Comparative analysis on propagation effects of flight delays: a case study of China airlines

Citation for published version (APA):

DOI:
10.1155/2018/5236798

Document status and date:
Published: 29/01/2018

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Comparative Analysis on Propagation Effects of Flight Delays: A Case Study of China Airlines

Weiwei Wu,1 Cheng-Lung Wu,2 Tao Feng,3 Haoyu Zhang,1 and Shuping Qiu1

1College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China
2School of Aviation, UNSW Sydney, Kensington, NSW 2052, Australia
3Department of the Built Environment, Eindhoven University of Technology, 5600 MB Eindhoven, Netherlands

Correspondence should be addressed to Weiwei Wu; nhwei@nuaa.edu.cn

Received 8 September 2017; Revised 12 November 2017; Accepted 3 December 2017; Published 29 January 2018

This paper aims to capture the interdependency among the sequence of flight delays due to airline operations in airports, weather, and air traffic control conditions. A copula function is used to determine the distribution of delay sequence and examine the propagation effects. Using the actual data sourced from an airline in Asia Pacific region, it is found that flight delays could propagate to downstream airports/airlines, where the strength of delays was decreased, passed on, or increased. Considering the possible effects of increased delays under air traffic control or airline factors, scenarios that adjust flight schedules with additional buffer time were created and analyzed. Results show that, by adding buffer time efficiently, flight schedules can become more reliable.

1. Introduction

Flight delays have been one of the important problems in airport management and flight scheduling, blurring the efficiency of air system operations and the choice of passengers. A significantly high number of passengers have suffered in choosing a reliable flight or airport. In Europe, more than 2.4 million flights are delayed or canceled each year due to various factors, such as weather, airlines, and air traffic control (ATC). In China, the delays can be even more serious in some airports because of the increasing demand of air travel. As found in FlightStats [1], among the 61 largest airports worldwide, seven of them from China have been reported at the bottom line based on the on-time performance (OTP) rates.

Although some airports and/or airlines have put efforts in airport/airline management to reduce the possible delays, flight delays become unavoidable in some airports. In reality, multiple factors that impact flight delay are in many cases independent. A part of the indicators are related to the departure delay such as aircraft type, flight schedule, and flight departure sequence (Dai and Liou [2]). Others may relate to external triggers, like weather and airport capacity. A comprehensive overview on the potential factors that influence flight delay has been given by Xu et al. [3], where more than 50 potential factors were identified based on a detailed airport analysis. To find a suitable solution of flight delays, Liu et al. [4] presented an optimized GDP strategy, where the operational efficiency, airline and flight equity, and ATC risks were taken into account. The simulation study showed that the proposed solution reduced the total delay time, unnecessary ground delay, and unnecessary ground delay flights.

In spite of the direct effects of main factors on flight delay, many delays are induced by the previous flight, where the delay may be propagated to the subsequent flights (Wu [5]; Hansen and Hsiao [6]). Vigneau [7] presented the propagation of sequential flight delays as the dependence between upstream arrival delays and downstream departure delays, which are caused by factors not related to air traffic control (ATC). Mott [8] presented a method for modeling air carrier departure delays by considering the correlation of a priori demand data to significantly reduce prediction error. The results demonstrated that the accuracy of the prediction
delay time can be improved. Kafle and Zou [9] examined the degree of delay propagation by considering the new delays generated by each sequence of flights based on a joint discrete-continuous econometric model. They found that connected resources can significantly influence the initiation and progression of delay propagation.

Meanwhile, studies also found that delay time may be magnified when two subsequent flights are sharing the same aircraft (Kondo [10]). Shervin et al. [11] analyzed the slack between sequential flights in the planned schedule when delay occurs and showed how delay propagation can be reduced by redistributing the existing slack in the planning process. Moreover, the econometric analysis from Kafle and Zou [9] quantified how much propagated delay will be generated out of the newly formed delays that occur to each sequence of flights and revealed the effects of various influencing factors on the initiation and progression of propagated delays. Pyrgiotis et al. [12] proposed an AND (Approximate Network Delays) model based on a queuing engine and a delay propagation algorithm to study the complex phenomenon of delay propagation between airports. The model was applied to the network composed of 34 busy airports in the United States. The results show that, in some major airports, especially in hub airports, the propagated delay tends to put off traffic demand. Apart from that, Xu et al. [13] found that departure delays were the primary reason for the over-one-hour delay at the destination airports, and the taxiing delay was related to the previous delay. However, departure delay could be absorbed by scheduled turnaround time. When the cascading delay exceeds 30 minutes, more than 80% of flights may reduce their actual turnaround time.

