

Spatial Delay Distributions in Chinese Airline Networks

Max Z. Li*, Karthik Gopalakrishnan*
Massachusetts Institute of Technology, Cambridge, MA, 02139

Yanjun Wang†
Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 211106, China

Understanding a variety of performance characteristics is a crucial first step in improving efficiency and managing delays in airline networks. Given the diversity of operating conditions and constraints, particularly across different national airspace, there can be important insights to be gained by analyzing different national airspace and airlines. In this work, we apply several data-driven techniques to characterize various airport delay distributions in Chinese airline networks. Specifically, we focus on examining two central characteristics of air transportation delays: its impact on the system and where the impacts are distributed spatially. We show that it is not enough to simply examine the former, as many operational impacts of delays can only be found through also considering spatial factors of airline delays. We focus on the airport delays at major Chinese airports incurred by the four largest domestic airlines in China (Air China, China Southern, China Eastern, and Hainan), and provide insights into aspects such as temporal trends in delay magnitudes versus spatial distribution, as well as outlying days in terms of delay impacts. We note that the focus of this work is on applications and operational impacts, rather than a methodological deep dive.

I. Introduction

Air transportation is responsible for connecting vast geographies by transporting over 4 billion passengers worldwide and approximately 60 million tons of freight annually [1]. Several airlines operate flights that help achieve connectivity at such a large scale. Thus, the air transportation system can be thought of as a multi-layer network, where each layer comprises a specific airline’s sub-network. Such a multi-layered network view of the system helps us understand and characterize a major source of inefficiency in the system: flight delays. While flight delays are very common (resulting in over 9000 cancelled flights every week worldwide [2]), analyzing, modelling, and predicting this highly dynamic and unpredictable system has remained challenging and is an active area of research. It should be noted that some delays are a sign that the system is working close to capacity, but excessive delays coupled with large variations in delay dynamics indicate inefficiencies that may require corrective measures. Furthermore, airlines can have very different operating practices such as the routes flown, crew scheduling practices, network structure (hubs, focus cities, etc.), and preference for cancellations versus delays, among other factors. Thus, our work in this paper analyzes the delay characteristics of the aviation system from the perspective of each airline, offering a fresh and potentially more insightful perspective.

Modeling and characterizing the behavior of air traffic delays post airline-specific decomposition can reveal delay dynamics that might be otherwise hidden in a system-wide analysis. For example, localized disruptions such as thunderstorms and airport-specific outages may affect individual airlines differently than the system as a whole. These observations and insights could be useful not only for dispatchers and flow managers at individual airlines, but also system-wide flow managers and air navigation service providers. Understanding the fundamental aspects of how delays behave within an air transportation network and its sub-networks is crucial for designing and deploying both strategic and tactical initiatives that seek to manage operations.

A. Motivation and problem description

In this paper, we will focus on two aspects of air traffic delays: how bad are delays in terms of their magnitudes, and how are delays at an airport relative to delays at other airports. More formally, the former analysis pertains to the *scale*, or *magnitude* of delays, whereas the latter analyzes the *spatial distribution* of delays within a network of airports [3]. Intuitively, the spatial distribution of airport delays describes *where* in the system the delay is located; this is distinctly

*PhD Candidate, Department of Aeronautics and Astronautics, 70 Vassar St., Cambridge, MA, 02139. AIAA Student Member.

† Associate Professor, College of Civil Aviation, 169 Shengtai Xilu, Nanjing 211106, China.

different from observing the impact, or magnitude of the delays within the network. For example, a total delay of D_{tot} minutes within a system of N airports could be exactly the same, but distributed in completely different ways. One extreme would be equal amounts of delay (D_{tot}/N minutes) at all airports, whereas another extreme would be zero delay at all but one airport, where that one airport has D_{tot} minutes of delay. The metric of total delay, or total system impact, or even aggregated delay in a small subset of airports (e.g., only focus on the hub airports of one particular airline) have been the focus of air traffic management. As our example with N airports shows, this total delay metric is not enough to differentiate between the two extreme scenarios. We depict in Figure 1 an equivalent scenario that demonstrates the inability of the total delay metric to distinguish between different spatial delay distributions.

Analyzing the relative spatial distribution of delays is particularly important. This stems from the fact that the air transportation system is highly interconnected, and thus delays at any one airport should not be viewed in isolation as it may be the result of delays elsewhere, and can potentially also propagate to other airports (i.e., independent versus dependent delays [4]). Thus, traffic flow managers would plan proactive measures not only based on the local delays at one airport, but also the network-wide distribution of delays. Furthermore, an airline-specific analysis of the distribution of delays will provide insights for airline schedulers and network managers regarding the current state of the system, as well as offer potential suggestions on modifying route structures, schedules, and buffer allocations.

B. Outline

For the remainder of this work, we focus on the analysis of airport delays at 30 major airports in the Chinese air transportation network over a span of 5 years (2012-2017) from the airline-specific perspective of four major carriers: Air China (CA), China Southern (CZ), Hainan Airline (HU), and China Eastern (MU). We highlight the reasons behind the choice of analyzing the Chinese air transportation system as well as the major airlines and stakeholders there in Section II. In Section IV, we give a high-level overview of the data-driven analysis techniques used. For the main results and ensuing discussions, we will compare airline performances in terms of the magnitude of delays (Section VI) as well as the spatial distribution of delays (Section VII). Finally, we present some conclusions and directions for future work in Section VIII.

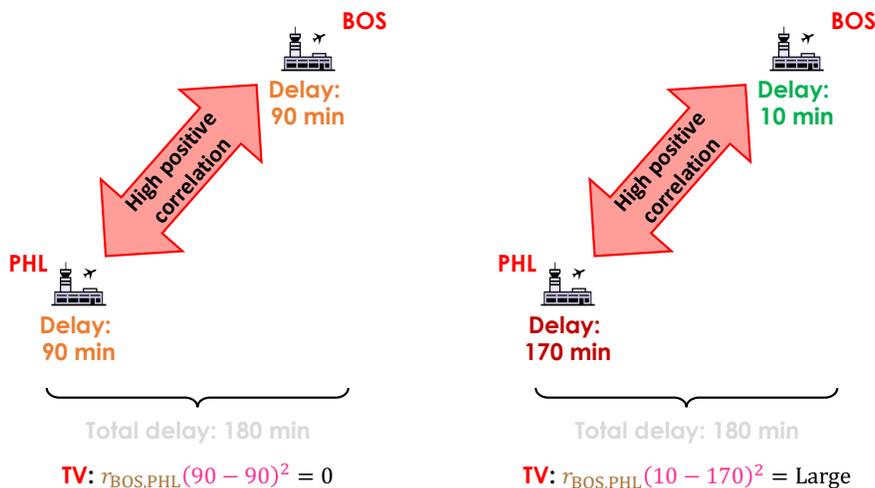


Fig. 1 Both scenarios have a total delay of 180 minutes, but the scenario on the left has a *smoother* spatial distribution of delay, whereas the scenario on the right does not. We use the *Total Variation* metric of graph signals to detect this difference in the spatial distribution of the delay.

