance of the flying time are affected by its departure time. The heteroscedastic regression model well shapes the characteristics of the flying time from two aspects, i.e., mean and variance. Secondly, due to the simple structure of the heteroscedastic regression model, the analytical closed-form of the flying time can be smoothly applied to the following modeling of departure and arrival times of each flight leg, which helps to obtain a tractable mathematical optimization framework.

Assumption 1. The flying time of leg j $FT_j(T)$ which depends on its departure time T can be modeled into a regression relationship as

$FT_j(T) = \mu_j(T) + \sigma_j(T)\epsilon.$

Here T is a random variable and is independent with ϵ, and ϵ is taken to be a standard normal random variable.

The flying time $FT_j(T)$ is assumed to be a heteroscedastic regression model (Yu and Jones 2004). Two aspects of the distribution of the flying time conditional on $T = t$ are expected to be estimated, i.e., the regression mean function $\mu_j(t) = E(FT_j(T)|T = t)$ and the regression variance function $\sigma_j^2(t) = Var(FT_j(T)|T = t)$.

To make the problem tractable and work well with the mathematical modeling in the following section, here we assume the regression mean function is in the form of simple linear function, i.e. $\mu_j(t) = \beta_0 + \beta_1 t$. The coefficients are first estimated by the ordinary least squares criterion (OLS) and refined by weighted least squares (WLS) procedures to remove the impact of heteroscedasticity on the parameters estimation. The main focus here is to have an efficient regression parameter estimation for the regression variance function $\sigma_j^2(t)$. There are two main streams to solve the heteroscedastic linear model, i.e., parametric and nonparametric approaches. For the parametric approach, the variance function $\sigma_{j,\theta}^2(t)$ is assumed to be known, where θ is the parameters vector to be estimated. For the nonparametric approach, kernel-based nonparametric local estimation for variance function is widely used in the statistics literature (e.g., Carroll 1982; Fan and Yao 1998; Heuchenne and Laurent 2017). In this study, both approaches are explored and the main procedures for the estimator of the variance function in conjunction with the estimator of the mean function are constructed as follows:

(1) Let $\hat{\beta}_L = \text{OLS}$ estimate, with $\hat{\beta}_L' = (\hat{\beta}_0, \hat{\beta}_1)$.

(2) Let r^2 be the vector of squared residuals, i.e., $r_i^2 = (FT(t_i) - \hat{\beta}_0 - \hat{\beta}_1 t_i)^2$.

(3a) (For parametric approach) Let $\hat{\theta}$ be the OLS estimator that minimizes the sum of squares errors, i.e., $\min_{\theta} \sum_{i=1}^{n} [r_i^2 - \sigma_\theta^2(t)]^2$. Here, n equals the sample size.

(3b) (For nonparametric approach) Let $\hat{v}(t) = \sum_{i=1}^{n} r_i^2 K\left(\frac{t_i - t}{b}\right) \{\sum_{i=1}^{n} \mathrm{K}(\frac{t_i - t}{b})\}^{-1}$, where $\mathrm{K}(\cdot)$ is a kernel function, b is the bandwidth, and n equals the sample size.

(4a) Define $\hat{\sigma}^2(t_i) = \sigma_{\hat{\theta}}^2(t_i)$, and compute the WLS estimate $\hat{\beta}_w = (\hat{\beta}_{0w}, \hat{\beta}_{1w})$ and corresponding squared residuals, i.e., $\tilde{r}_i^2 = (FT(t_i) - \hat{\beta}_{0w} - \hat{\beta}_{1w} t_i)^2$.

(4b) Define $\hat{\sigma}^2(t_i) = \hat{v}(t_i)$, and compute the WLS estimate $\hat{\beta}_w = (\hat{\beta}_{0w}, \hat{\beta}_{1w})$ and corresponding squared residuals, i.e., $\tilde{r}_i^2 = (FT(t_i) - \hat{\beta}_{0w} - \hat{\beta}_{1w} t_i)^2$.

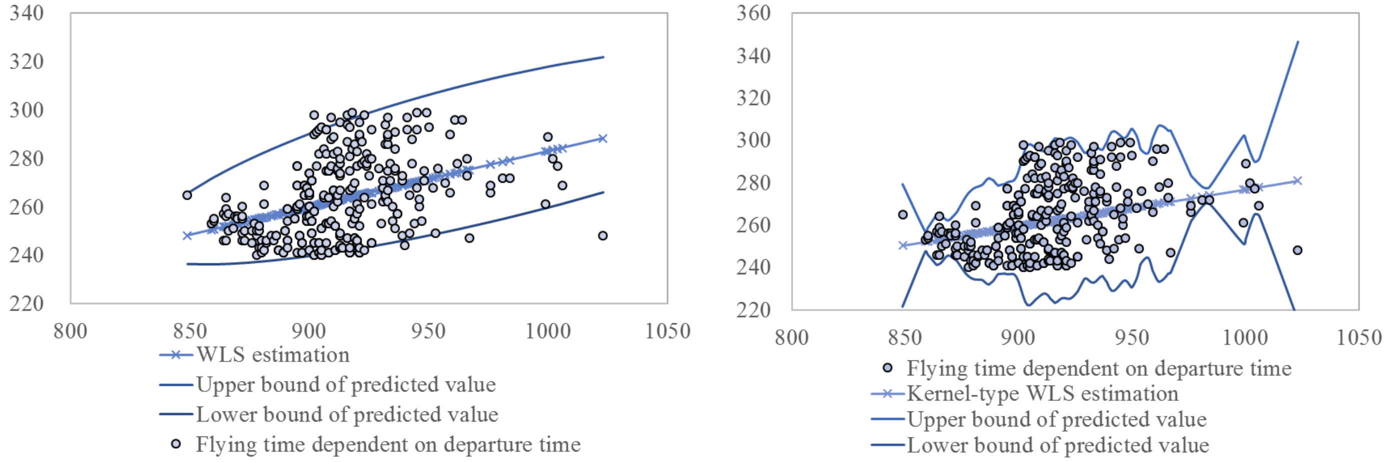
(5) Repeat (3) and obtain the conditional variance function $\sigma_{\tilde{\theta}}^2(t)/\ \tilde{v}(t)$.

Several researchers have discussed different rules of bandwidth selection. Please refer to Fan and Gijbels 1995 for more details. In this study, the selection of the bandwidth b is based on the mean squares errors (MSE). In addition, as most commonly used in the literature, we apply the Gaussian kernel function for the conditional variance estimate. In terms of parametric function $\sigma_\theta^2(t)$, different functions displayed in Table 1 are tested. The specific parametric variance function for each flight leg is selected according to the goodness-of-fit.

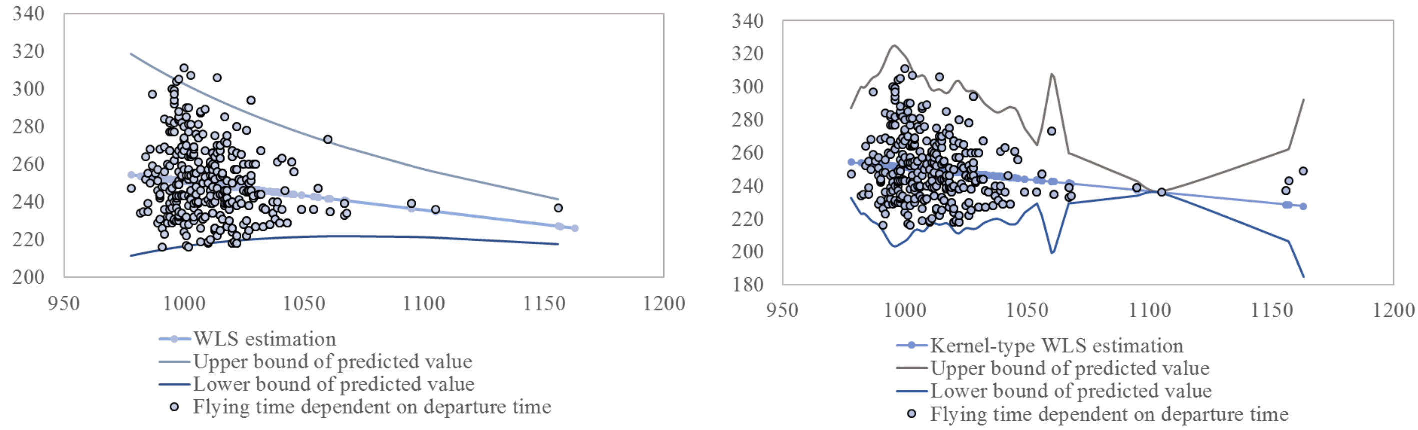
Table 1: List of parametric regression variance functions

1. $\hat{\sigma}^2(t) = \theta_0 + \theta_1 t$	6. $\hat{\sigma}^2(t) = \theta_0 + \theta_1 lnt$
2. $\hat{\sigma}^2(t) = \theta_0 + \theta_1 t + \theta_2 t^2$	7. $\hat{\sigma}^2(t) = \theta_0 e^{\theta_1 t}$
3. $\hat{\sigma}^2(t) = \theta_0 + \theta_1 t + \theta_2 t^2 + \theta_3 t^3$	8. $\hat{\sigma}^2(t) = \theta_0 \cdot {\theta_1}^t$
4. $\ln(\hat{\sigma}^2(t)) = \theta_0 + \theta_1 t$	9. $\hat{\sigma}^2(t) = \theta_0 \cdot t^{\theta_1}$
5. $\ln(\hat{\sigma}^2(t)) = \theta_0 + \theta_1 \frac{1}{t}$	10. $\hat{\sigma}^2(t) = \theta_0 + \theta_1 \frac{1}{t}$

Figure 5: illustration of the estimation under parametric and nonparametric approaches.



a. Regression mean function with positive correlation



b. Regression mean function with negative correlation.

