An Empirical Study of the Resilience of the US and European Air Transportation Networks

Barış Başpınar, Karthik Gopalakrishnan, Emre Koyuncu, and Hamsa Balakrishnan

***Abstract*—Air travel connects people and goods across vast geographical regions. However, operational inter-dependencies in the air transportation system due to factors such as aircraft, crew, and passenger connectivity also result in the spread of disruptions in the system. Our work uses tools from network science and control theory to characterize the relation between the inter-connectivity (i.e., network structure, both in terms of flights and delays) and the robustness of the air transportation system. These methods are applied to characterize the resilience of the air transportation networks in the United States (US) and Europe by considering the flight and delay network structures and delay dynamics. Our study reveals that stronger inter-connectivity in the US makes the system more susceptible to disruptions that spread rapidly. However, we also find that this higher connectivity enables greater flexibility and controllability while recovering from disruptions.**

***Index Terms*—Air traffic management, air transportation networks, United States, Europe, delay propagation, robustness, resilience.**

I. Introduction

Air transportation is a major driver of global connectivity, offering significant economic benefits. However, increasing traffic and limited investments in infrastructure have led to increased flight delays and cancellations, adversely affecting system performance. Modeling and analysis tools are therefore needed to develop insights into the air transportation system, and to efficiently serve the increasing traffic demand.

The air transportation system displays a high degree of inter-connectivity, both in terms of flights and delays. The physical movement of aircraft between airports is referred to as *flight connectivity*. The spread of delays between airports, known as *delay connectivity*, can occur not only due to flight connectivity, but also due to factors such as crew connections and correlated weather impacts across airports. The temporal evolution of delay connectivity is referred as the *delay dynamics*. Flight connectivity enables passengers and goods to reach their destinations, and is realized through multiple non-stop flight legs. By contrast, geographical proximity, traffic flows between airports, and airline operating practices lead to delay connectivity.

© 2022. This manuscript version is made available under the CC-BY-NC-ND 4.0 license https://creativecommons.org/licenses/by-nc-nd/4.0/

DOI 10.1016/j.jairtraman.2022.102303

Barış Başpınar and Emre Koyuncu are with the Department of Aeronautical Engineering, Istanbul Technical University, Istanbul, 34469 Turkey (e-mail: {baspinarb,emre.koyuncu}@itu.edu.tr)

Karthik Gopalakrishnan is with the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, 94035 USA, karthikg@mit.edu

Hamsa Balakrishnan is with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA, hamsa@mit.edu

Graphs or networks are natural abstractions of the interdependencies in the air transportation system. Airports represent the nodes of the network, and the meaning of the edges can vary depending on the quantity of interest. In case of flight connectivity networks, the presence of an edge between two airports may represent the existence of at least one daily non-stop flight. Weighted, directed graphs, in which each edge is associated with a weight and direction, can be used to capture metrics such as delays, traffic, or flight cancellations from one airport to another. Consequently, network abstractions can be used to model the propagation of delays, assess the susceptibility of airports to delays, and evaluate the resilience of the air transportation network to disruptions.

In this paper, we are interested in evaluating the effect of network connectivity on delay dynamics. In particular, we would like to know how flight and delay connectivity influence the susceptibility of airports to disruptions, their post-disruption recovery, and the ability to regulate the system after a disruption.

The US and European air transportation systems are among the largest in the world, but differ in important ways [1]. Firstly, Europe, unlike the US, has slot controls at nearly all its major airports, potentially ensuring the smoothing of peaks in demand [2]. Secondly, the number of airlines operating in Europe is significantly higher than the US, resulting in more distributed operations, and lower susceptibility of the system to disruptions that affect just one airline or its hub airports. These differences in their operational characteristics make the US and European aviation systems ideal candidates for a comparative analysis of their underlying connectivity structures, and the effects on delay dynamics.

A. Related literature

The role of network structure in driving flight connectivity [3], [4], delay dynamics [5], [6], and system resilience [7] has been considered in prior studies. Flight connectivity networks, typically represented as graphs with an edge between two nodes (airports) if there is a direct non-stop flight between them, have received the most attention among these [8], [9], [10]. Similar analysis has been extended to weighted graphs, where the magnitude of traffic flows are also considered [11], [12], [13]. There is, however, limited literature on estimating the delay connectivity graphs, since it may require information on airline schedule buffers, tail assignment, and crew assignments. Delay networks, where edges are weighted by the average flight delays on the route, have been used to represent the state of the system. Representative delay networks have been identified through clustering in prior works [14], [15], [16]. The temporal evolution of these delay networks have also been studied [17], [5], [18].

Prior work on network resilience has focused on flight connectivity, and the loss in connectivity due to node removal attacks [3], [19], [20]. More recently, the susceptibility of US airports to delay propagation [21], and the recovery of European airports from delay-disruptions [22], have been studied. More generally, dynamic measures of operational network resilience have been restricted to specific contexts (e.g., epidemics [23], communication networks [24], financial institutions [25]). These studies illustrate that the underlying graph structure can result in different levels of robustness depending on the process that evolves on the graph [26]. Similarly, the robustness of aviation networks must be evaluated with respect to the delay propagation context – a gap in the literature that we aim to fill in this work. Finally, we illustrate our methods through a comparative analysis of the US and European air transportation networks.

B. Contributions and findings

Our main contributions are as follows:

1) We use network-theoretic measures to compare the connectivity of multi-layer networks (the network layers represent flights and delays, in our case).
2) We propose a novel approach and different metrics from other domains to assess the characteristics of the delay networks.
3) We develop a method to compare the resilience of networked systems based on the dynamics of disruptions and recovery.
4) We compare the US and European air transportation networks to distinguish their delay spreading characteristics and obtain operational insights.

Our comparison of the US and European air transportation networks leads to the following findings:

1) The US has stronger flight and delay connectivity than Europe.
2) This increased connectivity makes the US more susceptible than Europe to the spread of delays, and subsequent losses in flight connectivity.
3) Although the stronger connectivity in the US increases the propagation of delays, it also enables a faster recovery and a greater ability to regulate system performance.