II. Air Transportation in China

China is the world's second-largest aviation market, with around 560 million passengers transported in 2017 [5]. This number is comparable in magnitude to the US, which is the largest market serving almost 890 million passengers. More interestingly, with a projected growth rate of 5.3% per year for the next 20 years, it is significantly higher than the

	Air China (CA)	China Southern (CZ)	Hainan Airlines (HU)	China Eastern (MU)
Passengers (2019, Mil.)	115	152	82	130
Fleet size	429	619	223	567
Primary hubs (IATA code)	PEK, PVG, CTU	CAN, PEK	PEK, HAK, XIY, SZX	PVG, SHA, KMG
Carrier type	Full service	Full service	Full service	Full service
Market share (2017)	18%	23%	14%	17%
Average cancellation (2017)	4.2%	5.2%	4.6%	5.4%

Table 1 Overview of network and operating characteristics for the four Chinese airlines we consider in this analysis. Please refer to Table 3 in the appendix for a list of the IATA codes and full airport names.

global average of 3.8% [5]. To give a sense of scale, China contributed to the largest increase in domestic air traffic among all parts of the world with an increase in 48.8 million passenger journeys in 2018 [6]. Such unprecedented growth, along with the massive scale of the system, motivates the analysis of inefficiencies such as flight delays in the system. Currently, the average arrival delay for flights at the top 10 airports (based on traffic volumes) in China is 26 minutes [7]. Note that this is a substantial amount of arrival delays compared to other mature markets like the US, which has an average arrival delay of 17 minutes per flight for their top 10 airports [8]. With the ever increasing traffic, the problem of excessive flight delays is only expected to grow in China.

Two main factors make the Chinese air transportation system more vulnerable to delays: First off, the limited availability of airspace resources creates a capacity-side constraint; secondly, high-traffic routes are concentrated in a relatively small geographical region. To expand upon the first constraint, we note that in China, a vast majority of the airspace is blocked for military use, making only about 25% of the airspace available for civilian use. As a consequence, flights do not get to follow optimal routes, and create congestion at airway choke points in order to navigate around restricted airspace.

Expanding upon the second point of concentrated traffic densities, most of the traffic are distributed in the east part of the country, resulting in collocated regions of congested airspace. Among the top 15 domestic passenger traffic origin-destination pairs, six of them belong to China (Beijing-Capital and Shanghai-Hongqiao; Chengdu and Beijing-Capital; Beijing-Capital and Shenzhen; Shanghai-Hongqiao and Shenzhen; Guangzhou and Shanghai-Hongqiao; Guangzhou and Beijing-Capital) [9]. All of these are routes in the eastern part of the country. Thus, systemic high traffic in that region, combined with adverse meteorological phenomena that typically affects the east part of the country, result in large disruptions of airline and airport operations. Additionally, it is also worth mentioning that certain social events like the Golden Week of national holidays have shown to significantly increase demand as well as delays and cancellations [10].

In this paper, we focus our analysis on four major air carriers in China. They are Air China (CA), China Southern (CZ), Hainan Airlines (HU), and China Eastern airlines (MU). Together, these four full-service carriers represent nearly 70% of the market share in the country. Of these four, China Southern is the largest air carrier, with almost a quarter of the market share in the country. A high level summary of these four airlines are presented in Table 1. An important point worth noting is that all four of these carriers operate a hub-based network structure, with Beijing Capital International Airport (PEK) being a hub for three of them. While each of the four carriers have on-time performance metrics that could be improved upon, our work highlights the differences among these four airlines in terms of the temporal and spatial extent of delays they experience. Several of our results also provide airline comparisons in terms of temporal trends (i.e., when outlying disruption days occur), as well as which group of airports are typically affected in such cases.

III. Literature review

Our analysis and characterization of delay distribution within Chinese airline sub-networks adds to the growing body of literature focused on improving the efficiency and performance of the aviation system in China. Furthermore, such analyses are of benefit to the wider air traffic management community, as broader, translatable insights could be shared and adapted. We will now provide a non-exhaustive overview of some existing literature related to data-driven characterizations of Chinese airline networks. [11] investigates the evolution of the Chinese airport network (CAN), including its topology and traffic. It was found that the traffic grows at an exponential rate with seasonal fluctuations, while the network topology does not change significantly. A k -core decomposition method was used in [12] to divide the CAN into multiple layers. The *core layer* contained 17.7% of airports within the CAN, most of which were located in economically developed cities. Furthermore, this core layer was found to be a densely connected network, sustaining most of the traffic flows. Comparisons between the Chinese air transportation network and its US counterpart have also been made [10, 13], albeit at different granularities: [13] focused on using flight-specific trajectory data to characterize air traffic networks, whereas [10] analyzed differences in airport-level delay distributions. From these comparative studies, some notable conclusions include the fact that compared with the US network, China has more restricted airspace availability for commercial flights and faces a greater risk of en route congestion [13]. At the level of all airlines, the airports responsible for unusual spatial delay distributions in China were found to be much less geographically clustered than those in the US [10]. A key extension that we pursue in this work is deriving airline-specific insights at the level of airport delay distributions for the CAN.

The use of graphs as a helpful abstraction for network analysis has been studied in the context of the CAN and the Chinese aviation system. [14] conducts community detection on Chinese airport delay correlation networks in which edges are built using correlations of the delay time series between two airports to reveal implicit delay propagation relationships. [15] found that inter-country differences in how scheduled block times (SBTs) are set can account for large differences in on-time performance between two countries. They indicate that there are restrictions on SBTs for flights passing through congested airways in peak hours, out of concern that longer SBTs will lead to excessive airspace occupancy. Finally, [16] developed a number of operational resilience measures, and used them to assess how different scales and intensities of disruptions at a given airport affect aviation operations throughout the network.

In terms of methodology, the delay magnitude analysis utilizes standard measures of performance, such as average airport delays and associated quantile-based measurements. The spatial distribution analysis has more of a frequency-based flavor, where we utilize a graph frequency approach. This spectral approach to characterizing air traffic behavior extends from a line of work focused on frequency-based [17] and spectral graph wavelet-based [18] methods for analyzing airspace and sector performance. Specifically, we rely on our previous work in [3] for the paradigm of analyzing disruptions through a spatial distribution lens, [19] for a theory of outliers in graph-supported signals, and [10] for non-parametric approaches to generating graph signal outlier detection bounds. Furthermore, an interpretation framework for spatial distribution outliers is also presented in [10], and grounds the results and discussions found in Section VII.