*Upper and lower bounds are determined by the departure-time-dependent standard deviation of the flying time.

Evaluating the predicted performance of models

Table 2 shows the statistics R^2 adjusted and mean squares error (MSE) of the estimated mean regression models under different models. It is verified that there is a strong linear correlation between the expected flying time and its departure time. The impact of heteroscedasticity is well removed by the weighted least squares procedures. The R^2 adjusted is increased to 0.971 from 0.145 by parametric WLS. However, the corresponding R^2 adjusted under kernel-type variance estimator is increased to 0.310, much less than 0.971. The main reason behind that is, the nonparametric model is much sensitive to the quantity of the given data. Once data deficiencies exist at some points, a large error will be incurred, which largely affects the estimation of the mean function. Fig. 5 illustrates the predicted value of the conditional variance function under parametric and nonparametric approaches. It is obvious that the goodness-of-fit by the kernel-type estimator is much better than that of the parametric estimator. However, there is no explicit expression of the kernel-based variance function which is not compatible with further modeling of the robust optimization. Besides, its data sensitivity may also affect the robustness of the final plans on the crew pairings. Therefore, we apply the parametric approach to estimate both mean and variance regression functions.

Table 2: Model performance.

Models	R^2 adjusted	MSE
OLS	0.145	244.6
Parametric WLS	0.971	240.1
kernel-type WLS	0.310	242.4

Modeling of the consecutive departure and arrival times

In this section, we model the consecutive departure and arrival times of the flights in each duty, which are expressed into the recursive formulas.

Firstly, we start from the arrival time of the first flight j leg $\mathrm{arr}(j)$ in a given duty. To obtain the arrival time of the leg j, we have the following recursive formula:

$$\mathrm{arr}(j) = dep(j) + FT_j\big(dep(j)\big),\ \forall\ j \in F. \qquad (1)$$

It equals the sum of the departure time denoted as $dep(j)$ and its corresponding flying time denoted as $FT_j\big(dep(j)\big)$. Accordingly, we take expectation on both sides of Eq. (1), then we have the expected arrival time of flight leg j as follows:

$$\begin{aligned}\mathrm{E}[\mathrm{arr}(j)] &= \mathrm{E}\big[dep(j) + FT_j\big(dep(j)\big)\big] \\ &= \mathrm{E}[dep(j)] + \mathrm{E}\Big[\mathrm{E}\big[FT_j\big(dep(j)\big)\big|\mathrm{dep}(j) = t\big]\Big] \\ &= \mathrm{E}[dep(j)] + \mathrm{E}[\mu_j(t)],\ \forall\ j \in F.\end{aligned}$$

(2)

Refer to Jula et al. (2006) and Fu and Rilett (1998), as here $\mu_j(t)$ is a linear function, the mean and variance of $arr(j)$ can be expressed exactly as:

$$E[arr(j)] = E[dep(j)] + \mu_j(E[dep(j)]) \quad (3)$$

$$Var[arr(j)] = \left(1 + \mu_j'(\mathrm{E}[dep(j)])\right)^2 Var[dep(j)] + \sigma_j^2(E[dep(j)]) \quad (4)$$

Obviously, they cannot directly be generated without the mean and variance of its departure time, i.e., $\mathrm{E}[dep(j)]$ and $Var[dep(j)]$. Hence, we further explore the explicit expression of the departure time $\mathrm{dep}(j)$. Accordingly, we have the following assumptions:

Assumption 2. The aircraft is always available.

Assumption 3. The first leg j_s^d in each duty d is expected to be an on-time departure, i.e., $\mathrm{E}[\mathrm{dep}(\ j_s^d)] = dep^s(j_s^d)$. A primary variance $Var[\mathrm{dep}(\ j_s^d)] = v_j$ is given to the first leg of each duty.

Note that, the superscript *s* refers to the scheduled time of departure. The primary variance is the time deviation which is not induced by the crew schedule. Assumption 2 does not hold in reality when the aircraft changes happen. The arrival time of the aircraft may also be affected by its previous crew schedules. In this case, the crew schedules get intersected due to aircraft changes and their interaction further affects the flight delay, which is not our purpose of the study. In order to make our study focus on the interdependency of propagated delay and primary delay, and make the problem tractable, we do not consider the case when the aircraft is not available and the crew has to wait.

Under Assumption 2, we get the expression of $\mathrm{dep}(j)$ as follows:

$$\mathrm{dep}(j) = \mathrm{Max}(dep^s(j), arr(j^-) + \mathrm{Minsit}),\ \forall\ j \in F. \quad (5)$$

where, j^- is the preceding leg of flight leg j.

By taking expectation on both sides of Eq. (5), we have:

$$E[\mathrm{dep}(j)] = \mathrm{E}[\mathrm{Max}(dep^s(j), arr(j^-) + \mathrm{Minsit})],\ \forall\ j \in F. \quad (6)$$

Under the Assumption 3, there is no need for the crew to wait for the connected flight, thus the departure time of the connected flight leg $dep(j)$ is the maximum time between its scheduled departure time $dep^s(j)$ and the sum of the arrival time of preceding flight leg $arr(i)$ and required minimum connection time $Minsit$. Thus, Eqs. (3), (4) and (6) are composed of the recursive formulas for the calculation of the expected departure and arrival time of each leg in a duty. Their calculation is critical as it is much related to the measure of robustness in the optimization. Here, the challenge is to find the mean and variance of $\mathrm{dep}(j)$. Firstly, given the probability distribution of $arr(j)$ with mean and variance equal to $\mathrm{E}[arr(j)]$ and $\mathrm{Var}[arr(j)]$, the mean and variance of the departure time of leg *j*, i.e., $\mathrm{E}[\mathrm{dep}(j)]$ and $\mathrm{Var}[dep(j)]$, can be calculated according to the following proposition. Before we develop the specific close-form of the mean and variance of departure time for leg *j*, we first present the general case in Lemma 1.

Lemma 1. The mean and variance of $X = \max\{X_1, c\}$, where X_1 follows normal distribution $\mathcal{N}(\mu, \sigma^2)$ and c are a constant, can be formulated into the following equations, respectively.

$$E(X) = \mu\Phi\left[\frac{\mu-c}{\sigma}\right] + c\Phi\left[\frac{c-\mu}{\sigma}\right] + \sigma\phi\left[\frac{c-\mu}{\sigma}\right], \tag{7}$$

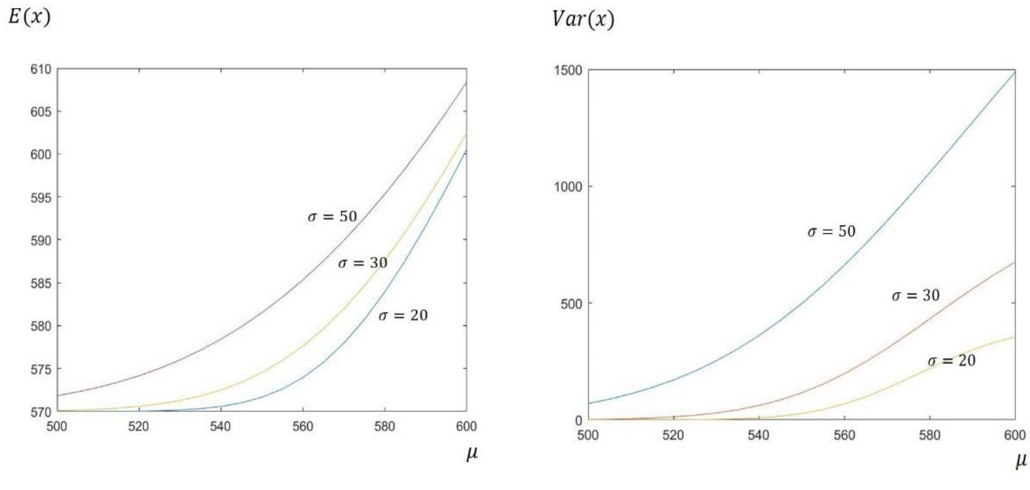
$$Var(X) = (\mu^2 + \sigma^2)\Phi\left[\frac{\mu-c}{\sigma}\right] + c^2\Phi\left[\frac{c-\mu}{\sigma}\right] + (\mu + c)\sigma\phi\left[\frac{c-\mu}{\sigma}\right] - [E(X)]^2, \tag{8}$$

where $\Phi[\cdot]$ is the cumulative distribution function of the standard normal distribution, $\phi[\cdot]$ is the probability density function of the standard normal distribution.
Proof: See appendix.

Fig. 6 presents the impacts of the statistical parameters of the arrival time of the preceding leg on those of the departure time of its successive flight leg. It is identified that there is an amplification effect of the arrival delay on the departure delay from two dimensions, i.e., mean and variance. This implies that a well-controlled flying time may help mitigate further propagated delay due to late arrivals.

For robustness checking, the analysis is further conducted for the case of Gamma distribution, which shows similar results as Normal distribution. It implies that the interrelationship between the arrival time and departure time of the consecutive flight legs is not affected by the properties of the probability distributions. The amplification effect between the arrival delay of the preceding leg and the departure delay of its successive leg is insensitive to the patterns of the probability distributions. For more details, please refer to the appendix.

Figure 6: Illustration of the relationship between (μ, σ^2) and $(E(x), Var(x))$.