C. Outline

For clarity of presentation, we analyze each aspect of connectivity separately, for the US and Europe. We discuss flight connectivity in Section II, delay connectivity in Section III, and delay dynamics in Section IV. The structure of the paper is constructed such that the case studies in different network levels support each other to reveal the correct operational insights. Additional details on the underlying network-theoretic measures are provided in the appendix.

II. Analysis of Flight Connectivity

First, we compare the structures of the traffic networks in Europe and US. The networks are represented by directed graphs, in which each node denotes an airport, and each edge is weighted by the average number of daily flights from one airport to the other. Flight data are obtained for a 1-year period from ALLFT+ [27] for Europe (2017), and from BTS [28] for the US (2015). The ALLFT+ dataset includes the transcontinental flights, thus allowing us to study the US and European networks separately, as well as jointly. We only include edges with at least one flight per day on average.

TABLE I: High-level comparison of the US and Europe air traffic networks

Graph Properties	European Network	US Network
# of Nodes	460	337
# of Edges	4781	3901
Avg. Degree	20.79	23.15
Avg. Strength	70.04	114.07
% of Core Nodes	21.1	19.3

Table I presents some basic properties of the European and US networks. The numbers of nodes and edges are higher in the European network, since there are more airports and routes served with at least one flight a day in Europe than in the US. However, the average strength, and to a lesser extent the degree, of nodes in the US is higher than that in Europe. In other words, the US, despite being a smaller network, not only has more connections per airport, but they are stronger (i.e., more flights between the nodes). This difference could be due to the dominance of regional operations and more operations to smaller airports in Europe. However, the percentages of core nodes in the US and Europe are similar, indicating that both networks are dominated by similar fractions of central nodes. Table VIII in the appendix lists the names and locations of the airports identified by their ICAO codes in our results and discussions.

The average number of flight operations, a commonly-used metric to capture the prominence of the airport, is captured by its node strength. A less frequently-used metric, the number of non-stop destinations from an airport, is captured in its degree. We present a list of the top 10 airports in the US and Europe based on their degree and strength centrality in Table II. We observe that the 5 airports with the highest degree centralities in the US all have higher degrees than any of the airports in Europe. Chicago (KORD) covers 41% $(277/(2 \times 337))$ of the US network with direct connections, whereas Amsterdam (EHAM) covers only 24% of the European network with direct connections. These results imply that the US network has higher reachability than the EU network, an observation that will be further supported when we consider the expansion metric.

A. Flight connectivity: Comparison of US and Europe

Firstly, we discuss the clustering coefficient and the rich-club parameter to understand the role played by airports in the entire network. These metrics have been selected to assess the local connection characteristics of the networks. While

TABLE II: Top 10 airports by degree centrality

European Traffic Network			US Traffic Network		
Airport	Degree	Strength	Airport	Degree	Strength
EHAM	223	1036.15	KORD	277	1968.15
EDDM	199	879.16	KDFW	273	1555.15
LTBA	186	731.94	KATL	270	2077.84
EDDF	179	893.96	KDEN	237	1287.84
LFPG	168	853.51	KCLT	224	1277.50
EGKK	168	588.81	KDTW	198	908.16
LEBL	157	704.16	KIAH	187	983.19
EGSS	156	359.91	KMSP	184	876.20
LEMD	154	757.54	KPHL	174	897.27
EKCH	142	593.80	KSEA	143	862.95

the clustering coefficient captures the tendency of forming triangles in the network, the rich-club parameter quantifies the tendency for high-degree nodes to be more densely connected among themselves than nodes of a lower degree. Specifically, the clustering coefficient is a relatively local metric for a node, and captures the tendency of the incident edges on a node to form triangles. Thus, if there are flights from airport A to B, and flights from C to A, this metric captures the likelihood of there being a direct flight from B to C. The route structure of airlines, and whether they operate a hub and spoke or point-to-point operation determines the clustering coefficient. More importantly, it has implications both for passenger flight connectivity as well as for delay propagation. Thus, when there is a reduction in airport capacity and limited schedule slack, a high degree of feedback in the system, captured by the clustering coefficient, can accentuate the spread of delays and magnify their impact. Edge weights are normalized to 1 for each of the respective networks before evaluating the coefficient, and the histogram of the node clustering coefficients are presented in Fig. 1. We observe that most of the nodes have a very small clustering coefficient, indicating a low tendency to form triangles. This is to be expected, since many peripheral airports will only be connected to one major hub, and not to each other through direct flights. Nevertheless, in the US, there is a slightly higher fraction of airports that tend to form connected triplets; this increases the potential for delay propagation in the absence of robust scheduling practices.

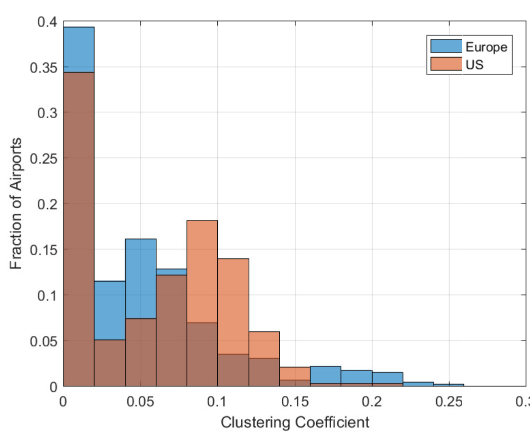


Fig. 1: Comparison of clustering coefficients of flight networks

The rich-club coefficients, which reflect the extent to which high degree nodes are connected to each other, of both flight networks are illustrated in Fig. 2. We see that the US network has a larger rich-club parameter than the European network, and both show a sharp decline around $d = 120$. This point can be interpreted as a transition to the high-degree nodes. Fig. 2 shows that the connections between the high-degree nodes in the US network are stronger than in the EU network. This also suggests that traffic flows in the US can be controlled through a smaller fraction of airports than in Europe.