IV. Methodology

In terms of characterizing delay magnitudes (Section VI) versus its spatial distribution (Section VII), the methodologies we use are distinct. The methods we use for the analysis of Chinese airline delay magnitudes can be considered standard ways of measuring system performance in an air traffic management context. For example, our airline- and airport-specific delay summaries are quite similar to how airlines and air navigation service providers might keep track of on time performance metrics. When we shift to analyzing the spatial distribution of delays across each Chinese airline’s networks, the core methodological framework we use leverages techniques from graph signal processing and the notion of a frequency-domain analysis over graphs, i.e., the graph Fourier transform (GFT). In particular, the GFT provides a way to decompose the airport delay signal into components of varying “frequencies” within our airport network. Intuitively, more unexpected spatial delay distributions have higher frequencies and variations, whereas more expected spatial delay distributions are “smoother” across the graph (see Figure 2 for a conceptual illustration).

To formalize the process of analyzing spatial delay distributions in airline networks, we consider airport delays to be a signal on a graph, where the nodes are airports, and the edge weights are the historical correlation in the delays at the two airports. The Total Variation (TV) of the airport delays with respect to the graph is a measure of the “smoothness” of the signal. Our prior work relates the TV of the airport delays to conformance of the particular delay distribution to historical trends [3]. Furthermore, our theoretical contributions in [19] provide mathematical as well as operational interpretability as to why a data point was classified as such. We define two independent dimensions along which we

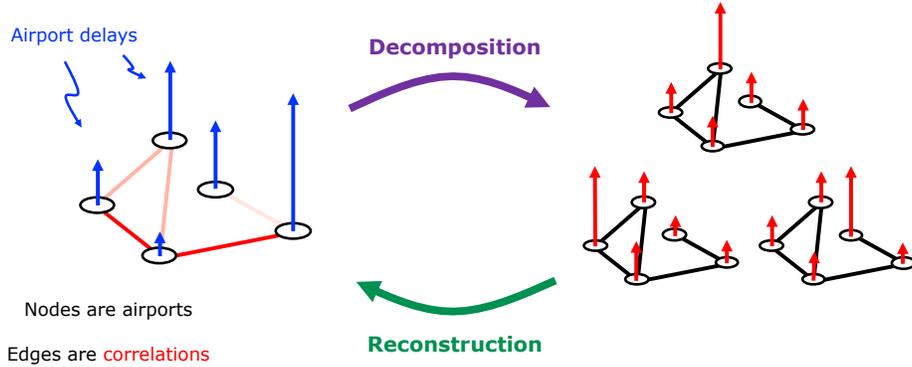


Fig. 2 A conceptual illustration of decomposing a graph signal into its graph frequency components.

can classify a delay distribution as an outlier: one based on a classical measure of delay magnitude, referred to as an *Outlier in Scale (OIS)*, and another based on our novel measure of distribution, namely *Outlier in Distribution (OID)*. Specifically, if the observed delays in an airline’s network can be classified as one the following three types of outliers (note that these classifications are not mutually exclusive), the interpretation would be as follows:

- (i) **OIS**: The total amount of delays incurred in the system is either abnormally high, or abnormally low, compared to historical occurrences.
- (ii) **Weak OID**: The spatial distribution of delays (i.e., the airports across the network that are experiencing and reporting flight delays) across the network is either abnormally smooth (i.e., we see a relatively even distribution of delays across the network), or abnormally uneven (i.e., there are large variations in delay magnitudes across pairs of highly correlated airports), *taking into consideration the magnitude of delays in the system*.
- (iii) **Strong OID**: The spatial distribution of delays across the network is either abnormally smooth or abnormally uneven, *without taking into consideration the magnitude of delays in the system*.

Note that the interpretations for weak versus strong OID are nearly identical, with the exception of whether or not the delay magnitudes comes into play. This is because the way we measure the spatial distribution of delay (i.e., the total variation) is a function of the delay magnitudes. Delay distributions marked as weak OID are outliers in the spatial distribution sense without taking into account that the delay magnitudes could have confounded the analysis, whereas the way we detect for strong OID takes into account this relationship, normalizes it, and removes the confounding factor. For a more technical description of these outlier definitions, as well as a more detailed overview of using spectral analysis for examining aviation delay networks, please refer to [3, 10, 19]. Our focus in this paper is on applying the notions and techniques developed in [3, 10, 19] to Chinese airline-specific delay data in order to gain a better sense of their spatial and magnitude characteristics.

V. Data processing

We obtain airport delay data, conditioned on the four Chinese airlines, for the 30 Chinese airports considered in our analysis (see Table 3 for a list of these airports and their IATA codes, as well as Figure 10 for a geographic plot of these airports) from the Operations Monitoring Center of the Civil Aviation Administration of China (CAAC). The time frame of the airline-specific airport delay data spans January 1, 2012 through December 31, 2015; the data granularity is one observation per day, equaling $N = 2,192$ observations per airline, or 8,768 daily airport delay observations in total. We can abstract each daily delay observation for an airline as a non-negative multidimensional vector $\mathbf{x}^{(t)} \in \mathbb{R}_{\geq 0}^{30 \times 1}$, where the i^{th} component x_i gives the delay (in minutes) at airport i for day t . Later on, particularly in Section VII when we discuss delay signal distributions on top of our airport network graph, we can think of each airport delay signal x_i as being supported on the i^{th} node in our graph $G = (V, E)$, where V is the set of all 30 airport nodes, and E is the set of all edges connecting each airport node.

For the delay magnitude analysis in Section VI, the data processing is straightforwardly performed on $\mathbf{x}^{(t)}$,

conditioning on the appropriate airline, airport, or time period. We will need to build out the graph abstraction in order to perform the spatial delay distribution analysis in Section VII, i.e., we need to construct how the dependencies within our airport network G are modeled. Following [3, 10, 19], we use the sample Pearson correlation coefficient as a measure of dependence between pairs of airports in V . Specifically, this gives us a first-order measurement of linear dependencies between airport delays. One way to represent this is through an *adjacency matrix* A that describes the 30×30 connections within the airport network G . This adjacency matrix A will have 30 rows and columns, with each element a_{ij} being the sample Pearson correlation coefficient between airport i and j . Note that since Pearson correlations are symmetric quantities, we have that $a_{ij} = a_{ji}$. After computing all a_{ij} 's, we can plot them geographically in Figure 3 to get a sense of high and low correlation areas within each airline's graph.