To make the problem tractable, we assume the departure time $\text{dep}(j)$, the arrival time $\text{arr}(j)$ and the flying time of each flight leg j $\text{FT}(\text{dep}(j))$ follow the same probability distribution. In other words, if the departure time is assumed to follow a normal distribution, then its arrival time is also approximated to follow a normal distribution with $E[arr(j)]$ and $Var[arr(j)]$, vice versa. As verified by Ehmke et al. (2015), the percent error by such probability distribution approximation is within 1%. Therefore, under the setting of the normal distribution, the mean and variance of flight leg j can be calculated according to the Eqs. (7) and (8), and the details have been presented in Proposition 1.

Proposition 1. In a given duty, the mean and variance of the consecutive flight leg j, given its preceding flight leg i, which follows normal distribution $\mathcal{N}(E[arr(i)],\ Var[arr(i)])$, the mean and variance of the consecutive flight leg j satisfy the equations as follows:

$$E[dep(j)] = dep^s(j)\Phi\left(\frac{dep^s(j)-E[arr(i)]-Minsit}{\sqrt{Var(arr(i))}}\right) + \mathrm{E[arr(i)]}\Phi\left(\frac{E[arr(i)]+Minsit-dep^s(j)}{\sqrt{Var(arr(i))}}\right)$$

$$+\sqrt{Var(arr(i))}\phi\left(\frac{dep^s(j)-E[arr(i)]-Minsit}{\sqrt{Var(arr(i))}}\right).$$

$$Var[dep(j)] = \mathrm{E}[\max(dep^s(j), \mathrm{arr(i)} + \mathrm{Minsit})]^2 - E^2[dep(j)],$$

where,

$$\mathrm{E}[\max(dep^s(j), \mathrm{arr}(i) + \mathrm{Minsit})]^2 = \left(dep^s(\mathrm{j})\right)^2\Phi\left(\frac{dep^s(j)-E[arr(i)]-Minsit}{\sqrt{Var(arr(i))}}\right)$$

$$+(Var[arr(i)] + E^2[arr(i)] + 2MinsitE[arr(i)] + Minsit^2)\Phi\left(\frac{E[arr(i)]+Minsit-dep^s(j)}{\sqrt{Var(arr(i))}}\right)$$

$$+(\mathrm{E[arr(i)]} + \mathrm{Minsit} + dep^s(\mathrm{j}))\sqrt{Var(arr(i))}\phi\left(\frac{dep^s(j)-E[arr(i)]-Minsit}{\sqrt{Var(arr(i))}}\right).$$

Bi-criteria optimization modeling for robust crew pairing problem (BCM)

The following summarizes the notation used in this paper and the modeling of the proposed bi-criteria optimization model.

Notation:

Sets:

D	set of duties;
D_b^s	set of duties stating from crew base ($D_b^s \subset D$);
D_b^e	set of duties ending at crew base ($D_b^e \subset D$);
F	set of flights;
P	set of pairings.

Parameters:

Minsit	minimum time needed between two consecutive legs within a duty;
Maxfly	maximum amount of flying time for each duty;
$Maxelapse(d)$	maximum amount of elapsed time for duty d;
$dep^s(j)$	scheduled departure time of leg j;
$arr^s(j)$	scheduled arrival time of leg j;
c^w	unit cost for connection time;
c^{w^+}	unit waiting cost during flight connections;
c^{w^-}	unit cost for short connection time;
c^r	unit rest period cost;
c^g	unit penalty for a short duty period;
c^h	deadhead cost;
m_{ij}	=1, if the connection of leg *i, j* is feasible; = 0, otherwise;
$\mathrm{s, e}$	dummy starting and ending nodes for each duty (pairing);
T_{min}	ideal flying time of each duty;
r	ideal rest period;
$i\ \ ,j$	flight leg $i\ \ ,j$;
d, k, l	duty *d, k, l.*

Decision variables:

x_{ij}^d	=1, if flight leg i precedes flight leg j in duty d; = 0, otherwise;
x_{kl}^p	=1, if duty *k* precedes duty *l* in pairing *p*; = 0, otherwise.

Random variables:

$dep(j)$	actual departure time of leg j;
$arr(j)$	actual arrival time of leg j;
$FT_j(t)$	actual flying time of leg j given the departure time at t.

Objective functions:

In the proposed bi-criteria optimization model, two objective functions are proposed. The first one is to minimize the total basic crew cost shown in Eq. (9). Each pairing which is composed of a legal sequence of duties is associated with a basic cost. Refer to Chung et al. (2017), the basic cost components include deadhead cost c^H, the waiting cost for flight connections c^W and rest periods between consecutive duties c^R, and cost for pairing minimum duty guaranteed (PMDG) c^G. The deadhead cost will only happen if the first flight leg of each duty is not started from its duty base, presented in Eq. (10). The connection cost is a nonmonotonic function of the waiting time during the flight connections defined in Eq. (11). If the connection time is bigger than the required sit time, waiting cost will be induced. While when the scheduled connection time is less than the required sit time, a large cost will be caused. So, $c^{w^+} \ll c^{w^-}$. The cost of PMDG is to penalize the situation when the planned duty period is shorter than the ideal period T_{min}, as shown in Eq. (12), which helps to increase crew utility. Eq. (13) states the details of the rest period cost which includes a fixed cost c_f^r for each rest period and the additional cost related to the time period exceeding the ideal rest time period r.

The other objective function (14) is to minimize the deviations of the actual flight schedule from the planned one induced by crew schedule and schedule disruption due to safety illegality, i.e., working hour exceeds the maximum as regulated. It is composed of four parts, which are the expected departure and arrival delays of each leg as well as the overwork time of each duty and each pairing, respectively. α, β, γ, and δ are the corresponding parameters which reflect their priorities and can be determined by the decision maker.

$$\mathbf{Z_1} = \mathbf{Min}\ c^H + c^W + c^G + c^R, \tag{9}$$

where,

$$c^H = c^h\left(\sum_{d\in D}\sum_{j\in F} x_{sj}^d - \sum_{d\in D_b^s}\sum_{j\in F} x_{sj}^d\right), \tag{10}$$

$$c^W = \sum_{d\in D}\left[c^{w^+}\left(\sum_{i\in F}\sum_{j\in F\backslash\{i\}}\left(arr^s(j) - dep^s(j)\right)x_{ij}^d - Minsit\right)^+ \right.$$
$$\left. + c^{w^-}\left(\sum_{i\in F}\sum_{j\in F\backslash\{i\}}\left(arr^s(j) - dep^s(j)\right)x_{ij}^d - Minsit\right)^-\right]$$

(11)

$$c^G = \sum_{d\in D} c^g\left(T_{min} - \sum_{i\in F}\sum_{j\in F\backslash\{i\}}\left(arr^s(j) - dep^s(j)\right)x_{ij}^d\right)^+, \tag{12}$$

$$c^R = \sum_{p\in P}\sum_{k,l\in D\cup\{k\neq l\}}\sum_{i\in F}\sum_{j\in F\backslash\{i\}} c^r\left(y_{kl}^p x_{je}^k arr(j) - y_{kl}^p x_{si}^l dep(i) - r\right)^+ + \sum_{p\in P}\sum_{k\in D}\sum_{l\in D\backslash\{k\}} c_f^r\left(y_{kl}^p - 1\right). \tag{13}$$

$$\mathbf{Z_2} = \mathbf{Min}\ \alpha\sum_{d\in D}\sum_{i\in F}\sum_{j\in F\backslash\{i\}} x_{ij}^d E[dep(i) - dep^s(i)] + \beta\sum_{d\in D}\sum_{i\in F}\sum_{j\in F\backslash\{i\}} x_{ij}^d E[arr(i) - arr^s(i)]^+ + \gamma\sum_{d\in D} E\left[\sum_{j\in F} x_{je}^d arr(j) - \sum_{j\in F} x_{sj}^d dep^s(j) - Maxelapse(d)\right]^+$$
$$+ \delta\sum_{p\in P} E\left[\sum_{d\in D}\sum_{j\in F} y_{de}^p x_{je}^d\, arr(j) - \sum_{d\in D}\sum_{j\in F} y_{sd}^p x_{sj}^d\, dep^s(j) - Maxelapse(p)\right]^+. \tag{14}$$

The weighted sum method is utilized here to convert the multicriteria problem into a single objective problem. To make it compatible between cost (Objective function $\mathbf{Z_1}$) and robustness measure (Objective function $\mathbf{Z_2}$), both are normalized into the same range, i.e. (0,1). Therefore, the objective function of the problem is converted into:

$$\mathbf{Z_3} = \mathbf{Min}\ (1-\omega)\frac{Z_1 - \underline{Z_1}}{\overline{Z_1} - \underline{Z_1}} + \omega\frac{Z_2 - \underline{Z_2}}{\overline{Z_2} - \underline{Z_2}} \tag{15}$$

where, $\underline{Z_1}$ is the minimum value without Z_2 and $\overline{Z_1}$ is the maximum value Z_1, which corresponds to the value that Z_2 obtained its minimum value $\underline{Z_2}$. ω is a robustness factor within the range of (0,1). When ω moves to 1, it intends to be an extremely high level of robustness, in which case that operations cost is ignored. When ω moves to zero, it refers to the case of no consideration of robustness. The objective function is subject to the following constraints.