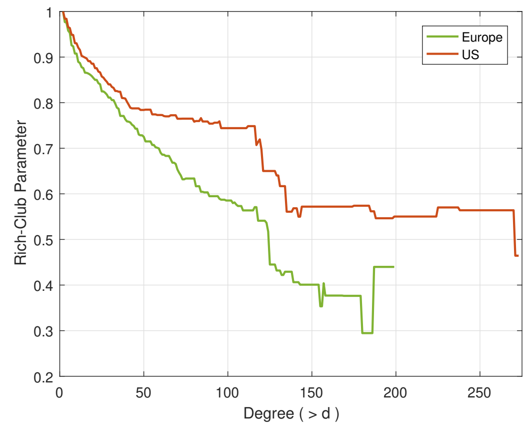


Fig. 2: Comparison of the rich-club parameters for the US and European flight networks

Next, we classify the nodes into communities by maximizing the modularity, and evaluate the gateway coefficient for each node. The community detection via modularity is a well-known algorithm. The presentation of the community detection is not a novel aspect of this study. But, there has been no study that compares the US and EU communities to figure out their differences in terms of the heterogeneity characteristics. We use an aggregated network in which the US and European networks and intercontinental flights are included. We perform a joint community detection on this aggregated network to distinguish the heterogeneity characteristics in the continents to provide an operational insight.

Fig. 3 presents the communities that are detected via modularity maximization. For a resolution threshold of 1 (the typically used hyperparameter value), we identify 16 communities in the aggregated network. Two of these communities cover the US airports exclusively, and the rest partition the European airports. The results show that the US has a more homogeneous traffic network structure than Europe. Even in Europe, there are only 3 major communities with a large number of airports; the others contain just 2 or 3 airports each. The heterogeneous structure of the European airspace network, although indicative of more fragmented connectivity among the airports, can be advantageous in limiting delay propagation across communities. Furthermore, the tight geographical proximity of most members of a community helps localize the delay and cancellation impacts of weather disruptions, and avoid cascading effects. The role of the gateway airports for such communities are paramount. If resources are directed towards gateway airports to maintain acceptable levels of

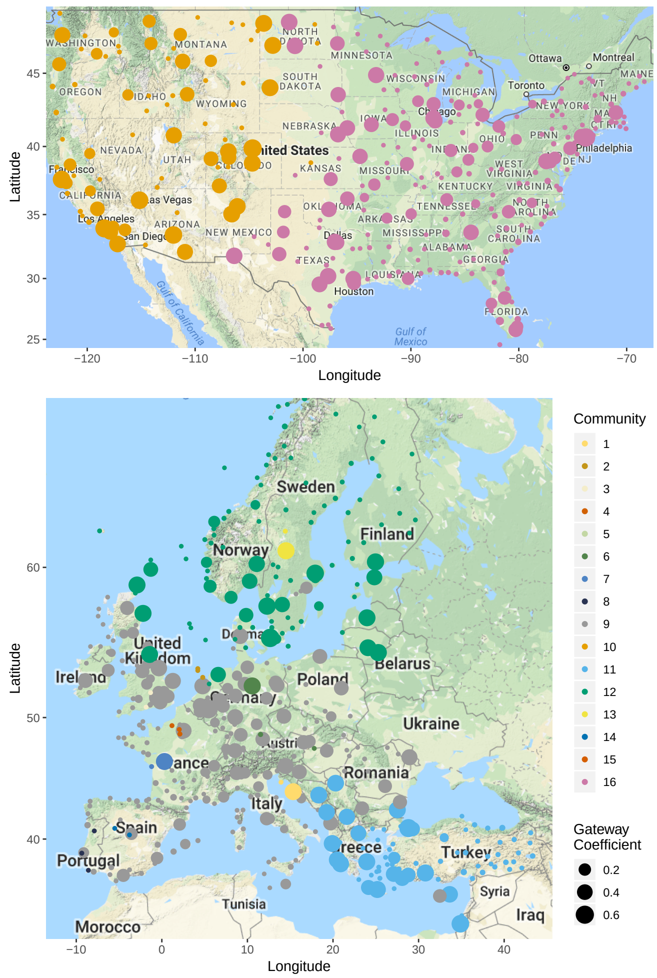


Fig. 3: Community structure and gateway coefficients in the combined US-Europe flight network

performance, then the chance of a disruption spreading from one community to another is minimized.

The airports with the highest gateway coefficients are presented in Table III. The right side of the table contains the gateway coefficients that are calculated according to the 16 communities in Fig. 3, whereas the left side includes the gateway coefficients that are computed assuming that there is only one community in each continent. The latter setup is used to identify key airports in either continents that not only are well connected within the continent, but also provide significant transcontinental connectivity. Naturally, the gateway coefficients are heavily affected by the community structure that is chosen. In the case of the join community detection (16 communities), Copenhagen (EKCH) and Istanbul (LTBA) have the highest gateway coefficients in Europe. While Copenhagen is the main gateway that connects the northern European community (marked in green) to the rest of Europe, Istanbul is the major gateway that connects the south-east European community (marked in blue) to the rest of Europe. In the US, Denver (KDEN) and Los Angeles (KLAX) have high gateway coefficients, and they are among the top gateways that connect the west coast community (yellow) to the east coast community (magenta). However, none of EKCH, LTBA, KDEN, and KLAX are among the top 10 ranked gateways between US and Europe (Table III, left). The top three airports connecting the US and Europe communities are New York's JFK (KJFK), London Heathrow (EGLL), and Shannon (EINN) airports. The node strength and gateway coefficient together provide a comprehensive method to identify airports that connect the two continents. Thus, we ignore the lower strength airports in the list (which are mainly cargo hubs), and conclude that KJFK and EGLL provide the best gateways into the US and European continents.