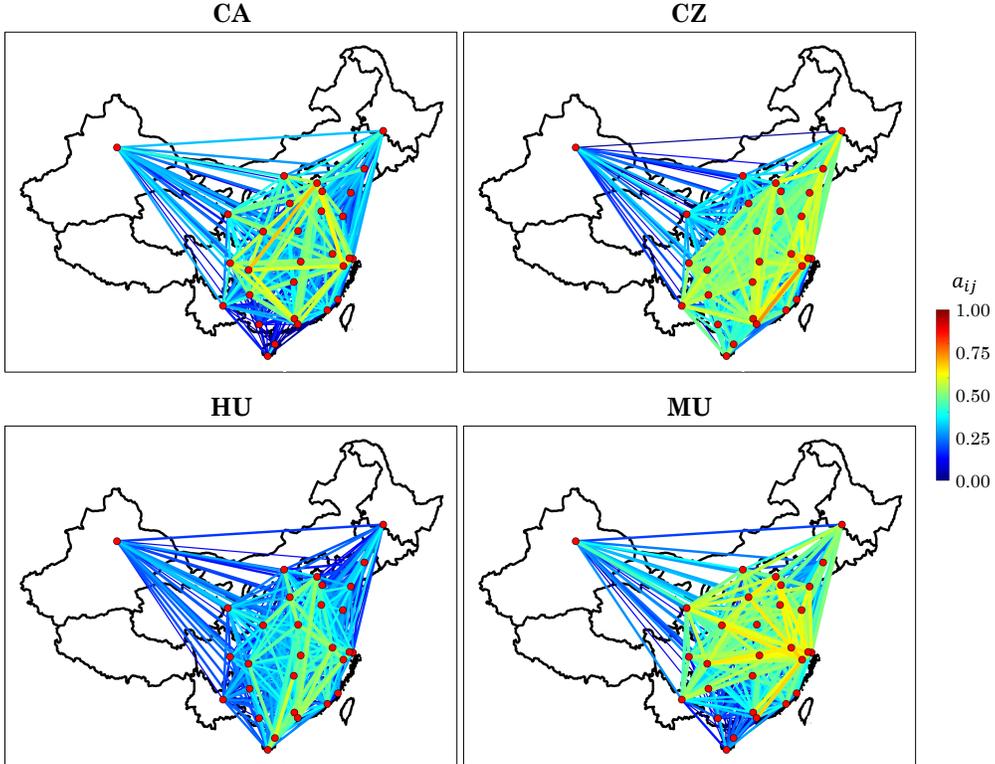


Fig. 3 Correlation networks superimposed on a geographical map of China for the four Chinese airlines we analyze. Note that higher correlation values are also emphasized with wider lines.

This adjacency matrix A (note that there will be four different A 's, one for each airline's network) is what will support the entirety of the spatial distribution analysis. Specifically, it is essential in computing the quantity of the *total variation* of $\mathbf{x}^{(t)}$; an illustration of computing one term within this total variation metric can be found in Figure 1. Note that in Figure 1, the pairwise term between one airport (e.g., PHL) and another airport (e.g., BOS) is given by $r_{\text{BOS, PHL}} \times (90 - 90)^2$. In other words, the contribution to the total variation is some measure of dependence $r_{\text{BOS, PHL}}$ multiplied by the *squared difference* of airport delays. In this analysis, we have that a_{ij} is precisely the dependence measure, and the total variation $\text{TV}(\mathbf{x}^{(t)})$ across an observation $\mathbf{x}^{(t)}$ can be written as:

$$\text{TV}(\mathbf{x}^{(t)}) = \frac{1}{2} \sum_{i \neq j} a_{ij} (x_i - x_j)^2. \quad (1)$$

The equation for total variation can be thought of as taking each pairwise combination of airports i and j , and summing over each pair's contribution, with a one-half scaling factor to account for double-counting. There are deep connections between the total variation and the adjacency matrix, as well as how specifically the total variation metric can be used for detecting outliers. Please refer to [3, 19] for a more in-depth technical discussion. With this concept of total variation in mind, we can actually rewrite the three outlier definitions from Section IV in a more quantitative manner as follows:

- (i) **OIS**: The (appropriate) norm of the delay observation $\mathbf{x}^{(t)}$, i.e., $\|\mathbf{x}^{(t)}\|_1$, is abnormally high or low compared to historical occurrences.
- (ii) **Weak OID**: The total variation of the delay observation $\mathbf{x}^{(t)}$ is abnormally high or low compared to historical total variation.
- (iii) **Strong OID**: The total variation of the delay observation $\mathbf{x}^{(t)}$, conditioning on the observed norm $\|\mathbf{x}^{(t)}\|_1$, is abnormally high or low compared to the historical conditional total variation.

Although Section VI focuses more on magnitude-based, or scale-based outliers (i.e., OIS) and Section VII focuses on strong OID, we give contingency table summaries of all outlier occurrences for each of the four Chinese airlines in Table 2.

VI. Delay magnitude analysis

In this section we analyze the delay magnitudes for each of the four airlines at the system-wide as well as airport-specific level. This approach is similar to the more conventional outlier analysis, where the magnitude of delays, and the deviation of the delay magnitude from the norm is used to infer properties and state of the system.

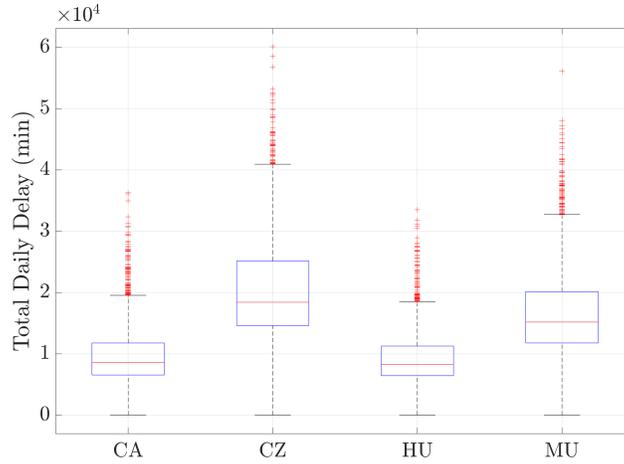


Fig. 4 Box plot of total daily delays in minutes for all four Chinese airlines.

First, we compare the the total system-wide delays for each of the four airlines. The resulting box plots are shown in Figure 4. For each airline, 2,192 total daily delay observations are used to create a box plot, which denotes the median delay, the 25th and 75th percentile, as well as outliers (points outside of the canonical $1.5 \times \text{IQR}$, where IQR stands for the inter-quartile range). Note that each outlier point actually correspond to a day that was an outlier. From Figure 4, we observe that HU marginally has the lowest median delay among the four carriers, and CZ has the highest. Interestingly, the 25th and 75th percentile boxes are almost equidistant from the mean, indicating that the distribution is nearly symmetric around the median. The width of the box plots indicates the variability in the daily delays – CZ remains the most variable airline in terms of delay performance in the 2012-2017 period. Each of the red crosses are days when that particular airline was an outlier in terms of the total daily delay. In the next section, we present more detailed results on the identification of these outlier days and compare their numbers across airlines. Also, any particular day may result in more than one airline being classified as an outlier due to the total system delay. This is more common than not due to externalities such as severe weather affecting all airlines that fly within a given region. Finally, it is worth pointing out that such an analysis of days with abnormally high delays, while insightful, leaves us with more questions than answers. Specifically, these precipitating questions include when do these outliers occur, which airports are affected, and what factors are common triggers.