Constraints:

The constraints for this crew pairing problem are divided into four categories according to their functions, which are formulated as follows:

i) Flow balance constraints:

$$\sum_{d\in D}\sum_{i\in F\cup\{s\}\backslash\{j\}} x_{ij}^{d} = m_{ij}, \quad \forall\ j\in F. \tag{16}$$

$$\sum_{d\in D}\sum_{j\in F} x_{sj}^{d} = 1. \tag{17}$$

$$\sum_{d\in D}\sum_{i\in F\cup\{s\}\backslash\{j\}} x_{ij}^{d}\, m_{ij} - \sum_{d\in D}\sum_{n\in F\cup\{e\}\backslash\{j\}} x_{jn}^{d}\, m_{jn} = 0, \quad \forall\ j\in F. \tag{18}$$

$$\sum_{d\in D}\sum_{j\in F} x_{je}^{d} = 1. \tag{19}$$

$$\sum_{p\in P}\sum_{k\in D\backslash\{D_b^e\cup l\}} y_{kl}^{p} = x_{ij}^{l}, \quad \forall\ l\in D\backslash D_b^s\ ; \forall\ i,j\in F; i\neq j\ . \tag{20}$$

$$\sum_{p\in P}\sum_{d\in D\backslash D_b^s} y_{sd}^{p} = 1. \tag{21}$$

$$\sum_{p\in P}\sum_{k\in D\backslash D_b^e} y_{kd}^{p} - \sum_{p\in P}\sum_{l\in D\backslash D_b^s} y_{dl}^{p} \leq 1 - x_{ij}^{d}, \quad \forall\ d\in D\backslash\{D_b^s\cup D_b^e\};\ \forall\ i,j\in F; i\neq j. \tag{22a}$$

$$\sum_{p\in P}\sum_{k\in D\backslash D_b^e} y_{kd}^{p} - \sum_{p\in P}\sum_{l\in D\backslash D_b^s} y_{dl}^{p} \geq x_{ij}^{d} - 1, \quad \forall\ d\in D\backslash\{D_b^s\cup D_b^e\};\ \forall\ i,j\in F; i\neq j. \tag{22b}$$

$$\sum_{p\in P}\sum_{d\in D\backslash D_b^s} y_{de}^{p} = 1. \tag{23}$$

ii) Time restriction constraints:

$$\sum_{i\in F}\sum_{j\in F\backslash\{i\}} E[arr(j) - dep(j)]\, x_{ij}^{d} \leq Maxfly, \quad \forall\ d\in D. \tag{24}$$

$$\mathrm{E}[dep(j) - arr(i)]x_{ij}^{d} \geq Minsit, \quad \forall\ d\in D; j\in d. \tag{25}$$

iii) Definitional constraints:

$$arr(j) = dep(j) + FT_j\big(dep(j)\big), \quad \forall\ j\in F. \tag{26}$$

$$dep(j) = \sum_{d\in D}\sum_{i\in F\backslash\{j\}} x_{ij}^{d}\ \mathrm{Max}(dep^{s}(j), arr(i) + \mathrm{Minsit}), \quad \forall\ j\in F. \tag{27}$$

iv) Integer constraints:

$$x_{ij}^{d},\ x_{kl}^{p}\in\{0,1\},\ \forall\ i,j\ \in F;\ \forall\ k,l\in D.$$

The first part is for flight and duty flow balance, which is to make sure the feasible connections for consecutive flight legs in each duty and consecutive duties in each pairing. Constraints (16) state that, given a flight leg, it must belong to one duty in which there is a connected preceding flight leg. Constraints (17) and (19) present that, for each duty there must be one flight as the first leg and another flight as the last leg in the given duty. Constraints (18) ensure that for each flight leg in a duty, it has one preceding flight leg and one successive flight leg. Constraints (20)-(22) correspond to the cases of duties, which are similar to the cases for flight legs as shown in (16)-(19). For instance, constraints (20) state that the duty works if and only if there are flight legs assigned to the duty. The second part is for time restriction. Constraints (24) guarantee the expected flying time of each duty does not exceed the regulated maximum amount of flying time $Maxfly$. Constraints (25) restricts the connection time between consecutive flight legs to be greater than the regulated minimum $Minsit$. Lastly, the third part is the definition of the departure and arrival times of each flight leg, which is explored in the former section. The expectation of the arrival and departure times of each flight leg in a duty can be calculated according to Proposition 1.

COLUMN GENERATION

Crew pairing problem is known as a complex combinatorial optimization problem and is classified as NP-hard problems. While the model can have tremendous number of variables and feasible solutions, column generation is identified as a promising approach which considers only a set of promising pairings to deal with the problem (Lavoie et al. 1988, Saddoune et al. 2012; 2013, Cacchiani et al. 2016, Chung et al. 2017). In this section, a column generation based approach with modified multilabel correction algorithm is developed to solve the problem.

Restricted master problem for the column generation approach

The objective of the master problem in the crew pairing problem is to determine a set of crew pairings covered all the flights with minimum cost. Let P be the set of all feasible pairings, and c_p be the cost of crew pairing $p \in P$. Here, we define a binary variable a_{fp} and a binary decision variable x_p. a_{fp} equals to 1 if the flight $f \in F$ is covered by the pairing $p \in P$, and equals to 0 otherwise. x_p equals to 1 if pairing $p \in P$ is selected in the solution and equals to 0 otherwise. The problem is formulated as the following set covering model:

$$Z = \text{Min} \sum_{p \in P} c_p x_p \tag{28}$$

Subject to:

$$\sum_{f \in F} a_{fp} x_p \geq 1, \qquad \forall\ p \in P \tag{29}$$

$$x_p \in \{0, 1\}, \qquad \forall\ p \in P \tag{30}$$

The objective function (28) is to minimize the total cost of the selected pairings, while the constraint (29) ensures that each flight is covered by at least one selected pairing.

The master problem is converted into a restricted master problem (RMP) with a subset of feasible crew pairings. The RMP is solved iteratively and the set P is updated in every iteration. In the initial set P, one pairing covers only one flight and a deadhead is utilized to ensure that the pairing starts and ends at the same base. At each iteration, the dual prices of the RMP are determined and transferred to the pricing subproblem for generating a new column (pairing). By solving the subproblem, the pairing with the most negative reduced cost is identified and added to set *P* as a new column. The column generation procedures will stop when all the reduced cost are positive.

Pricing problem

The goal of the pricing subproblem is to determine a column with a minimum reduced cost. A column corresponds to a feasible pairing, which is composed of a number of duties. To generate a feasible pairing, a duty network is constructed. Let G = $(\mathcal{N}, \mathcal{A})$ be the completed duty network where *N* is the set of nodes and *A* is the set of arcs for all feasible connections of the nodes. Each duty d $\in$ D corresponds to a duty node $n \in \mathcal{N}$. In the network, a resource variable r_n^1 corresponding to the values of the duty period in a pairing is associated with each duty node, and a resource window [min, max] defining the minimum and maximum values of the resource in a pairing is given. To construct a resource-feasible path, the value of the resource variable will be accumulated along the path, and to be ensured within the corresponding resource window. A problem to determine a resource-feasible path with minimum reduced cost is regarded as a shortest path problem with resources constraints (SPPRCs). Now, we explain the computation of the reduced cost.

A feasible pairing p is represented by a resource-feasible path ρ = {o, n_1, n_2,…, e} from the start node $\{\boldsymbol{o}\}$ to the end node $\{\boldsymbol{e}\}$ in *G*. An arc (n^-, *n*) $\in \mathcal{A}$ is associated with a fixed arc cost $t^f_{n^-,n}$, and a variable arc cost $t^v_{n^-,n}$. $t^f_{n^-,n}$ represents the crew cost in Eq. (31), which depends on the duty *n* and its prior duty n^-. $t^v_{n^-,n}$ represents the robustness cost, which depends not only on the duty nodes n^- and *n*, but also on the preceding duty nodes of n^- in the path. Thus, it needs to be determined by constructing the path. A pairing cost c_p equals to the sum of the arc costs in path ρ as shown in Eq. (32), where *K* is a constant to denote the basic fixed cost for a pairing.

$$t^f_{n^-,n} = c^H_{n^-,n} + c^W_{n^-,n} + c^G_{n^-,n} + c^R_{n^-,n}, \forall (n^-,n) \in \mathcal{A} \tag{31}$$

, where $c^H_{n^-,n}$, $c^W_{n^-,n}$, $c^G_{n^-,n}, c^R_{n^-,n}$ are referred to Eq.(10-13).

$$c_p = K + \sum_{n \epsilon \rho}(t^f_{n^-,n} + t^v_{n^-,n}) \tag{32}$$

, where n^- is the node prior to *n* in path ρ.

The pairing cost is further modified to a reduced cost by subtracting the dual prices from the master problem as shown in Eq. (33). Given the reduced cost $\bar{c}_p$ of the pairing p in Eq. (33), the aim of the sub-problem is to find the path ρ* (pairing) with the most negative reduced cost as shown in the complete objective function (34).

$$\bar{c}_p = (c_p - \sum_{f \in F} \pi_f \cdot a_{fp} \quad)$$

(33)

, where $\pi_f, \forall f \in F$ is the dual price associated with the constraints (29) in the master problem.

$$\bar{c}_{p^*} = \min K + \sum_{n\in\rho}(t^{Z}_{n^-,n} - \sum_{f\in D_\rho}\pi_f) \tag{34}$$

, where D_ρ is a set of flights covered in the duties of path ρ, $t^{Z}_{n^-,n}$ is the normalized value of $(t^{f}_{n^-,n} + t^{v}_{n^-,n})$ based on Eq. (15).

A modified multilabel correction algorithm is developed for searching the shortest path ρ^* with the most negative $\bar{c}_p$ in the network.