TABLE III: Top 10 airports based on their gateway coefficients (flight networks)

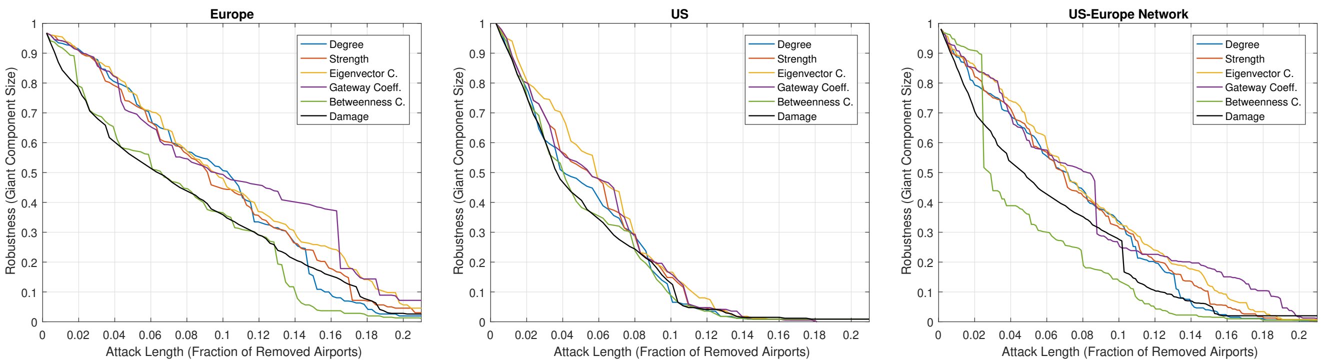
Coeff. between two Continents			Coeff. between 16 Communities		
Airport	Gateway Coeff.	Strength	Airport	Gateway Coeff.	Strength
KJFK	0.36	778.69	EBLG	0.64	14.92
EGLL	0.35	955.30	KJFK	0.63	778.69
EINN	0.35	18.22	EKCH	0.62	596.61
EBLG	0.26	14.92	KDEN	0.61	1287.84
KORD	0.21	2026.58	KLAX	0.60	1310.28
KEWR	0.21	821.74	LTBA	0.60	739.47
LFPG	0.20	929.69	EGLL	0.56	955.30
KATL	0.20	2110.08	KSFO	0.56	922.09
EDDF	0.18	958.07	KLAS	0.56	805.58
EHAM	0.18	1097.11	LLBG	0.55	254.89

B. Implications of higher flight connectivity on robustness

Robustness analysis via targeted node removal has been studied from flight's and passenger's perspectives in the literature [19], [20], [29]. But, there has been no study addressing the differences in the US and EU networks. Our intention is not to do the similar robustness analysis in the literature or identify only the most effective metric, but to compare the both networks and identify the differences in their flight networks' characteristics using the robustness analysis via targeted node removal strategy. We measure the robustness of a network with respect to a targeted node attack strategy. In this process, a node is removed from the network according to a specific network metric, and the impact of this removal on the size of the giant component is presented as a measure of robustness of the network. After each node removal, the network metric is recomputed, and the process repeated until the entire network is dismantled. We evaluate six different attacking strategies to determine which one is the most effective, and to evaluate if there are any differences between the US and Europe networks. The six metrics used are degree centrality, strength, eigenvector centrality, gateway coefficient, betweenness centrality, and damage (evaluated using a locally greedy heuristic).

The robustness of the EU, US and combined US-EU networks with respect to different node removal attack strategies are shown in Fig. 4. The three plots show the relative sizes of the giant components as a function of the fraction of nodes removed. Smaller the size of the giant component, the lower the robustness of the network to a given attack strategy.

We make four observations from the results in Fig. 4. First, for both networks where Europe was considered, the



(a) European network. (b) US network. (c) Combined US-Europe network.

Fig. 4: Size of the giant component of flight networks under targeted node removal attacks

best attack strategy among the predefined metrics, i.e., the one that leads to the smallest size of the giant component, varies depending on the attack length (the number of nodes removed). For instance, in the Europe-only network, betweenness centrality and damage lead to the greatest reduction in the size of the giant component until about 75% of the network is dismantled; after that, the betweenness centrality alone is the most effective. In the joint Europe-US network, damage is initially the most effective metric between the predefined metrics, and then for more intense attacks, the betweenness centrality strategy dominates. This is not the case in the US, where several of the metrics lead to similar decreases in the size of the giant component, with the betweenness centrality, damage, and degree being marginally better than others. A factor contributing to this similarity is the homogeneous nature of the US network. This structure is associated with the high connectivity between the airports, and thus multiple centrality metrics (degree, betweenness, damage) yield the same ordering of airports for targeted attacks.

A further consequence of the homogeneous structure of the US network (and the consistent ordering of airports for different attack strategies) is that it is easier to identify critical airports in the US, as compared to Europe. The first 10 airports that are removed from the network according to damage-, betweenness- and degree-based attacks are presented in Table IV. In the US, 8 of the top 10 airports are the same, independent of the attack strategy used. However, this is not the case for the European network, where only 4 of the 10 airports are the same for attacks based on betweenness centrality and damage. In other words, the risks in Europe are more distributed (with different airports being critical depending on the choice of the metric), making the European network robust to a larger number of attack strategies.

Our third observation is that the US network is also less robust than the European network with respect to the attack length. We consider a network as being *dismantled* when the size of the giant component is less than 1% of the total nodes. In Europe, 21% of the nodes have to removed for the network to be dismantled, while only 14% of the airports need to be removed to dismantle the US network. Even for a fixed fraction of node removals, say 0.1 (i.e., 10%), the size of the giant component is significantly higher in Europe (35-55% depending on the attack strategy) compared to the US (5-20%). The dismantling the whole network or a specific percent of a network leads to similar conclusions when comparing the focused networks. The observation implies that, in US, the minority of the network has great impact on the overall network. We will see that this observation also supports the results obtained in the control effectiveness analysis.

Finally, we note that the US-Europe combined network, because of its natural partition into two primary sub-networks, is highly susceptible to attacks based on the betweenness centrality metric. There is a sharp decline in the size of the giant component of the combined network when about 2.5% of the nodes are removed based on the betweenness centrality, due to the removal of gateway nodes that serve to connect the two continents. Beyond this threshold, all of the airports in the giant component belong to the one of the continents. Interestingly, the number of airports needed to obtain this sharp decline in robustness (i.e., size of the giant component) is only 19. The 19 airports that divide the network into two parts are KORD, KJFK, EHAM, EDDF, KBOS, LFPG, ESSA, LSZH, EIDW, EGLL, EGKK, KSFO, EGCC, EDDK, KEWR, KIAD, KATL, LEMD, KCVG.