In spite of the studies mentioned above, existing literature, which involves regression analysis, neural networks, Bayesian network models, and simulation methods, mostly analyzes the flight delays and delay propagation by focusing on specific flight legs. Studies on the relations between arrival delay of upstream flight and departure delay of its downstream flight, considering the total delay time of the entire flight legs, are rare. In order to fill this gap, this paper contributes to the analysis of propagation effects of flight delays by proposing a copula-based approach. In this regard, the correlation between sequential flight delays under the influence of different delay factors is explored. The magnitudes of delays on the subsequent flights are compared and possible scenarios are designed to examine the possible reduction of delay propagation by the adjustment and improvement of flight schedules.

2. Methodology

As we consider the situation of flight delay propagation, which is the correlation between arrival and departure delay, that is, late arrival of one flight causes late departure for the next flight on the itinerary of the same aircraft, we use the concept of propagated delay in line with the definition given by Lan et al. [14]: the delay induced by its prior flight delay. More specifically, we focus on the sequential flight delays \( d_{i-1}^d \) and \( d_{i-1}^a \), which are the departure delay of flight \( f_i \) and arrival delay of its upstream flight \( f_{i-1} \), respectively.

To measure the correlation between sequential flight delays induced by different delay factors, this paper uses a copula function. Copula function has been typically used in the study of finance and hydrological events with an emphasis to examine the interdependency between various random variables (Nelsen [15]). It has been also used to measure the predictability of on-time gate arrivals using the degree of concordance between gate arrival delays and block delays (Diana [16]).

In general, a copula function combines the joint distribution function of random variable, \( F(x_1, x_2, \ldots, x_n) \), and the marginal distribution function, \( F_{X_i}(X_1), F_{X_2}(X_2), \ldots, F_{X_n}(X_n) \). A copula function \( C \) satisfies

\[
F(x_1, x_2, \ldots, x_n) = C \left( F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n) \right).
\]

If the delay time of the upstream \( f_{i-1} \) was a continuous random variable \( x \) and the delay time of the downstream \( f_i \) that shares the same aircraft was a continuous random variable \( y \) with a joint distribution function \( F \), then (1) can be transformed as follows:

\[
F \left( d_{i-1}^a, d_i^d \right) = C \left( F_{X_1}(d_{i-1}^f), F_{X_2}(d_i^d) \right),
\]

where \( F(d_{i-1}^a) \) and \( F(d_i^d) \) are the marginal cumulative distribution functions of \( d_{i-1}^a \) and \( d_i^d \) and \( C \) is a bivariate function. Then, we determine the dependence structure of the random variables \( d_{i-1}^a \) and \( d_i^d \) by specifying a meaningful copula function to calculate the relativity between sequential flight delays.

The main issue related to the dependence between sequential flight delays is the magnitude of the influence of the upstream delay. It is necessary to understand how the downstream flight delay time changes according to the change of upstream delay time. In this particular case, the tail dependence analysis of copulas would be useful.

Let \( \lambda_{up} \) and \( \lambda_{lo} \) describe the upper-tail and lower-tail dependence of the random variables (dependence coefficient); then

\[
\lambda_{up} = \lim_{\alpha \rightarrow -1} \Pr \left[ Y > F_2^{-1}(\alpha) \mid X > F_1^{-1}(\alpha) \right],
\]

\[
\lambda_{lo} = \lim_{\alpha \rightarrow -1} \Pr \left[ Y < F_2^{-1}(\alpha) \mid X < F_1^{-1}(\alpha) \right],
\]

\[
\lambda_{up} = \lim_{\alpha \rightarrow 0} \Pr \left[ Y > F_2^{-1}(\alpha) \mid X > 0 \right] = \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha},
\]

\[
\lambda_{lo} = \lim_{\alpha \rightarrow 0} \Pr \left[ Y < F_2^{-1}(\alpha) \mid X < 0 \right] = \frac{C(\alpha, \alpha)}{\alpha},
\]

where \( \alpha \) is a probability value and \( F_i^{-1}(\alpha), i = 1, 2 \), represents the corresponding quantile of \( \alpha \).