We attempt to answer one such question pertaining to the airports that are typically affected using Figure 5. Here, we plot the airport-specific delays for each of the four airlines. It is not surprising to note wide variability in delay distributions across airports (due to varying degrees of demand-capacity imbalances) as well as airlines (due to their network structure and operational practices). The x -axis (airports) is sorted according to the traffic volumes (number of enplanements), where the airports on the left have higher traffic than the ones on the right. As one moves right along the

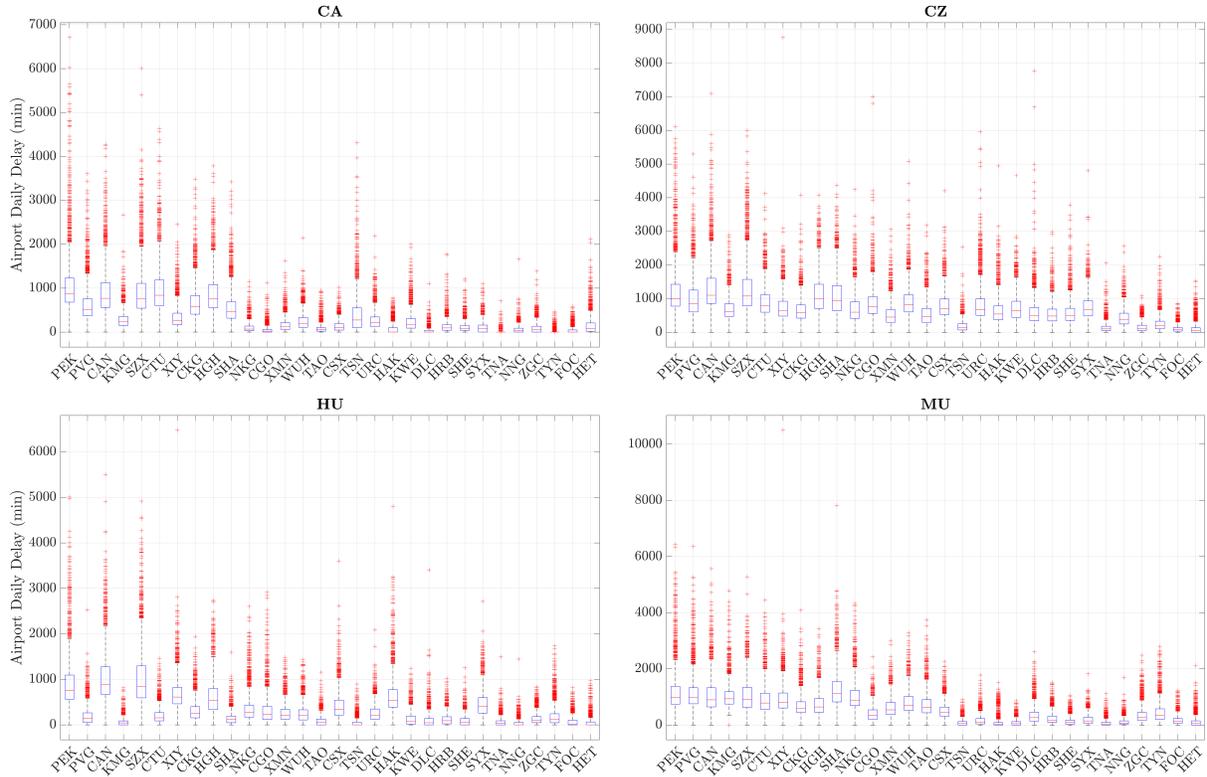


Fig. 5 Box plot of daily delays in minutes for all four Chinese airlines, split by airport.

x -axis, one would expect the airports to be less busy, less congested, and thus incur lower delays. However, this is not monotonically the case, as we see several notable exceptions to this trend (e.g. TSN having high delays for Air China, or HAK having a high delay for Hainan Airlines). In Figure 5, the outlier points denote days where that particular airport had unusually high delays. While this analysis helps identify days when a particular airport has unexpectedly high delays, it hides the fact that many of these outlier occurrences across airports are *correlated* and interdependent. This is a result of the strong network effects in the aviation system, where it is typically not just one airport, but rather a group of airports, that are affected by flight delays. This can happen either because an external disruption such as severe weather was geographically widespread (e.g., independent delays), or due to the spread and propagation of delays within the system (e.g., dependent delays). Our analysis in the next section highlights some of these network effects and captures outliers in the system while accounting for these inherent correlations between the delays at different airports.

The last findings of this section is presented in Figure 6, and shows the monthly distribution of delays. In Figure 6, we plot the monthly distribution of average daily system delays for each airline. The data indicates that summer months, in particular June, July, and August, are associated with higher delays for each airline. The increase is very prominent, with the delay minutes increasing almost by 100%. While increased summer traffic might dictate that the average delay per flight may not experience as dramatic of an increase, this still places a tremendous amount strain on the system as a whole. It is also reasonable to assume that a lot of outlier days (in terms of delay magnitudes) that are identified also fall within these three months. In the next section, we will present a surprising finding from the perspective of the geographical distribution of delays within the system – in particular, while the summer months typically have higher delay magnitudes than other months, the relative spatial distribution of delays is very much expected. In fact, it turns out that the winter months are the ones where the delays are lower than usual in terms of total magnitude, but are spatially distributed in unexpected ways. A deeper discussion on this is presented in Section VII.

To conclude, our analysis of delay magnitudes helps identify days, airlines, and airports with unusually high delay impacts, where “impacts” in this case is in the context of *how much* delay has accrued. However, a major limitation is the inability to capture correlations and dependencies between different airports, which limits our capabilities in finding network outliers in the system. This is a more nuanced question, and the next section is devoted towards identifying and analyzing outliers in distribution within the individual Chinese airline sub-networks.

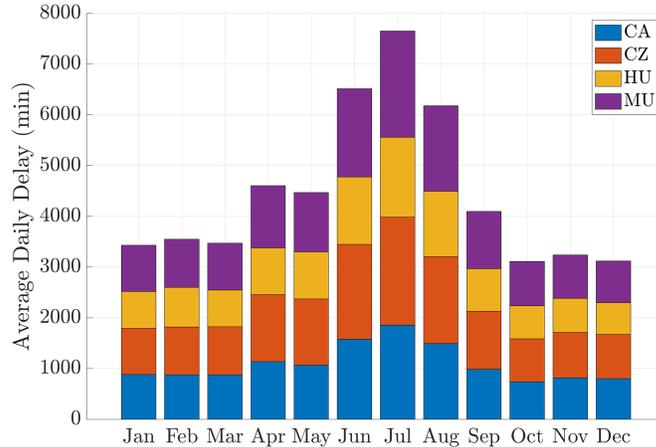


Fig. 6 Monthly distribution of average daily delays for all four Chinese airlines.