TABLE IV: First 10 airports that would be removed under different attack strategies on flight networks

European network			US network		
Damage	Between.	Degree	Damage	Between.	Degree
LGAV	ESSA	EHAM	KDFW	KDEN	KORD
EFHK	EGKK	EDDM	KORD	KORD	KDFW
ESSA	ENGM	LTBA	KDEN	KDFW	KATL
LFPO	EGPD	EDDF	KMSP	KMSP	KDEN
EDDM	EHAM	EGKK	KSLC	KSEA	KCLT
ESSB	EKCH	LFPG	KSEA	KSLC	KDTW
EKCH	ENTO	EGSS	KANC	KDTW	KIAH
ENGM	EGLL	LEBL	KDTW	KCLT	KMSP
ENTC	EGSS	LEMD	KATL	KATL	KPHL
ENBR	EFHK	EKCH	KCLT	KIAH	KSEA

C. Discussion

The analysis of the traffic networks highlights some key differences between the US and European air transportation systems. The US has fewer airports, and fewer direct links between airports (i.e., fewer nodes and edges); however, it has significantly more operations between these airports compared to Europe (higher average strength). Complementing this denser connectivity, the US also has stronger connections between the high-degree airports (rich-club parameter), a greater tendency to form locally-aggregated clusters (clustering coefficient), and is more homogeneous (smaller number of distinct communities). The greater connectivity in the US, however, also affects network robustness: local disruptions at a few highly-central airports have the ability to significantly impact system performance.

III. Analysis of Delay Connectivity

Next, we investigate the flight delay networks in Europe and US. The delay networks are modeled as weighted directed graphs, in which each node represents an airport, and the weight on each edge denotes the magnitude of delay for the corresponding Origin-Destination (OD) pair. The OD pairs are originated from direct flights. Since delays vary at timescales of the order of hours, having a single representative network for the US (or Europe) would hide the subtleties of the delay dynamics. We therefore consider hourly delay networks, where each edge weight corresponds to the median delay of all flights that took off on that route in that hour. We only consider those OD pairs with at least seven flights a day, along with the associated airports.

We use a network to describe the state of the system for each hour; however, this still results in a very large number of networks over the course of a year ($24 \times 365 = 8,760$). We therefore identify representative delay networks using clustering. A feature vector consisting of all the edge weights is used to identify clusters, or groups, of hours with a similar network delay structure (see [15] for details on the clustering) using the k-means clustering algorithm. The centroid of each cluster is then considered as the representative delay state for that group of hours. The number of clusters in both the US and Europe is chosen to be eight, since a further increase does not result in qualitatively different network structures.

The delay states in the EU network and the US network are illustrated in Figs. 5 and 6, respectively. Although the networks are directed, for visualization purposes, they are symmetrized. Tables V and VI present the occurrence frequencies and high-level descriptions of the delay centroids. Henceforth, we use these eight representative networks to understand delays in the system, as they characterize the main trends that are observed, and provide a tractable set of graphs for analysis. As expected, the most common state of the system (i.e., the centroid with the highest frequency of occurrence) is characterized by low delays. We note that the scales for the color bars differ considerably between the US and European networks. The low delay state in the EU network corresponds to an average OD delay of 5 min, whereas the low delay in the US network corresponds to an average OD delay of 10 min. In general, the low delay states are more frequent in Europe than in the US. 87% of hours in Europe correspond to some low delay state (states 5, 6, and 1), whereas this fraction is only about 76% in the US (corresponding to states 2 and 3). The high delay states in both Europe and the US see major airports being disrupted – London, Amsterdam and Paris in Europe, and Chicago, Atlanta, Dallas, and the New York/Northeast airports in the US. Further analyses of the networks corresponding to these delay states are presented in the subsequent sections.

TABLE V: Description of delay states for the European network

Delay State	Frequency of occurrence (%)	Qualitative description
1	20.2	Medium network delay
2	1.6	EHAM very high delay
3	3.4	EGLL, LFPG, EDDF medium delay
4	2.7	LFPG, EHAM high delay
5	34.4	Very low network delay
6	32.0	Low network delay
7	4.1	EDDF, LFPG high delay
8	1.6	EGLL very high delay

TABLE VI: Delay states for the US network

Delay State	Frequency of occurrence (%)	Qualitative Description
1	1.9	KORD very high, KATL medium delay
2	31.8	Low network delay
3	44.2	Very low network delay
4	2.4	KLGA, KORD, KATL high delay
5	2.4	KDFW high, KORD, KIAH, KATL medium
6	9.3	KORD, KATL, KLGA medium delay
7	1.2	KATL high, KORD medium delay
8	6.8	KLAX, KORD medium delay

A. *Delay connectivity: Comparison of US and Europe*

Firstly, we present a novel approach to assess the characteristics of delay networks. The first set of analyses we conduct on delay networks helps us identify and understand which groups of high delay edges are connected. For this purpose, we use the size of the giant component as a metric of connectivity, and compute it for varying thresholds of delay. In this setup, the size of the giant component corresponds to the volume of the delay spreading in the network, and the delay thresholds refer to the possible amounts of delay compensations. Specifically, we compute the size of the giant component for the delay network while only considering edges with weights higher than the threshold. The threshold represents the compensation of delays through some external action; this metric therefore reflects the ability of the system to absorb delays. The threshold has an operational meaning. The delays originated from a problem in an airport can spread along the network. But, the impact of a departure delay on the destination airport or consecutive flights can also be mitigated by intervening the in-flight operation. For example, the cost index can be increased to reduce flight time or direct routes can be requested to reduce the travel distance, so the travel duration. The delay thresholds refer to how much delay can be compensated via these types of interventions. And, by changing the delay threshold, the analysis shows how the volume of delay spreading or cascading effect change as a function of delay threshold. In this way, the impact of varying delay thresholds on the delay spreading in the network is assessed, and the required delay thresholds to prevent the delay