Tail dependence coefficients and relativity measurements with extreme values can measure the tail dependence relativity of the delay times for sequential flights. That is, they measure the change in probability of departure delay \( Y \) for the downstream flight when the arrival delay \( X \) of the
upstream flight changes to a great extent. The tail dependence coefficients are within $[0, 1]$. If $\lambda = 0$, then the propagated delays are progressively tail-independent; otherwise, they are tail-related. When $\lambda$ is close to 1, the probability of changes in the downstream delays will correspondingly become larger when the upstream delays reach a certain value.

### 3. Data and Analysis

Flight data was sourced from an airline based in the Asia Pacific, and the data span is one year. In total, 9,325 flight delays were recorded. The data include flight data (e.g., time of departure and arrival), flight delays (e.g., time of the delay), and delays factors (e.g., delay causes).

#### 3.1. Delays Induced by Airlines, Weather, and ATC

Figure 1 shows how propagated delay happened between sequential flights, and Figure 2 shows the distribution of delay causes. Around 70% of the delays are attributable to the factors related to airlines (air carrier delay and aircraft late arrival), weather, and ATC, where the air carrier delay takes the largest share (25%).

Figure 3 shows that factors related to airlines account for the largest share during 9:00 a.m.–1:00 p.m., 11:00 a.m.–1:00 p.m., and 1:00 p.m.–3:00 p.m. This indicates that flights during these periods may be disturbed frequently by airline operations. The peak period of delay happens during 11:00 a.m.–1:00 p.m. from the morning and 1:00 p.m.–3:00 p.m. in the afternoon.

Figure 4 represents the delay distribution induced by weather. It shows that the largest proportion of delay happens during 7:00 a.m.–9:00 a.m. and 5:00 p.m.–7:00 p.m.

Figure 5 represents the delays due to ATC. It shows that the delays during 7:00 a.m.–9:00 am, 9:00 a.m.–11:00 a.m., 11:00 a.m.–1:00 p.m., and 5:00 a.m.–7:00 p.m. are higher than other times. To narrow down the scope of this study on flight delays, this paper will focus on major delay causes including airlines, weather, and ATC and examine how they affect flight delays and delay propagation.

In summary, it can be found that different delay factors during different periods within one day can have diverse effects on flight delays, which result in different levels of delay propagation through aircraft routing for the subsequent flights.

#### 3.2. Exploring Sequential Flight Delays

In order to examine the characteristics of delays in sequential flights and delay propagation, sequential flights that share the same aircraft
null
Figure 6

(2) As $\alpha$ increases to a certain value, $\lambda^{up}$ reaches a limit. That is, when the upstream arrival delays are large, the dependence between arrival and departure delays remains constant. In Figure 8, when the upstream delay probability is close to 0.95, the dependency coefficient $\lambda^{up}$ remains the same in all three states. Note that the effect of upstream delay under ATC factor is the highest with $\lambda^{up}$ close to 60%.

4.4. Conditional Probability Analysis Using Copulas. According to the above analysis, departure delay time of a downstream flight is related to how long its upstream flight was delayed. Therefore, we can calculate the probability of the downstream flight delay based on the upstream delay caused by different delay factors as follows:

$$P(\gamma_1 \leq Y < \gamma_2 \mid x_1 \leq X < x_2) = \frac{P(\gamma_1 \leq Y < \gamma_2, x_1 \leq X < x_2)}{P(x_1 \leq X < x_2)},$$

(4)

where

$$P(\gamma_1 \leq Y < \gamma_2, x_1 \leq X < x_2) = C(F_1(x_2), F_2(\gamma_2)) - C(F_1(x_1), F_2(\gamma_2)) + C(F_1(x_1), F_2(\gamma_1)) - C(F_1(x_1), F_2(\gamma_2))$$

$$P(x_1 \leq x < x_2) = F_1(x_2) - F_1(x_1).$$

(5)

The conditional probability $P(\gamma_1 \leq Y < \gamma_2 \mid x_1 \leq X < x_2)$ means that the probability of the delay time of the downstream flight is $[\gamma_1, \gamma_2)$ when the delay time of its upstream flight is $[x_1, x_2)$. As flight delay that is longer than 15 minutes is always considered to be delay time (Nelsen [15]), here we categorize the arrival and departure delays into five levels (between 15 minutes and 165 minutes) and added the extreme ranges [0, 15] and [165, $+\infty$) for comprehensive investigation. Then the value of $P$ was calculated under the conditions of airline, weather, and ATC factors, respectively, using (4).