VII. Spatial delay distribution analysis

While characterizing the magnitude of airport delays is important and its operational implications are clear, given that air transportation operates across large geographical distances, simply observing the magnitude of delays leaves out a crucial component: its spatial distribution. In this section, we provide a complementary analysis to the previous section by analyzing *where* the delays are located within the Chinese airport network, with a focus on *unexpected* spatial delay distributions. Through the course of this analysis, we will utilize the graph signal processing-based disruption analysis techniques described in [3, 10], which will explicitly account for *network effects*, something that is difficult to do from the standpoint of analyzing delay magnitudes. We will first identify and examine sub-groups of Chinese airports that are commonly implicated in unexpected spatial distributions of delays, providing insights into which airports to strategically monitor the delays at. We then zoom out and identify individual outlier days in terms of their spatial delay distribution, through the use of *total variation* and total delay-based outlier detection methods formalized in [10]. Finally, we examine monthly distributions of days that are considered spatial delay distribution outliers, and contrast this with the monthly distribution of airport delay magnitudes (i.e., Figure 6). In particular, we will highlight the differences in monthly distributions when airport delays are observed from a magnitude versus spatial distribution standpoint, and comment on its operational implications.

When it comes to delay magnitudes, one could monitor the total delay in the system in comparison to some threshold to determine whether or not the system is in a high-delay state. Alternatively, an individual airline could monitor its subset of hubs and focus city airports. The selection of airports is less obvious when it comes to the question of monitoring for unexpected spatial distribution of airports. One approach is through decomposing a particular delay network into pieces, and analyzing individual *modes* that contribute to the observed airport delays on any given day. Recall from Section IV that we could examine the airline-specific *eigenvector modes* corresponding to each airlines' correlation network to get a sense of airline-specific airport subsets whose delays are often implicated in unexpected spatial delay distributions. In particular, we are interested in the *higher energy modes*, which we plot in Figure 7 for each of the four Chinese airlines.

These modes describe how the delays at particular airports are behaving with respect to delays at other airports. We provide an example of how to interpret eigenvector modes through examining Figure 7(a), i.e., the highest-energy mode for Air China (CA). The two airports highlighted in Figure 7(a) are PEK and HGH, the former is CA's main hub, whereas the latter is a focus city for CA. While it is not surprising that hubs and focus cities would be implicated in any sort of CA-specific disruptive event, what the eigenvector mode analysis tells us is that an extremely unusual spatial delay distribution in CA's airline network consists of delays at PEK trending *opposite to* delays at HGH. Specifically, it is an unusual situation when delays at PEK are increasing, but the delays at HGH are decreasing, or vice versa. Hence, for CA, monitoring changes in the relative differences between delays at PEK versus HGH is an indicator for whether or not the delays are becoming more (or less) unusually distributed across CA's network. The second-most energetic mode for CA is the state where PEK and HGH delays are trending opposite to CKG, another focus city for CA. Thus,

these eigenvector modes can be used to determine which (potentially overlapping) subsets of airports an airline should monitor, if that airline wishes to also measure the geographic variance of delays across their network.

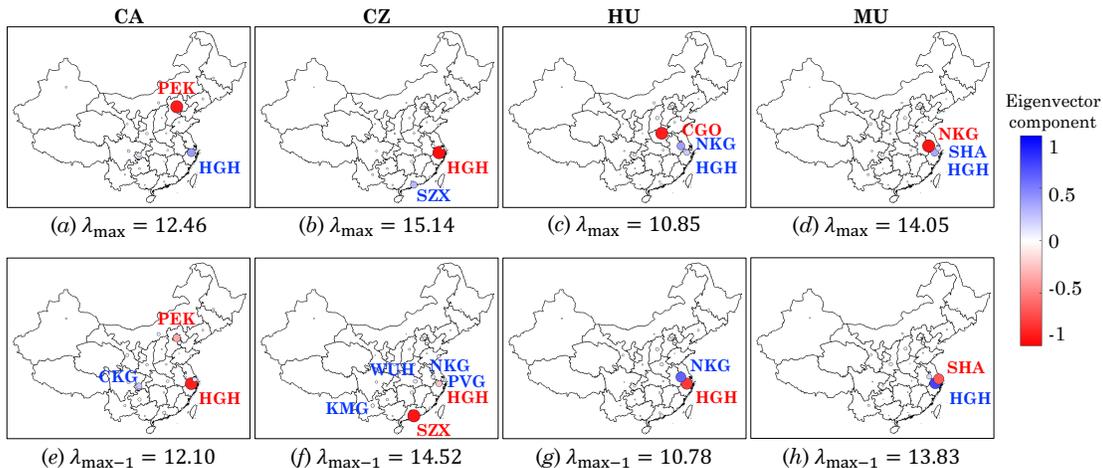


Fig. 7 Most (a)-(d) and second most (e)-(h) energetic eigenvector modes for each of the four Chinese airlines.

Note that besides CA, all high-energy modes for the other three airlines are concentrated in the southern and southeastern region of China. This is reflective of the network and hub structures of China Southern (CZ), Hainan (HU), and China Eastern (MU). One interesting observation is that HGH appears in every high-energy mode depicted in Figure 7, regardless of the airline. This indicates that regardless of the airline sub-network that one considers, it is likely that delays at HGH are either unusually high or unusually low if the network is in a state of having unexpected spatial delay distributions. This might be important at a central traffic management level, particularly since the importance of HGH seems to not be airline-specific. We note that HGH is a rare case where it is a focus city for all four airlines we consider in this study; it is possible that this attribute of being such a shared resource is what is being reflected in the eigenvector mode analysis. Finally, the airport subsets for HU and MU’s high-energy modes are quite closely collocated, whereas some of CZ’s airports are quite distributed throughout the southern regions of China.

Now that we have a sense of airports specific to each airline that are often implicated in unusual spatial delay distributions, we will examine the actual day-by-day occurrences of unexpected spatial delay distributions for each airline’s sub-network. We utilize the skewed interquartile range-based non-parametric outlier bound detection method detailed in [10] to classify the three different types of outliers summarized in Section IV: outliers in scale (OIS), weak outliers in distribution (weak OID), and strong outliers in distribution (strong OID). Recall that OIS describe days where the magnitude of delay – now taking into account network effects – is higher (or lower) than normal, whereas both weak and strong OID measure the spatial distribution of airport delays across an airline’s sub-network, with the latter being a magnitude-independent measure of spatial variation. We plot the total delay and total variation for each of the 2,192 days between 2012-2017 for each airline in Figure 8, along with the bounds demarcating whether or not a day is OIS, weak OID, or strong OID. Note that these outlier classifications are not mutually exclusive; for example, a day could be both an OIS as well as a strong OID, indicating that an extremely severe disruption occurred in that airline’s sub-network, affecting not only the magnitude but also the spatial variance of delays.