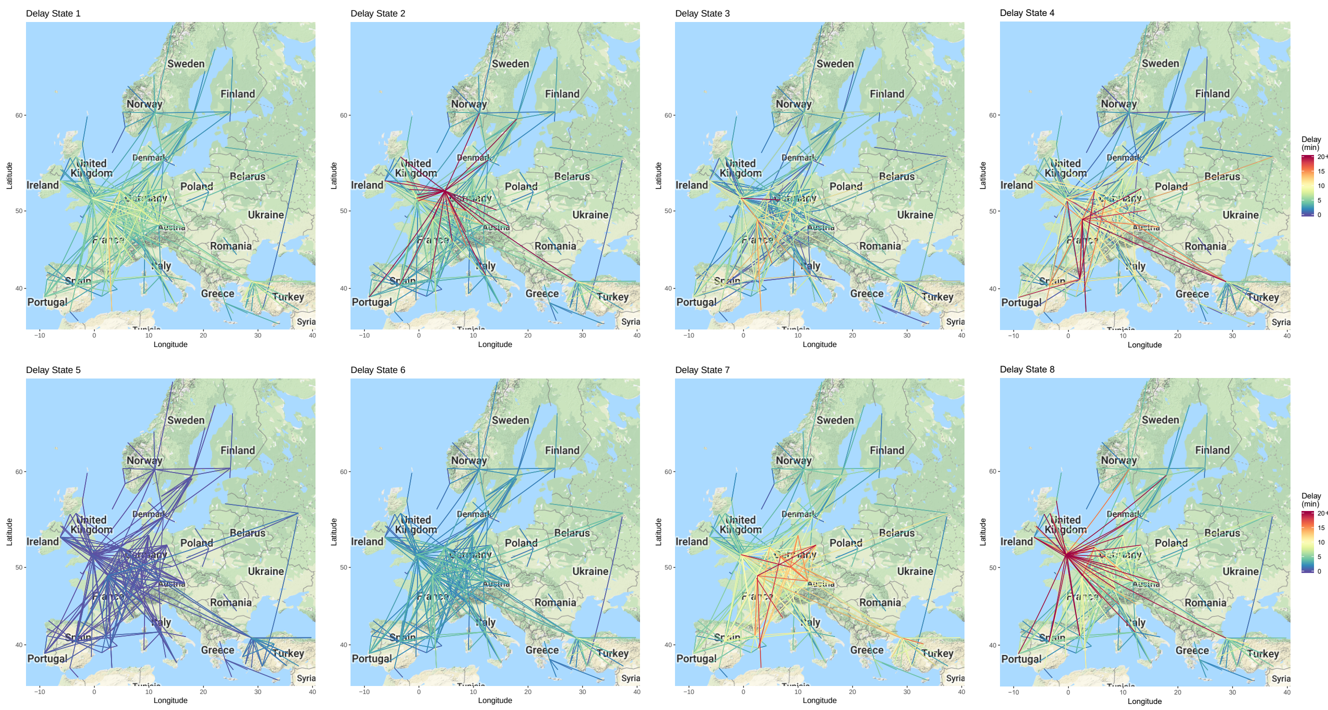


Fig. 5: Visualization of the delay states for the European Network

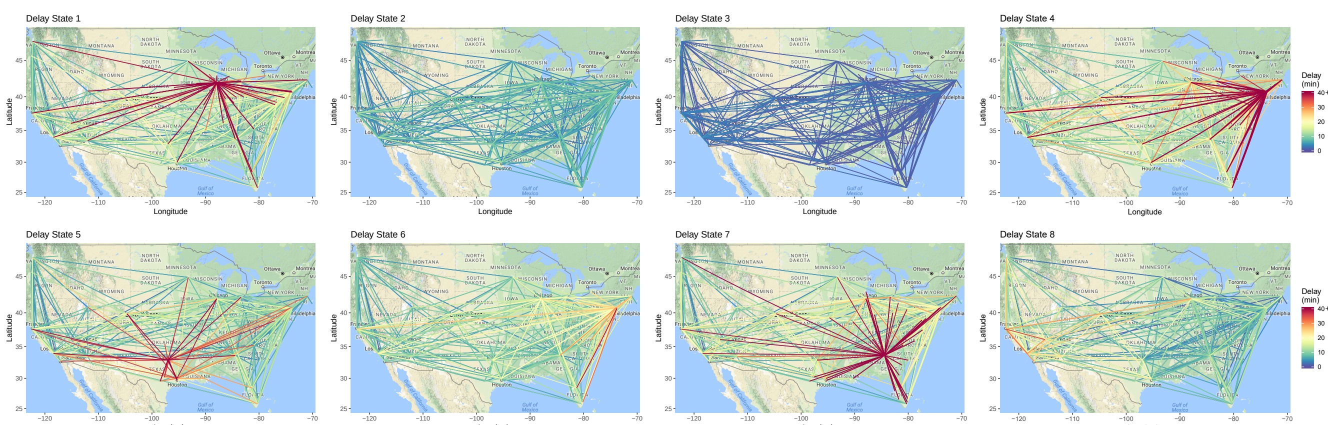


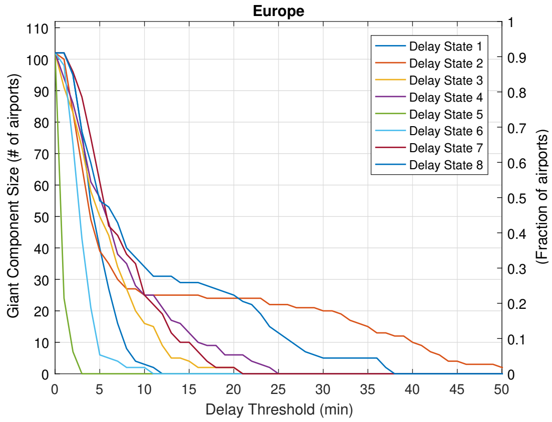
Fig. 6: Visualization of the delay states for the US Network

spreading or to restrain delay spreading into a smaller network are identified. Then, the delay spreading and compensation characteristics of the EU and US networks are compared.

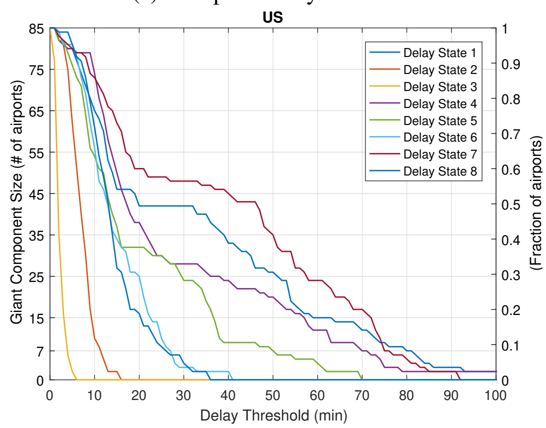
The size of the giant component with respect to varying delay thresholds for the network corresponding to each characteristic delay state is shown in Fig. 7. We observe that the very low delay states in both Europe and the US can be dismantled at even small delay thresholds. However, the specific thresholds – 3 min for Europe and 6 min for the US – reflect the higher baseline delay levels seen in the US.