Let L be a set of nodes to be handled in the algorithm, a set of attributes (labels) $\{tc_n, tr^1_n\}$ represent the accumulated total cost and the accumulated resource variable at each nodes $n \in N$ along the path. The values of these labels are propagated forward through the network by adding a new node *n* to the previously computed partial path by using the Eqs. (35) and (36) respectively, where $\rho(n^-)$ is the prior node to node *n* in path ρ.

$$tc_n = tc_{p(n^-)} + t^{Z}_{\rho(n^-),n} - \sum_{f\in D^\rho_n}\pi_f \tag{35}$$

, where D^ρ_n is a set of flights covered in the duty of node n in path ρ.

$$tr^1_n = tr^1_{\rho(n^-)} + r^1_n \tag{36}$$

To execute the searching, first, all the attributes in the network are set as $+\infty$ except those with the start node $\{\boldsymbol{o}\}$ are set as zero. Starting from node *o*, $L = \{\boldsymbol{o}\}$. Removing a node from *L* (i.e. node $\boldsymbol{o}$ in this case*)*, for each connection of ($\boldsymbol{o}$, *n*) $\in \mathcal{A}$, compute the attributes $\{\widetilde{tc_n}, \widetilde{tr^1_n}\}$ where the tilde variables denote the temporary values of the attributes at node *n*. Hence, to determine the $t^{Z}_{\rho(n^-),n}$, variable arc cost $t^{v}_{\rho(n^-),n}$ is computed by Eq. (18) simultaneously. However, this robustness cost may vary in different paths in which the preceding nodes (duties) are different. It is because the deviation of a flight's departure time and arrival time, overtime of a duty, and a pairing depend not only on its preceding flights' deviations but also the connection time and the rest period of these flights.

Subsequently, $\widetilde{tc_n}$ and $\widetilde{tr^1_n}$ can be determined by Eqs. (37) and (38) respectively. If the value of $\widetilde{tc_n}$ is less than the current tc_n, and the value of $\widetilde{tr^1_n}$ is within the resource window, tc_n and tr^1_n at node *n* will be updated as $\widetilde{tc_n}$ and $\widetilde{tr^1_n}$ respectively. The prior node of n in path ρ, $\rho(n^-)$, is set as $\boldsymbol{o}$. Then, put node *n* in *L*, $L = \{n\}$. For every node $n \in$ L, if arcs $(n, n^+) \notin \mathcal{A}$, where $n^+ \in N$, $n^+ \neq n$, or the value of $\widetilde{tr^1_{n+}}$ is out of the resource window, node *n* will be removed from set *L* without adding any newly connected node n^+ to *L*. The processes will be executed iteratively by removing a node from *L* and stop until $L = \{\emptyset\}$. Finally, the shortest path $\bar{\rho} = \{o,\ldots, \overline{\rho(\rho(e^-)^-)}, \overline{\rho(e^-)}, e\}$ with the minimum reduced cost $\bar{c}_{p^*}$ can be identified.

$$\widetilde{tc_n} = tc_{\rho(n^-)} + t^{Z}_{\rho(n^-),n} - \sum_{f\in D^\rho_n}\pi_f \tag{37}$$

$$\widetilde{tr^1_n} = tr^1_{\rho(n^-)} + r^1_n \tag{38}$$

When the shortest path $\bar{\rho}$ with negative $\bar{c}_{p^*}$ is determined, the path $\overline{\rho}$ is treated as a new pairing *p* and added to the solution pool of the master problem. The iterative process will stop when no pairing with negative reduced cost can be identified. Then, the variables determined in the master problem are the optimal set of pairing to cover the flights.

COMPUTATIONAL EXPERIMENTS

In this section, we conduct experiments which demonstrate (1) the significance of the interdependency of the departure and arrival times in improving the robustness of the crew planning, and (2) the impact of the robustness factor on the operations cost. The proposed approach and other comparison approaches are coded in Java, and implemented on IBM ILOG CPLEX 12.5/ Concert Technology in solving the linear programs of the RMP on a 2.4 GHz PC with 4 GB RAM.

Data Source and Experiment Settings

Our experiments are conducted based on two-years flight data for one of the major airlines headquartered in Hong Kong. The set of data is divided into two parts. One of which from one of the week in August is used for the creation of the three instances, which includes 64 to 122 flights and with the attributes of flight number, origin, destination, scheduled departure and arrival time, and its corresponding set of parameters from the regression analysis. The optimal solution by the proposed model BCM is obtained by the proposed algorithm. To obtain the parameters estimation for both the regression mean and variance functions of the flying times of the flights in that week, the rest of the data as mentioned in the previous section is used for heteroscedastic regression analysis. It is common practice to divided flights into long haul and short haul by regions for crew assignment in the airline industry. Here, we focus on the short haul flights. In addition, it is assumed that all the crew bases are the same, i.e., Hong Kong.

Results discussion

To demonstrate the performance of the proposed model BCM, we set the model of the case when $\omega = 0$ as the benchmark, whose objective is to minimize the total operations cost Z_1 for crew pairing. In addition, the parameters in the robustness measure objective Z_2 are set to be equal. To make the results perspicuous and easily compared with the benchmark model, we let them equal to 1, i.e., $\alpha = \beta = \gamma = \delta = 1$.

Firstly, the optimal solutions under two different settings of the robustness factor ω are derived by the proposed model BCM, i.e. $\omega = 0.5$ and 1. Here, we denote the corresponding optimal results as Z(0.5) and Z(1). Their corresponding comparison results with the benchmark Z(0) are summarised in Table 3. It is shown that the total time deviated from the schedule one is decreased by the proposed bi-criteria model. Besides, the corresponding reduction percentage is increased significantly with the number of flights involved in crew pairing. For the situation when the two criteria are equally weighted (i.e., $\omega = 0.5$), the reduction percentage regarding the total deviated time from the schedule one is 12.4%, 14.2% and 19.5% for the cases of 62, 96, 122 involved flights, respectively. On the other hand, the increase percentage of the basic crew cost is decreased slightly with the number of flights involved in crew pairing. For the case of 122 flights, the increase percentage regarding the basic crew cost is 1.5%. The reason behind the lower cost reduction is intuitive. For large-scale flights involved, the possible combination for crew pairing is exponentially increased. There is a much larger chance to derive more reliable connections between flights without extra cost compared to a small-scale problem. In particular, the extra cost incurred by deadhead or extra crew can be offset due to the reduction in waiting time and unnecessary long rest hours between flight connections. Moreover, it indicates that, once the economic savings from the reduction of the actual flight schedule deviation, such as flight delays, extra waiting times, over-work time, is no less than 1.5% of the basic crew cost, the total profit risk can be decreased by the total operations cost reduction. It is reported by one European aviation consultancy *inmarsataviation*, crew scheduling issues are estimated to account for 3% of all global flight delays, equating to an annual loss of US$3.6 billions[3]. Accordingly, there is a high possibility to achieve annual savings for the airlines by the proposed robust model, especially for the large-scale crew pairing.

Table 3: The performance of the proposed robust model for different problem scales.

No. of flights	$\omega = 0.5$		$\omega = 1$	
	Basic crew cost increase (%)	**Reduction of deviation from the planned schedule (%)**	**Basic crew cost increase (%)**	**Reduction of deviation from the planned schedule (%)**
62	2.7	12.4	28.0	23.4
96	1.5	14.2	35.6	30.4
122	1.5	19.5	21.5	32.7

For the extreme case when the robustness of the crew pairing get the highest priority while the basic cost is ignored, we obtained the solutions with least total time deviation and over-work hours. For the case of 122 flights involved, the deviated time can be reduced by more than 32%, compared with the benchmark model. However, the corresponding basic crew cost for operations are dramatically increased by 20%. While, for the case with fewer flights involved, there is a high chance to lose money for the airlines with pure robustness consideration. In other words, it is not wise for the airline company to consider a high level of robustness during their operational planning for the crew unless the economic loss incurred

[3] For the more details, please refer to the website https://www.inmarsataviation.com/en/benefits/operational-efficiencies/the-sky_s-the-limit-2--the-future-of-airline-operational-cost-savings-revealed.html.

due to the flight schedule deviation is extremely high and is comparable with the basic crew cost.

Next, to explore how the robustness factor influence the optimal crew pairing in terms of both robustness and total basic crew cost, the Pareto analysis is carried out and shown in Fig. 6. Firstly, the reduction amount of total deviated time is increased rapidly first and smoothly then with the robustness factor ω, which is ranged from 0 to 1. The largest reduction of total time deviation is achieved when $\omega = 1$. Simultaneously the expenses for the basic crew operations as defined in the objective Z_1 reaches its worst case. Table 4 presents the corresponding details regarding different problem scales. It is demonstrated that, regardless of the number of flights involved, the additional consideration of robustness can also improve the schedule reliability with a slight increase in the basic operations cost. It also indicates that the operational cost efficiency is greatly achieved by the analysis of the flying time characteristics as well as its accurate estimation based on the historical data. However, when the robustness factor ω is determined to be relatively large (e.g., $\omega = 0.9$), the amount of reduction regarding the deviated time becomes relatively small compared to the amount of increase in the basic crew operations cost.

In fact, it is always difficult for the decision maker to determine the value of ω to make a trade-off between the basic crew cost and the schedule robustness. Table 5 shows the ratio of the reduction of the deviation from the schedules to increase of the basic crew cost along with different robustness factor (ω). In these cases, we can see the ratio decrease dramatically when $\omega > 0.2$, especially in the large scale problem with 122 flights. It means, with an appropriate value of ω, a higher deviated time reduction can be achieved with a unit increase in basic crew cost. This analysis provides a good reference to the decision maker to determine the value of ω for the achievement of cost-effective solutions.