Based off of how many data observations (each observation represents a day in an airline’s network) lie outside of the various outlier bounds from Figure 8, we record the number of OIS, weak OID, and strong OID days for each airline, and report the outlier statistics in Table 2. Note that the figures quoted in Table 2 are in number of days, with percentages given by dividing the number of days by 2,192 days (i.e., the total number of observations between 2012-2017). We see that HU has the highest number of OIS days and weak OID days, as well as the second-highest number of strong OID days. If we condition out the magnitude aspect of delay and focus solely on its spatial distribution, then CZ claims the highest number of strong OID days, with approximately 5% of all days between 2012-2017 being classified as a strong OID. Referring back to Figure 3, we note that CZ has several edges with high correlations, concentrated in the southeastern region of China. Recall that since “unsmooth” distributions of delays can only occur across edges with high correlation (i.e., two highly correlated airports where one airport has low delays and the other has high delays, or vice versa), this indicates that CZ often experiences uneven delay distributions across highly correlated airport cliques,

Outlier Type	Air China (CA)	China Southern (CZ)	Hainan Airlines (HU)	China Eastern (MU)
OIS	17 (0.8%)	44 (2.0%)	52 (2.4%)	13 (0.6%)
Weak OID	37 (1.7%)	24 (1.1%)	61 (2.8%)	32 (1.5%)
Strong OID	47 (2.1%)	116 (5.3%)	76 (3.5%)	54 (2.5%)

Table 2 Airline-specific outlier statistics; the number of outlier days as well as the percentage out of 2,192 days (2012-2017) are provided.

leading to a high number of strong OID days. This is true for CZ independent of how high the delays were (i.e., the magnitude of the delays).

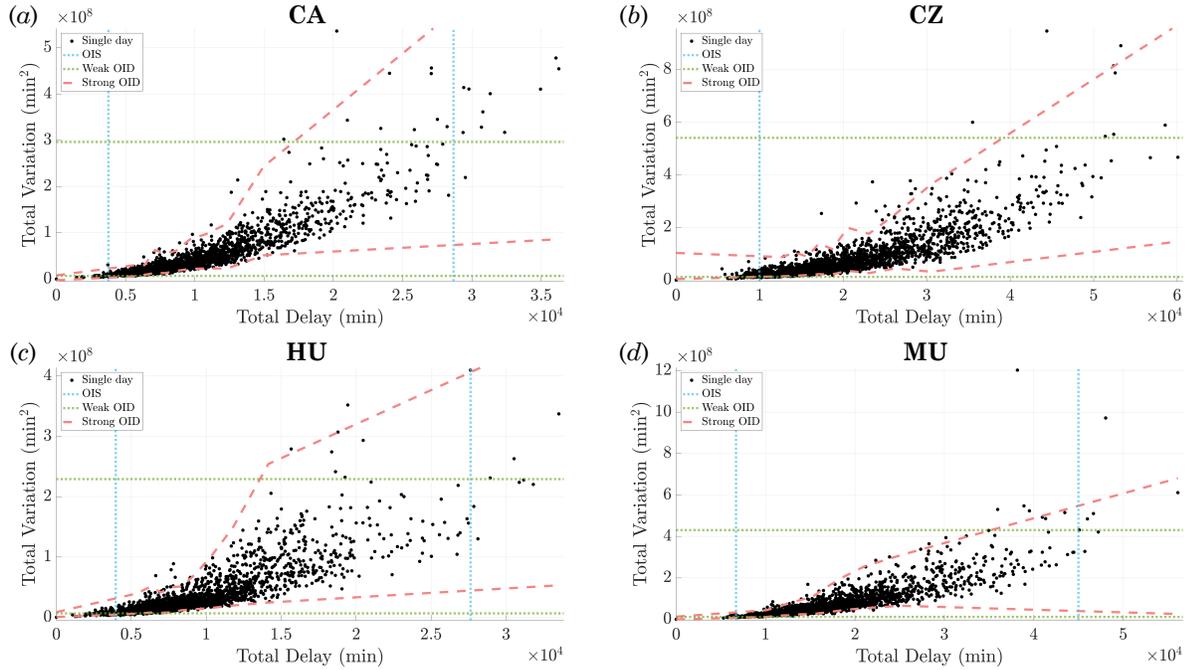


Fig. 8 TV versus TD for all days in 2012-2017 for (a) Air China (CA), (b) China Southern Airlines (CZ), (c) Hainan Airlines (HU), and (d) China Eastern Airlines (MU).

We can also assess the temporal trends of unexpected spatial delay distributions occurring in airline sub-networks; specifically, we will examine the number of strong OID days by months, and contrast this with the average delay magnitudes by month as shown in Figure 6. Note that we will only consider the strong OID counts in this analysis, because we want to focus solely on the effects of spatial variance, and explicitly exclude the effects of delay magnitudes. Recall that for delay magnitudes, we saw a unimodal distribution with peak average delays in the summer months. In particular, for the 2012-2017 time period, the month of July saw peak average delays across the four airline sub-networks. Interestingly, this is not the case with strong OID counts, as shown in Figure 9. This contrasting observation immediately suggests that characterizing delays according to their magnitude versus their spatial variance can result in highly orthogonal measurements of two different kinds of network performance: for the Chinese airline sub-networks, the periods of high delay magnitudes do not seem to coincide with periods of unexpected spatial delay distributions. In other words, whereas delays are typically high during summer months, they are spatially located at “expected” airports throughout each of the airline’s network. On the other hand, during fall and winter months, in particularly October, even if the magnitude of delays may be decreasing from its peak summer months, the *location* of these delays within each airline’s network is increasingly unexpected, and “unsmooth” across the network. One potential operational factor is that in the summer months, heavy convective activity cause severe levels of disruptions, but diffused in the network

in an expected manner. On the other hand, unusual demand profiles in months such as October could coincide with less severe disruptions, but result in delays that appear in unexpected locations throughout each airline’s network. An example of such an unusual demand profile could be during the Golden Week holidays in October, where the Chinese aviation system typically experiences large surges in travelers [10].

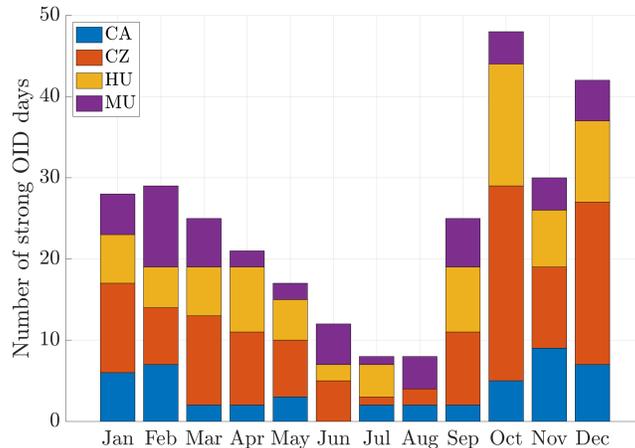


Fig. 9 Monthly distribution of strong OID days for all four Chinese airlines.