The next set of delay states, ones with moderate delays (states 6 and 8 in the US, and states 1,3,4, and 7 in Europe), show a similar behaviour. The decrease in the size of the giant component in these delay states is more gradual, and tends to have a longer tail with an elbow-like feature. In other words, dismantling the last few nodes requires a significant increase in the delay threshold. In practice, this means that there is a set of highly-delayed airports that can be isolated from the majority of the network if a reasonable amount of delay (about 25 min for the US delay states, and 15 min for the European delay states) is compensated.

The delay states representing the most severe network impacts are different in Europe and the US, primarily due to the strength of high delay edges in the US. As a consequence, although the number of connected airports is higher in Europe to begin with, most of the airports can be isolated from the giant component with a modest amount of delay compensation. In the most severe delay states, the size of the giant component in Europe decreases sharply with an increase in the delay threshold, whereas it decreases slowly in US. For instance, let us consider the following high-delay states: state 2 in Europe, and state 7 in the US. Initially, more airports are a



(a) European delay network



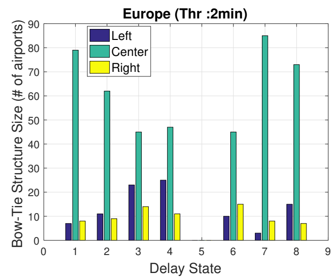
(b) US delay network

Fig. 7: Analysis of giant components in the characteristic delay states

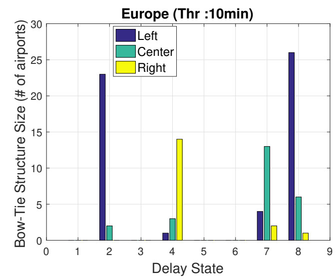
part of the giant component in Europe compared to the US. However, with just 8 min of delay compensation (i.e., a delay threshold of 8 min), the size of the giant component in Europe reduces to 27 airports; by contrast, the US has 77 airports still connected for a 8 min delay threshold. The decrease in giant component size comes at much higher thresholds for the US: with a delay compensation of 19 min, the size reduces to 51 airports. Not surprisingly, a significantly higher delay threshold is required to completely isolate all the airports in the US. Another interesting characteristic is the occurrence of relatively flat regions in the curve, indicating that there is a group of airports which experience similar delays. While this trend is sometimes repeated twice in the US at different ranges of the delay threshold, such repetition is not common in Europe. The repeated pattern indicates that there may be different sets of delay sources.

The results in Figure 7 pertain to the eight characteristic delay networks (each, for the US and Europe) that are observed. Similar trends hold when we consider all the actual realized delay networks (a total of 8,760 networks). While we do not show those plots for brevity, the same information, namely, the distribution of giant component sizes for each delay threshold, can be calculated. We find that half the network is dismantled for a delay threshold of 30 min in Europe, while the corresponding threshold is 75 min for the US. Furthermore, the delay threshold needed to dismantle 95% of the network is 60 min in Europe and 170 min in the US. We also observe that the size of the giant component in Europe decreases sharply with respect to delay threshold, but is more gradual in the US. One can therefore conclude that not only does the US have more high delay routes than Europe, but that these routes are more tightly connected to each other and form a larger giant connected component.

The second analysis focuses on the direction of the delay propagation. The aim here is to not just identify groups of high-delay edges and their connectivity, but rather understand the directions in which these high delay edges are pointed. The bow-tie structure metric will be used as a proxy to identify airports that serve as sources or sinks of delays in the system.

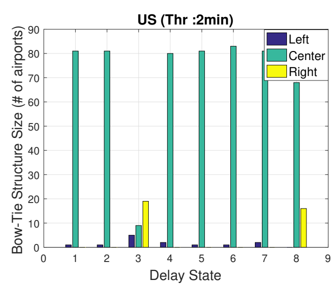


(a) 2-min threshold

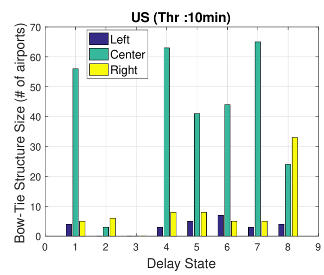


(b) 10-min threshold

Fig. 8: Bow-tie structures in Europe.



(a) 2-min threshold



(b) 10-min threshold

Fig. 9: Bow-tie structures in US

We wish to divide the airports in the giant component into three groups. The first group of airports has only outbound delay edges, the second group only has inbound delay edges incident on it, and the third group has both outbound and inbound edges. These are referred as the left, right, and center regions in the bow-tie structure, respectively. We restrict our analysis to the eight characteristic delay networks for Europe and the US. Fig. 8 and Fig. 9 present the number of airports in each of the three identified regions for the Europe and US characteristic delay networks, at threshold levels of 2 min and 10 min. A network that has several airports in all three groups is said to have a classic bow-tie structure. Operationally, this means that there are a reasonable number of airports that act as buffers or sources of delay in the system. Note that there should be some delays in the system to discuss the existence of the bow-tie structure. In case the delays are very low (e.g., the delay state 5 in Europe), there will be no discussion about the bow-tie structure.