Moreover, the operations manager may also select an appropriate robustness level according to the budget level of the company. The extreme high robustness level may do more harm than good to the airline, especially for the situation when the compensation fees for flight delay and overtime work for the passengers and crew are relatively low. In addition, the influence of the factor ω on the optimal results not only depends on the risk-averse level of the decision maker but relates to the cost ratio between the basic crew cost and the compensation expenses for passengers and crew. For instance, a relatively high ω $(\text{i.e.}, \omega > 0.5)$ may be preferred to obtain a cost-effective decision when the cost incurred due to risk is low. In our experimental example, we set the cost parameters of the deviated time as 1 which is much lower than that of the basic crew operations. Thus, a high increase of time deviation reduction still can be attained when $\omega \geq 0.8$. When the cost parameters for robustness becomes large, it becomes much more sensitive to the increase of ω. Consequently, the same cost efficiency may be achieved with a smaller ω.

To further show the significance of the analysis of the flying time characteristics and the analytical modeling of the expected departure and arrival time in the robust optimization, we compared the optimal result of the BCM with a classical buffer model, which is widely used in the robust crew pairing literature (Ehrgott and Ryan 2002, Lu and Gzara 2015). It gives a fixed time buffer, i.e., 3 times standard deviations of the expected delay, to each flight leg to avoid the missing connection between flights. To make it reasonable, we exclude the extreme case (i.e., $\omega = 1$), and compare the results of the buffer model with the proposed BCM under the setting of $\omega = 0.9$. Table 6 shows the details of the comparison results in terms of basic crew cost and total time deviation. It is verified that the proposed BCM is superior in both aspects. A higher reliability of the flight schedule with a lower basic crew cost is achieved by BCM. The reason behind it is because the overestimate of the flight delay can be avoided by the proposed robustness model which is embedded with the explicit formulations of the interdependent departure and arrival times. It is good to observe the significance of the new consideration of the departure time based flying time and the development of the explicit modeling of the departure/arrival times for the consecutive flight legs. However, the way to model the characteristics of the flying time and its estimation can be further improved. Possibilities for future study are further discussed in the conclusion.

Due to the simplicity, many airlines still consider a fixed time buffer to improve the reliability of their operational planning. However, with the increasing demand and complexity of the flight schedule, a more reliable schedule is needed. The results demonstrated that the proposed approach can significantly reduce the schedule disruption because of the more accurate estimation of the flight arrival times. This implies that the proposed departure time based flying time and the explicit modeling of the departure/arrival times for the consecutive flight legs would be a promising way for an airline to estimate flight arrival time. In airlines' point of view, the flight arrival time estimation is not only important to crew scheduling, but also to many other robust operational planning, such fuel consumption estimation, aircraft routing, aircraft maintenance scheduling, etc.

Figure 6: Pareto front in terms of the percentages of operations cost increase and risk cost reduction.

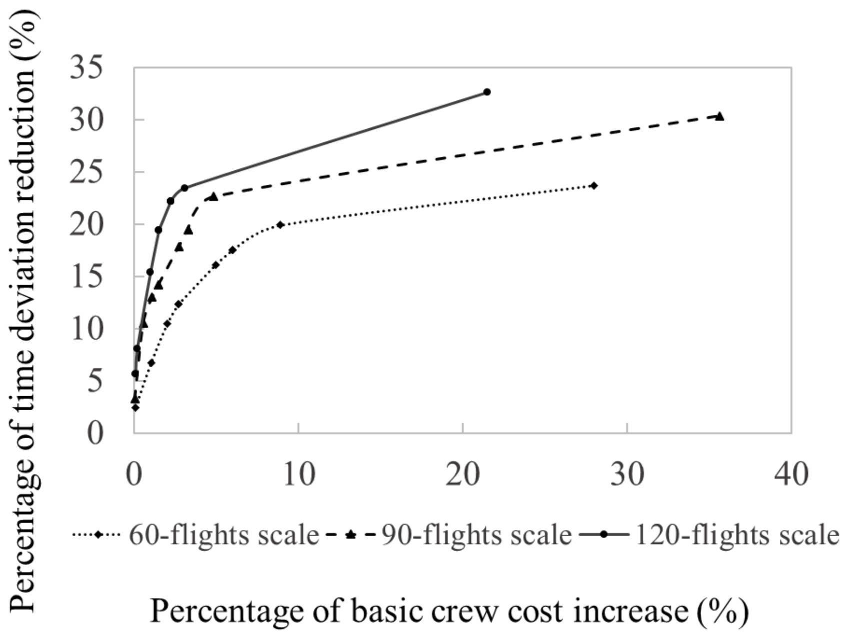


Table 4: Cost percentage pairs regarding different robustness factor ω.

No. of flights		ω									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
62	**Basic crew cost increase (%)**	0.1	1.1	2.1	2.7	2.7	5.0	6.0	6.0	8.9	28.0
	Deviated time reduction[4] (%)	2.5	6.7	10.5	12.4	12.4	16.2	17.6	17.6	19.9	23.7
96	**Basic crew cost increase (%)**	0.1	0.6	1.1	1.5	1.5	2.7	3.3	3.3	4.9	35.6
	Deviated time reduction (%)	3.3	10.6	13.0	14.2	14.2	17.8	19.5	19.5	22.6	30.4
122	**Basic crew cost increase (%)**	0.1	0.2	1.0	1.5	1.5	2.2	2.2	2.2	3.1	21.5
	Deviated time reduction (%)	5.7	8.1	15.4	19.5	19.5	22.3	22.3	22.3	23.5	32.7

Table 5: Ratio of deviated time reduction (%) to basic crew cost increase (%).

No. of flights	**ω**

[4] Note that, here deviation time reduction is short for the reduction percentage regarding the total deviated time from the schedule one.

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
62	25.0	6.1	5.0	4.6	4.6	3.2	2.9	2.9	2.2	0.8
96	33.0	17.7	11.8	9.5	9.5	6.6	5.9	5.9	4.6	0.9
122	57.0	40.5	15.4	13.0	13.0	10.1	10.1	10.1	7.6	1.5

Table 6: Performance of the proposed BCM *vs* a standard deviation based buffer model.

No. of flights		BCM	Classical buffer model	Difference (%)
122	**Basic crew cost increase (%)**	3.1	4.0	-0.9
	Deviated time reduction (%)	23.5	21.0	2.5

CONCLUSION

The profit risk incurred due to flight delays can be largely reduced through the efficient crew operations by the airlines. In particular, the development of information technology facilitates the operations efficiency. The big data analytics can help the airlines to further improve the robustness and cost efficiency of their limited resources planning. By data analytics on hundreds of short-haul flights in the past two years, the characteristics of the flying time are first explored and predicted by both parametric and nonparametric predicted models. The recursive formulations for the interdependent departure and arrival times of each flight and its connective flight legs are then explicitly modeled. Based on the data analytics of the flying time and its corresponding explicit formulations of the departure and arrival times, a new data-driven bi-criteria robustness optimization model is proposed. The computational experiments show how the consideration of interdependency of the departure-arrival times results in high cost efficiency, with the increase of the reduction amount of flight schedules deviation. The realistic consideration of the departure-arrival times interdependency can help reduce the actual schedule time deviation with a slight increase in basic crew cost, because it helps avoid underestimate on the cascading risks and overestimate on the flight delays due to unnecessarily long buffer times.

Managerial implications for robust crew pairing

From a managerial perspective, the discovered characteristics improve the flying time prediction so as to avoid the operational risk of flight cancellation due to underestimate on the cascading delays, and the risk of high basic crew cost due to unnecessarily planned buffer times. In fact, flying time duration is a critical component to determine a valid crew pairing. Unexpected extension of the flying time may induce the risk of exceeding the maximum duty hour limits of the crew members. It may cause flight disruption or cancellation of the successive flights by shortage of available crew members or aircrafts. By exploring the characteristic of flying time which is significantly influenced by the actual flight departure time, we can first improve the accuracy of flying time prediction in order to obtain a more reliable crew pairing at the planning stage. Moreover, at the operational stage, airlines can take a better reaction to any flight departured with a large deviation from its mean value, for example, to resechdule or to position reserved crew to the destination for replacement.

On the other hand, the robustness factor plays an important role in the optimal decision as well. It is encouraging that a significant reduction on the amount of expected departure and arrival delays, as well as overtime, can be achieved while only a slight increase of the basic crew cost is required. The results suggest that an appropriate level of the robustness factor may help the airlines attain the best cost-effective operations decisions. However, extreme risk-aversion for the flight delays may do more harm than good to the airlines. Moreover, the cost ratio of the basic crew cost and compensation expenses due to flight delays for both passengers and crew may affect the impact of the robustness level on the optimal decisions. When the cost ratio is low, a much higher value of the robustness factor can be set for the achievement of the robustness level expected by the decision maker.

Contributions and limitations

The implications of this research should be viewed in light of its limitations. Although our computational experiments conducted in this study are based on the real data from the airline industry, it could not cover the whole range of the settings of crew pairing, e.g., long-haul flights schedule. The flying time, which plays a critical role in the modeling of the proposed robust optimization is assumed to behave according to the heteroscedastic linear regression model. The formulations of the recursive departure and arrival times for each flight leg requires probability distribution assumptions to make the problem tractable. Consequently, the implications of the results must be viewed in light of those assumptions.

Despite its limitations, this study has made a contribution by developing a new optimization model for crew pairing with flying time forecasts and analytical expression of crew-induced flight delays. The computational studies evaluate the benefits of proactive planning considering the flying time characteristics. The results are discussed with a focus on managerial implications related to the trade-off between the robustness of the crew operations and the basic crew cost. Based on the data analytics on the flying time characteristics, the modeling of the consecutive departure-arrival times that affects each other is a key feature included in this study that has not been seen before in prior research. According to its realistic nature, the modeling framework developed in this study could motivate the development of the robust resource planning that includes this dimension for airlines, which may not be limited to crew planning. Besides, the research's results suggest that the data collection regarding the actual departure and arrival times could be critical for the airlines to determine the use of the predicted methods.