The conditional probability examines the delays propagated to downstream flights. The propagated delay could be decreased ($x_1 \geq \gamma_2$), passed on ($x_1 = \gamma_1, x_2 = \gamma_2$), or increased ($x_2 \leq \gamma_1$). Figure 9 shows the probability histograms in terms of delay decrease, pass-on, and increase in sequential flights caused by the three delay factors. Figure 9(a) shows that the probability of the delay reduction effect under weather factors is larger than the delay reduction effect under the airline and ATC factors, and this increases with an increase in upstream delay. Results show that delays
(a) Density function and contour distribution of sequential flight delay time due to airline-related factors

(b) Density function and contour distribution of sequential flight delay time due to weather

(c) Density function and contour distribution of sequential flight delay time due to ATC

Figure 7
Figure 8: Tail dependence coefficients for three groups sample data.

Table 2: Probability distribution parameters and test statistics of the sample data.

<table>
<thead>
<tr>
<th>Sample groups</th>
<th>Flights</th>
<th>Probability distributions</th>
<th>Mean value</th>
<th>Standard deviation</th>
<th>Test statistic ($W^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>Normal</td>
<td>9.16</td>
<td>12.32</td>
<td>172.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lognormal</td>
<td>1.62</td>
<td>1.08</td>
<td>7.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weibull</td>
<td>8.76</td>
<td>0.92</td>
<td>21.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gamma</td>
<td>9.40</td>
<td>0.98</td>
<td>58.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
<td>20.52</td>
<td>25.83</td>
<td>167.06</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>Lognormal</td>
<td>2.55</td>
<td>0.95</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weibull</td>
<td>20.76</td>
<td>1.02</td>
<td>22.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gamma</td>
<td>17.02</td>
<td>1.21</td>
<td>24.86</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Normal</td>
<td>24.67</td>
<td>32.36</td>
<td>15.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lognormal</td>
<td>2.54</td>
<td>1.20</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weibull</td>
<td>23.09</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gamma</td>
<td>27.84</td>
<td>0.89</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
<td>32.31</td>
<td>37.66</td>
<td>10.42</td>
</tr>
<tr>
<td>2</td>
<td>D</td>
<td>Lognormal</td>
<td>2.88</td>
<td>1.21</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weibull</td>
<td>31.58</td>
<td>0.95</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gamma</td>
<td>33.31</td>
<td>0.97</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>Normal</td>
<td>9.31</td>
<td>14.60</td>
<td>64.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lognormal</td>
<td>1.60</td>
<td>1.09</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weibull</td>
<td>8.63</td>
<td>0.88</td>
<td>7.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gamma</td>
<td>10.12</td>
<td>0.92</td>
<td>11.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal</td>
<td>13.28</td>
<td>19.27</td>
<td>55.29</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>Lognormal</td>
<td>2.06</td>
<td>1.01</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weibull</td>
<td>13.03</td>
<td>0.97</td>
<td>5.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gamma</td>
<td>12.24</td>
<td>1.09</td>
<td>6.84</td>
</tr>
</tbody>
</table>
due to weather from flight D tend to be lower than those from flight C, especially when the length of the delays from flight C increases.

Figure 9(b) shows that the probabilities of delay propagation under the three delay factors are similar: they first decrease with the increase in the upstream delay time; then in the category of $[165, -]$, the probability increases abruptly. Therefore, if the upstream delay is long enough (at least 165 minutes) without any flight delay recovery actions, the delay is likely to be passed on or increased. Results show that, in terms of the delays related to airline and ATC factors, when the delay time of flights A and E exceeds 165 minutes, there is a high probability that the delay time of flights B and F will also exceed 165 minutes.

Figure 9(c) shows the probability of delay effects on the downstream flight influenced by airline, weather, and ATC factors. The probability of the delay effects under the ATC factor has an overall increasing trend. That is, delays due to ATC from flight F tend to be higher relative to flight E, especially when delays from flight E increase. In contrast, the trend for the probability under the airline factor remains relatively constant.