VIII. Conclusions

When it comes to improving on-time performance and becoming more efficient at designing and implementing air traffic management initiatives, a crucial first step is to gather as much information as possible regarding the historical performance characteristics of the delays in the system. In this paper, we leverage two broad perspectives to examine delays within individual airline sub-networks in the Chinese air transportation system. Specifically, the first perspective we take when analyzing Chinese airline delay distributions is that of characterizing the *magnitude* of incurred flight delays. This is the classical approach, and seeks to answer how much the system was impacted by delays on any given day and disruptive event. We then analyze the *spatial distribution* of delays across a network, an important characteristic given that air transportation naturally operates across large geographic extents. Intuitively, it is not only important to understand how much impact the delays had on the system, but *where* in the system the impacts were. In order to formalize the second class of analysis, we use a graph frequency approach, and examine both airport-specific decompositions of delay signals, as well as specific days with outlying delay distributions in each airline’s network. Throughout the paper, we emphasize the complementary nature of these two approaches, as well as their orthogonality. The latter point was particularly prominent when examining monthly trends in terms of average delays versus spatial delay distributions in Chinese airline networks: a disruption in terms of delay magnitudes may not necessarily indicate a disruption in terms of where the network delays were, and vice versa.

It is clear that there is a wide range of data-driven approaches to characterizing the behavior of delays within air transportation systems. The merits of this work not only highlights the need to perform analyses from different network perspectives, but also to conduct network performance analyses across different national airspace in order to facilitate cross-country learning and insights. An interesting future extension could be to compare airline-specific network performance measures across different national airspace and Flight Information Regions, potentially revealing advantageous operational nuances that could be more broadly adopted. Another smaller-scoped analysis could combine the system-wide delay distribution insights for the Chinese air transportation system derived in [10] with airline-specific sub-network performances derived in this study, allowing us to then comment on deleterious network interactions and potentially how to mitigate them.

References

- [1] “ICAO Annual Report,” <https://www.icao.int/annual-report-2018/>, 2018. Accessed: 2020-05-14.
- [2] “OAG statistics,” <https://www.oag.com/>, 2020. Accessed: 2020-05-14.
- [3] Li, M. Z., Gopalakrishnan, K., Pantoja, K., and Balakrishnan, H., “A spectral approach towards analyzing airport performance and disruptions,” *13th Air Traffic Management Research and Development Seminar*, 2019.
- [4] Chiraphadhanakul, V., and Barnhart, C., “Robust flight schedules through slack re-allocation,” *EURO Journal on Transportation and Logistics*, Vol. 2, No. 4, 2013, pp. 277–306.
- [5] “The Importance of Air Transport to the People’s Republic of China,” <https://www.iata.org/en/iata-repository/publications/economic-reports/china--value-of-aviation/>, 2018.
- [6] “IATA statistics,” <https://www.iata.org/en/publications/annual-review/>, 2019. Accessed: 2020-05-14.
- [7] Operation Monitoring Center, the Civil Aviation Administration of China, “Civil flights operation performance report,” Tech. rep., Civil Aviation Administration of China, 2018.
- [8] “USA Airport delay statistics,” <https://www.visualcapitalist.com/visualizing-the-u-s-airports-with-the-worst-flight-delays/>, 2020. Accessed: 2020-05-14.
- [9] “WATS World Air Transport Statistics 2019,” <https://www.iata.org/contentassets/a686ff624550453e8bf0c9b3f7f0ab26/wats-2019-mediakit.pdf>, 2019.
- [10] Li, M. Z., Gopalakrishnan, K., Wang, Y., and Balakrishnan, H., “Outlier Analysis of Airport Delay Distributions in US and China,” *IEEE Artificial Intelligence and Data Analytics in Air Transportation Conference*, 2020.
- [11] Zhang, J., Cao, X.-B., Du, W.-B., and Cai, K.-Q., “Evolution of Chinese airport network,” *Physica A*, Vol. 389, No. 18, 2010, pp. 3922–3931.
- [12] Du, W.-B., Zhou, X.-L., Lordan, O., Wang, Z., Zhao, C., and Zhu, Y.-B., “Analysis of the Chinese Airline Network as multi-layer networks,” *Transportation Research Part E: Logistics and Transportation Review*, Vol. 89, 2016, pp. 108 – 116.
- [13] Ren, P., and Li, L., “Characterizing air traffic networks via large-scale aircraft tracking data: A comparison between China and the US networks,” *Journal of Air Transport Management*, Vol. 67, 2018, pp. 181–196.
- [14] Chen, S., Wang, Y., Hu, M., Zhou, Y., Delahaye, D., and Lin, S., “Community Detection of Chinese Airport Delay Correlation Network,” *2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT)*, 2020, pp. 1–8.
- [15] Wang, Y., Zhou, Y., Hansen, M., and Chin, C., “Scheduled block time setting and on-time performance of U.S. and Chinese airlines—A comparative analysis,” *Transportation Research Part A: Policy and Practice*, Vol. 130, 2019, pp. 825 – 843.
- [16] Wang, Y., Zhan, J., Xu, X., Li, L., Chen, P., and Hansen, M., “Measuring the resilience of an airport network,” *Chinese Journal of Aeronautics*, Vol. 32, No. 12, 2019, pp. 2694 – 2705.
- [17] Drew, M. C., and Sheth, K., *A Frequency Analysis Approach for Categorizing Air Traffic Behavior*, chapter and pages.
- [18] Drew, M. C., and Sheth, K., *A Wavelet Analysis Approach for Categorizing Air Traffic Behavior*, chapter and pages.
- [19] Gopalakrishnan, K., Li, M. Z., and Balakrishnan, H., “Identification of Outliers in Graph Signals,” *2019 IEEE Conference on Decision and Control (CDC)*, 2019.

Appendix

IATA	Airport Name
CAN	Guangzhou Baiyun International
CGO	Zhengzhou Xinzheng International
CKG	Chongqing Jiangbei International
CSX	Changsha Huanghua International
CTU	Chengdu Shuangliu International
DLC	Dalian Zhouzhuizi International
FOC	Fuzhou Changle International
HAK	Haikou Meilan International
HET	Hohhot Baita International
HGH	Hangzhou International
HRB	Harbin Taiping International
KMG	Kunming Changshui International
KWE	Guiyang Longdongbao International
NKG	Nanjing Lukou International
NNG	Nanning Wuxu International
PEK	Beijing Capital International
PVG	Shanghai Pudong International
SHA	Shanghai Hongqiao International
SHE	Shenyang Taoxian International
SYX	Sanya Phoenix International
SZX	Shenzhen Bao'an International
TAO	Qingdao International
TNA	Jinan Yaoqiang International
TSN	Tianjin Binhai International
TYN	Taiyuan Wusu International
URC	Ürümqi Diwopu International
WUH	Wuhan Tianhe International
XIY	Xi'an Xianyang International
XMN	Xiamen Gaoqi International
ZGC	Lanzhou Zhongchuan International

Table 3 IATA three-letter code and corresponding full airport name of the Chinese airports included in our analysis.



Fig. 10 Geographic locations of the Chinese airports (IATA code given) included in our analysis.