Future research direction for robust crew planning

A possibility for future research is to explore alternative models with further forecast updating to see how sensitive results are to changes in data as well as the model assumptions regarding flying time (Campbell 2017). The investigtaiton for ultra-high-dimensional regression relationships between the flying time and other factors will be an interesting but challenging direction for future resarch (Sun et al. 2019). Regarding the optimization algorithm, crew pairing problem is considered as an NP-hard problem. The problem associates with many operational constraints which are usually formulated as set covering problem or set partitioning problem. Column generation approach is widely adopted for solving such large scale linear programming problem. We suggest that in future, more works can be done to develop meta-heuristic approaches to deal with even larger scale problems. Besides, future research could also include the development of the back-up strategy embedded with the proposed proactive planning to tackle with the real-time severe disruptions which have been applied in supply chain management（Chen et al. 2015）. In addition, the capital constrained airlines (Shen et al. 2019) and risk-averse passengers (Ma et al. 2015) can be further considered in the robust optimization modeling to explore its impacts on the optimal robust decision.

REFERENCES

Ahmadbeygi, S., Cohn, A., & Lapp, M. (2010). Decreasing airline delay propagation by re-allocating scheduled slack. *IIE Transactions*, 42(7), 478-489.

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.

Ang, E., Kwasnick, S., Bayati, M., Plambeck, E. L., & Aratow, M. (2015). Accurate emergency department wait time prediction. *Manufacturing & Service Operations Management*, *18*(1), 141-156.

Bayliss, C., De Maere, G., Atkin, J. A., and Paelinck, M. (2019). Scheduling airline reserve crew using a probabilistic crew absence and recovery model. Journal of the Operational Research Society, 1-23.

Bayliss, C., De Maere, G., Atkin, J. A., & Paelinck, M. (2017). A simulation scenario based mixed integer programming approach to airline reserve crew scheduling under uncertainty. *Annals of Operations Research*, 252(2), 335-363.

Cacchiani, V. & Salazar-Gonz´ alez J.J., (2016). Optimal solutions to a realworld integrated airline scheduling problem. *Transportation Science*, 51(1), 250-268.

Chan, H. K., Wang, X., Lacka, E., & Zhang, M. (2016). A mixed-method approach to extracting the value of social media data. *Production and Operations Management*, *25*(3), 568-583.

Campbell, G. M. (2017). Time-Staged Overtime Staffing for Services with Updated Forecasts and Availabilities. *Decision Sciences*, 48(4), 691-722.

Carroll, R. J. (1982). Adapting for heteroscedasticity in linear models. *The Annals of Statistics*, 1224-1233.

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.

Chen, J., Zhao, X., & Shen, Z. J. (2015). Risk mitigation benefit from backup suppliers in the presence of the horizontal fairness concern. *Decision Sciences*, 46(4), 663-696.

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, *27*(10), 1868-1883.

Chung, S. H., Ma, H. L., & Chan, H. K. (2017). Cascading delay risk of airline workforce deployments with crew pairing and schedule optimization. *Risk Analysis*, 37(8), 1443-1458.

Cohen, Maxime C. "Big data and service operations." *Production and Operations Management* 27.9 (2018): 1709-1723.

Demir, E. (2014). A decision support tool for predicting patients at risk of readmission: A comparison of classification trees, logistic regression, generalized additive models, and multivariate adaptive regression splines. *Decision Sciences*, 45(5), 849-880.

Dunbar, M., Froyland, G., & Wu, C. L. (2012). Robust airline schedule planning: Minimizing propagated delay in an integrated routing and crewing framework. *Transportation Science*, 46(2), 204-216.

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.

Ehmke, J. F., Campbell, A. M., & Urban, T. L. (2015). Ensuring service levels in routing problems with time windows and stochastic travel times. *European Journal of Operational Research*, 240(2), 539-550.

Ehrgott, M., & Ryan, D. M. (2002). Constructing robust crew schedules with bicriteria optimization. *Journal of Multi-Criteria Decision Analysis*, 11(3), 139-150.

Ettl, M., Harsha, P., Papush, A., & Perakis, G. (2019). A data-driven approach to personalized bundle pricing and recommendation. *Manufacturing & Service Operations Management*. In Press.

Fan, J., & Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society:*

Series B (Methodological), 57(2), 371-394.

Fan, J., & Yao, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85(3), 645-660.

Guo, S., Choi, T. M., Shen, B., & Jung, S. (2017). Coordination and enhancement schemes for quick response mass customization supply chains with consumer returns and salvage value considerations. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, (99), 1-13.

Heuchenne, C., & Laurent, G. (2017). Parametric conditional variance estimation in location-scale models with censored data. *Electronic Journal of Statistics*, 11(1), 148-176.

Ionescu, L., Gqiggner, C., & Kliewer, N., (2016). Data Analysis of Delays in Airline Networks. *Business & Information Systems Engineering*, 58(2), 119-133.

Lan, S., Clarke, J.P. & Barnhart, C., (2006). Planning for Robust Airline Operations: Optimizing Aircraft Routings and Flight Departure Times to Minimize Passenger Disruptions. *Transportation Science*, 40(1), 15-28.

Lavoie, S., Minoux, M., & Odier, E. (1988). A new approach for crew pairing problems by column generation with an application to air transportation. *European Journal of Operational Research*, 35(1), 45–58.

Li, C., Qi, X., & Song, D. (2016). Real-time schedule recovery in liner shipping service with regular uncertainties and disruption events. *Transportation Research Part B: Methodological*, 93, 762-788.

Lu, D., & Gzara, F. (2015). The robust crew pairing problem: model and solution methodology. *Journal of Global Optimization*, 62(1), 29-54.

Ma, S., Li, G., Sethi, S. P., & Zhao, X. (2019). Advance selling in the presence of market power and risk-averse consumers. *Decision Sciences*, 50(1), 142-169.

Mercier, 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.

Muter, İ., Birbil, Ş. İ., Bülbül, K., ŞAhin, G., Yenigün, H., Taş, D., & Tüzün, D. (2013). Solving a robust airline crew pairing problem with column generation. *Computers & Operations Research*, 40(3), 815-830.

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.

Ruppert, D. (1997). Empirical-bias bandwidths for local polynomial nonparametric regression and density estimation. *Journal of the American Statistical Association*, 92(439), 1049-1062.

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.

Saddoune, M., Desaulniers, G., & Soumis, F. (2013). Aircrew pairings with possible repetitions of the same flight number. *Computers & Operations Research*, 40(3), 805-814.

Şafak, Ö., Çavuş, Ö., & Aktürk, M. S. (2018). Multi-stage airline scheduling problem with stochastic passenger demand and non-cruise times. *Transportation Research Part B: Methodological,* 114, 39-67.

Schaefer, A. J., Johnson, E. L., Kleywegt, A. J., & Nemhauser, G. L. (2005). Airline crew scheduling under uncertainty. *Transportation Science*, 39(3), 340-348.

Schaefer, A. J., & Nemhauser, G. L. (2006). Improving airline operational performance through schedule perturbation. *Annals of Operations Research*, 144(1), 3-16.

Shebalov, S., & Klabjan, D. (2006). Robust airline crew pairing: Move-up crews. *Transportation Science*, 40(3), 300-312.

Shen, B., Wang, X., Cao, Y., & Li, Q., (2019). Financing Decisions in supply chains with a Capital-Constrained Manufacturer: Competition and Risk. *International Transactions in Operational Research*. In press.

Sun, L., Zheng, X., Jin, Y., Jiang, M., & Wang, H. (2018). Estimating Promotion Effects Using Big Data: A Partially Profiled LASSO Model with Endogeneity Correction. *Decision Sciences*.

Wei, K., & Vaze, V. (2018). Modeling crew itineraries and delays in the national air transportation system. *Transportation Science*, 52(5), 1276-1296.

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

Yen, J. W., & Birge, J. R. (2006). A stochastic programming approach to the airline crew scheduling problem. *Transportation Science*, 40(1), 3-14.

Yu, K., & Jones, M. C. (2004). Likelihood-based local linear estimation of the conditional variance function. *Journal of the American Statistical Association*, 99(465), 139-144.

APPENDIX

Proof of Lemma 1

$\mathrm{E}[\mathrm{Max}(X_1, \mathrm{c})]$

$= \mathrm{E}[\mathrm{c}|X_1 \leq \mathrm{c}] + \mathrm{E}[X_1|X_1 > \mathrm{c}]$

$= \mathrm{cPr}(X_1 \leq \mathrm{c}) + \int_c^{+\infty} x_1 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X_1-\mu)^2}{2\sigma^2}} dx_1$

$= c\Phi\left(\frac{c-\mu}{\sigma}\right) + \int_c^{+\infty} (x_1-\mu) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} d(x_1-\mu) + \int_c^{+\infty} \mu \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} dx_1$

$= c\Phi\left(\frac{c-\mu}{\sigma}\right) - \frac{\sigma}{\sqrt{2\pi}} \int_c^{+\infty} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} d\left(-\frac{(x_1-\mu)^2}{2\sigma^2}\right) + \mu\left[1-\Phi\left(\frac{c-\mu}{\sigma}\right)\right]$

$= c\Phi\left(\frac{c-\mu}{\sigma}\right) - \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} |_c^{+\infty} + \mu\Phi\left(\frac{\mu-c}{\sigma}\right)$

$= c\Phi\left(\frac{c-\mu}{\sigma}\right) - \frac{\sigma}{\sqrt{2\pi}} \left[0 - e^{-\frac{(c-\mu)^2}{2\sigma^2}}\right] + \mu\Phi\left(\frac{\mu-c}{\sigma}\right)$