4.5. Sensitivity Analysis. In order to examine the influence of ATC and airline factors on the level of flight delay propagation, a sensitivity analysis was conducted, focusing on the possible effects of increasing the buffer time in flight schedules. Two scenarios that adjust the original flight schedule with an increase of turnaround buffer time by 10 minutes and 20 minutes, respectively, are created. The degrees of delay reduction of the adjusted schedules relative to the original schedule without changing buffer time are then compared. Concretely, the original flight delay data are processed as follows: if the turnaround time increases by 10 minutes, then the downstream flight delay time decreases by 10 minutes. The result is assumed to be zero if the original downstream flight delay time is less than 10 minutes, assuming that delays can be fully absorbed by scheduled buffer time.

Figure 10 shows the delay reduction probability for an increase in the buffer time under ATC factors. The average delay reduction probability decreases from 30% to 25% when the buffer time increases by 10 minutes, while the average probability increases from 30% to 54% when the buffer time increases by 20 minutes. This suggests that the 10-minute extra time would have a small negative effect on the conditional probability, while the 20-minute extra time would have a positive effect. Therefore, it is suggested that 20 minutes should be added to the turnaround buffer time of flights that are frequently influenced by airline factors in order to reduce the downstream delays.

Table 3: Parameters and goodness-of-fit statistics of the five copulas.

<table>
<thead>
<tr>
<th>Sample group</th>
<th>Parameters and OLS</th>
<th>Gaussian</th>
<th>t-Copula</th>
<th>Gumbel</th>
<th>Clayton</th>
<th>Frank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Copula parameter</td>
<td></td>
<td>(k = 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.676</td>
<td>0.721</td>
<td></td>
<td>1.901</td>
<td>1.506</td>
<td>5.8942</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>2.882</td>
<td>2.139</td>
<td>2.297</td>
<td>8.854</td>
<td>2.158</td>
</tr>
<tr>
<td>2</td>
<td>0.381</td>
<td>0.469</td>
<td></td>
<td>1.402</td>
<td>0.919</td>
<td>2.854</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>0.159</td>
<td>0.141</td>
<td>0.052</td>
<td>0.497</td>
<td>0.142</td>
</tr>
<tr>
<td>3</td>
<td>0.685</td>
<td>0.713</td>
<td></td>
<td>1.993</td>
<td>1.461</td>
<td>5.398</td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>0.634</td>
<td>0.658</td>
<td>0.317</td>
<td>2.492</td>
<td>0.786</td>
</tr>
</tbody>
</table>

5. Conclusion

This paper investigated the correlation between sequential flight delays under the delay factors of airline, weather, and ATC. It also examined the effects of buffer time on delay propagation in flight scheduling in order to reduce delay propagation. The main findings of this paper can be summarized as follows.

First, different from the previous studies, the correlation of delay propagation between sequential flights by using the copula function can be analyzed by dependence coefficients $\lambda_{up}$ and $\lambda_{lo}$. The lower tail $\lambda_{lo}$ showed that when delay from upstream flights is short, the factors of weather, airlines, and ATC have small impacts on the delay of downstream flights. In addition, the impact of airline factor is found to be the largest; it can be understood that when the delay between sequential flights is short, flight scheduling (especially the scheduling of buffer time) from airline plays an important
role in delay propagation. Meanwhile, the upper-tail dependence coefficients $\lambda^{\uparrow}$ showed that, under all three factors, delay from upstream flights may have an increasing effect on its downstream flights and will finally become a constant impact, which indicates limited influence of the three factors. As the delay from upstream flights gets larger, the factor of ATC represents the largest impact on delay propagation, which means that ATC has the largest influence from external environments.

Second, according to the conditional probabilities of delay propagation, flight scheduling can be effectively improved through reducing the uncertainty under different delay factors. With different scheduled buffer time, it is found that, for the flights that are frequently influenced by airline factors, 20 minutes should be added to the turnaround buffer time for delay reduction. For flights that are frequently influenced by ATC factors, adding 10-minute buffer time would be sufficient for delay reduction.

As presented above, the copula function used in this study is useful and capable of analyzing the causal structure of sequential delays. Conclusions were drawn from each of the influential factors. However, this may not identify the possible joint effects of multiple factors. Future works therefore are necessary to further investigate flight delay correlation under the combined delay factors, such as weather, airline, and ATC.

Conflicts of Interest

The authors declare that they have no conflicts of interest.
Acknowledgments

Financial support from the National Natural Science Foundation of China (71201081) and the Fundamental Research Funds for the Central Universities (NZ2016109) is gratefully acknowledged.

References