$= c\Phi\left(\frac{c-\mu}{\sigma}\right) + \mu\Phi\left(\frac{\mu-c}{\sigma}\right) + \sigma\phi(\frac{c-\mu}{\sigma}).$

$\mathrm{Var}[\mathrm{Max}(X_1, \mathrm{c})] = \mathrm{E}[\mathrm{Max}(X_1, \mathrm{c})^2] - \mathrm{E}^2[\mathrm{Max}(X_1, \mathrm{c})]$
$= \mathrm{E}[c^2|X_1 \leq \mathrm{c}] + \mathrm{E}\left[{X_1}^2\middle|X_1 > \mathrm{c}\right] - \mathrm{E}^2[\mathrm{Max}(X_1, \mathrm{c})].$

$\mathrm{E}[c^2|X_1 \leq \mathrm{c}] = c^2 \Pr(X_1 \leq c) = c^2\Phi\left(\frac{c-\mu}{\sigma}\right).$

$\mathrm{E}\left[{X_1}^2\middle|X_1 > \mathrm{c}\right] = \int_c^{+\infty} {x_1}^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} dx_1$

$= \frac{1}{2}\int_c^{+\infty} x_1 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} d({x_1}^2 - 2\mu x_1 + \mu^2) + \mu \int_c^{+\infty} x_1 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} dx_1$

$= -\sigma \int_c^{+\infty} x_1 \frac{1}{\sqrt{2\pi}} de^{-\frac{(x_1-\mu)^2}{2\sigma^2}} + \mu E[x_1|x_1 > c]$

$= -\frac{\sigma x_1}{\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} |_c^{+\infty} + \frac{\sigma}{\sqrt{2\pi}} \int_c^{+\infty} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} dX_1 + \mu E[x_1|x_1 > c]$

$= \frac{\sigma c}{\sqrt{2\pi}} e^{-\frac{(c-\mu)^2}{2\sigma^2}} + \sigma^2 \int_c^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} dx_1 + \mu E[X_1|X_1 > c]$

$= \mathrm{c}\sigma\phi\left(\frac{c-\mu}{\sigma}\right) + \sigma^2\left[1-\Phi\left(\frac{c-\mu}{\sigma}\right)\right] + \mu\left\{\mu\Phi\left(\frac{\mu-c}{\sigma}\right) + \sigma\phi\left(\frac{c-\mu}{\sigma}\right)\right\}$

$= (\sigma^2+\mu^2)\Phi\left(\frac{\mu-c}{\sigma}\right) + (\mathrm{c}+\mu)\sigma\phi(\frac{c-\mu}{\sigma}).$

Therefore, $\mathrm{Var}[\mathrm{Max}(X_1, \mathrm{c})] = c^2\Phi\left(\frac{c-\mu}{\sigma}\right) + (\sigma^2+\mu^2)\Phi\left(\frac{\mu-c}{\sigma}\right) + (\mathrm{c}+\mu)\sigma\phi\left(\frac{c-\mu}{\sigma}\right) -$

$\mathrm{E}^2[\mathrm{Max}(X_1, \mathrm{c})].$ ■

Proof of Proposition 1

It is similar to the proof of Lemma 1.

■

Lemma 2. The mean and variance of $X = \max\{X_1, c\}$, where X_1 follows gamma distribution $\Gamma(\alpha, \beta)$ and *c* is a constant, can be formulated into the following equations, respectively.

$$E(X) = cF(c; \alpha, \beta) + \frac{1}{\beta}\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}[1 - \mathrm{F}(\mathrm{c}; \alpha + 1, \beta)],$$

$$Var(X) = c^2F(c; \alpha, \beta) + \frac{1}{\beta^2}\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)}[1 - \mathrm{F}(\mathrm{c}; \alpha + 2, \beta)] - [E(X)]^2,$$

where $\Gamma(\alpha)$ is the gamma function, and $\mathrm{F}(\mathrm{x}; \alpha, \beta)$ is the cumulative distribution function of gamma distribution $\Gamma(\alpha, \beta)$ at point x.

Proof:

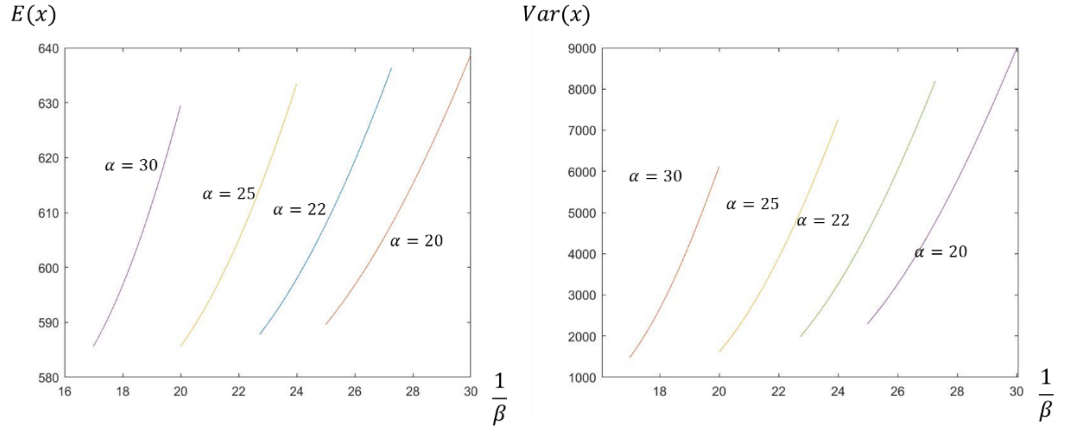
$$\mathrm{E}[\mathrm{Max}(X_1, \mathrm{c})]$$
$$= \mathrm{E}[\mathrm{c}|X_1 \leq \mathrm{c}] + \mathrm{E}[X_1|X_1 > \mathrm{c}]$$
$$= \mathrm{cPr}(X_1 \leq \mathrm{c}) + \int_c^{+\infty} x_1 \frac{\beta^\alpha {x_1}^{\alpha-1} e^{-\beta x_1}}{\Gamma(\alpha)} dx_1$$
$$= cF(c; \alpha, \beta) + \int_c^{+\infty} \frac{\beta^{\alpha+1} {x_1}^{\alpha} e^{-\beta x_1}}{\Gamma(\alpha+1)} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{1}{\beta} dx_1$$
$$= cF(c; \alpha, \beta) + \frac{1}{\beta}\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}[1 - \mathrm{F}(\mathrm{c}; \alpha + 1, \beta)].$$

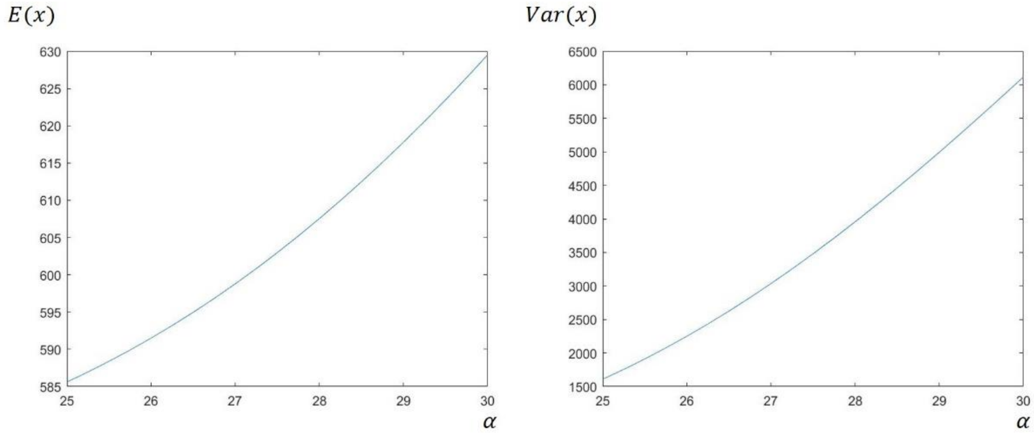
$$\mathrm{Var}[\mathrm{Max}(X_1, \mathrm{c})] = \mathrm{E}[\mathrm{Max}(X_1, \mathrm{c})^2] - \mathrm{E}^2[\mathrm{Max}(X_1, \mathrm{c})]$$
$$= \mathrm{E}[c^2|X_1 \leq \mathrm{c}] + \mathrm{E}\left[{X_1}^2\middle|X_1 > \mathrm{c}\right] - \mathrm{E}^2[\mathrm{Max}(X_1, \mathrm{c})]$$
$$\mathrm{E}[c^2|X_1 \leq \mathrm{c}] = c^2F(c; \alpha, \beta).$$
$$\mathrm{E}\left[{X_1}^2\middle|X_1 > \mathrm{c}\right] = \int_c^{+\infty} \frac{\beta^{\alpha+2} {x_1}^{\alpha+1} e^{-\beta x_1}}{\Gamma(\alpha+2)} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{1}{\beta^2} dx_1 = \frac{1}{\beta^2}\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)}[1 - F(c; \alpha + 2, \beta)].$$

Therefore, $\mathrm{Var}[\mathrm{Max}(X_1, \mathrm{c})] = c^2F(c; \alpha, \beta) + \frac{1}{\beta^2}\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)}[1 - F(c; \alpha + 2, \beta)] - \mathrm{E}^2[\mathrm{Max}(X_1, \mathrm{c})].$

■

Figure 1A: Illustration of the relationship between the statistical parameters (α, β) and $(E(x), Var(x))$.





$E(x)$
630
625
620
615
610
605
600
595
590
585
25
26
27
28
29
30
α
$Var(x)$
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
25
26
27
28
29
30
α