An important observation from the analysis of the directional properties of these delay edges is that the US network does not form a strong bow-tie structure, whereas the Europe network does. This means that in Europe, there is a tendency for a reasonable fraction of airports to either be *sources* or *sinks* of delays. In the US however, there is no such distinction about the roles played by airports, and all of them tend to have outbound and inbound delays. This conclusion is driven by the results in Fig. 8 for Europe and Fig. 9 for the US. We present the bow-tie structures at two different threshold levels, with the higher level specifically chosen to bring out the structural trends more prominently. In the US, except delay state 8, which has some airports in the right region, almost all the airports are in the central region of the bow-tie structure. Even though a lot of the networks only tend to have airports belonging to the central region, the airports that actually belong to the left, or right may also be interesting to analyze. For instance, consider state 5 in the US. This is the state where Atlanta (KATL) and Dallas (KDFW) are the worst affected airports. In the corresponding bow-tie structure for 30-min threshold, Atlanta (KATL) gets classified into the center region, indicating that its the source of delays in this network, and most of the airports have inbound as well as outbound delays as a result. Interestingly, Dallas (KDFW), Chicago (KORD), and Houston (KIAH), which are major hub airports in the US get classified into the right region, indicating that they are acting as sinks for delays in this particular case. However, in Europe, there are several delay state networks where the distribution of airports is not focused on the center region. For instance, there are significantly more airports that are sources of delays in state 2 and 8, whereas many airports act as sinks in state 4. On the other hand, as expected from the US example, states where the delays are more concentrated, tend to have the affected airport as the sources.

B. Impact of higher delay connectivity on robustness

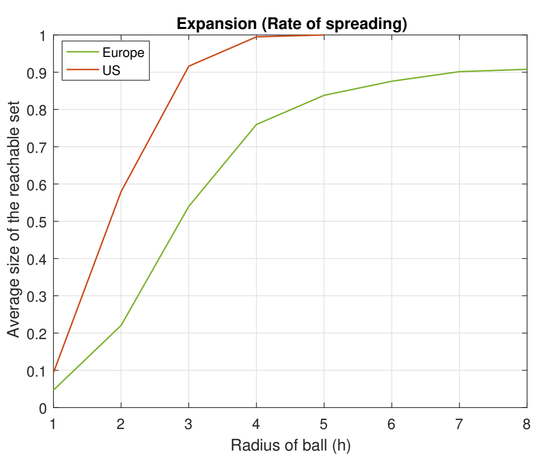


Fig. 10: Analysis of expansion (rate of spreading) metric

While the giant component informs us that the network is more connected, and has higher delay magnitude in the US as compared to Europe, it does not quantify how close or far away these airports are from each other. In other words, the same n airports can be a part of the giant component because they are all linearly connected, with the degree of the nodes being 1 or 2, or they can all be connected to one hub airport. This property is important in understanding the temporal dynamics in the propagation of delays- if there are fewer hops between two connected airports, delays can propagate much faster than if there are multiple edges that need to be traversed to travel from one airport to another. Essentially, when there is a disrupted airport that has flight delays, firstly the delays are propagated to its neighbors, and they spread it to their neighbors. The expansion metric captures this process, and is defined as the the average size of reachable set in the network for a fixed number of hops from a node.

The expansion functions for Europe and US are presented in Fig. 10. The expansion metric curves are relatively close for all the 8 delay state networks, and we present a weighted average line for US and Europe. As evident from the figure, all airports in US can be reached within a specific radius (of 4), whereas 10% of the airports in Europe cannot be reached on an average. To cover 90% of the airports in Europe, the radius (or number of hops) should be 7, whereas the corresponding metric for the US is just 4. Therefore, the US network has a higher rate of spreading than the European network.

C. Discussions

It is known through several studies and statistics that delays in the US are higher than in Europe. However, our finding is that delay spreading characteristic in the US is also different than in Europe. The giant component analysis shows that the US has a stronger cascading effect, and even with high delay thresholds some network delay states contain majority of the airports as a connected giant component. But, the cascading effect is not stronger in Europe in most delay states, and the delay spreading can be prevented using relatively low delay thresholds. It is observed that while Europe shows a bow-tie structure, this is not the case in US.

The strong connectivity of the US network in terms of both the traffic and delays in relation to Europe could be a result of the high traffic volumes, lower buffers in flight schedules, lack of slot controls, and other factors. These factors not only end up resulting in stronger directional trends in delay propagation (as evident by the bow-tie structure) but also a faster, and more expansive spreading of delays from disrupted airports (as suggested by the expansion metric)

IV. Modeling and Control of Airport Delay Dynamics

In this section, we present novel metric of resilience of networked systems grounded in network models and their control theoretic properties. The primary advantage of such measures is that they are based on the dynamics of the process on the network, rather than just the structure and connectivity of the system. In this section, we show how a data-driven model for a networked system can be interpreted as a measure of resilience. Further, we use the system model to quantify the controllability of delays spreading to present a practical measure of the ability of the system to recover from disruptions.

While this network-model based approach to quantifying resilience is independent of the model per-se, we use a specific class of data-driven network model called the Markov Jump Linear System (MJLS) to perform our analysis. It is however important that the model can be learnt from data, in order to capture the actual dynamics of the system, and not just rely on idealized first principles to quantify resilience. We also wish to emphasize that while there are several approaches to modeling, simulating, and predicting flight delays, analytically tractable models would be preferred for this analysis as the scale of the network may be prohibitively large for reliable simulations. In the remainder of the section, we describe the MJLS model along with the optimal controller, and a comparison between the European and US airport delay models.

A. Markov Jump Linear System (MJLS) Model

The MJLS model (presented in [21], validated in [30]) is used to represent the hourly evolution of total inbound and outbound delays at airports. We develop the model, and provide intuition for the terms in it. In terms of notations, let there be N airports, and $x_i^{out}(t)$ and $x_i^{in}(t)$ represent the total outbound and inbound delays along all incident OD pairs for airport i in the one-hour interval starting at t. A simple linear model for the delays at the next time step is given by

$$x_i^{out}(t+1) = \alpha_i x_i^{out}(t) + \sum_j \beta_{ji}\bar{w}_{ji}x_j^{in}(t) \quad (1)$$

$$x_i^{in}(t+1) = \alpha_i x_i^{in}(t) + \sum_j \beta_{ji}\bar{w}_{ij}x_j^{out}(t) \quad (2)$$

where $\bar{w}_{ji}$ are the elements of a normalized weight matrix, and α, β are model parameters. The weight matrix is associated with the delay network corresponding to the hour t, but is row-normalized. It represents the relative strength of interaction between any two airports in terms of the delays for that hour. The equations qualitatively describe the delay evolution using two factors: the persistence of delays due to queuing effects, and network spreading effects. The queuing effects results in persistence of delays at an airport and is represented with the first term in the equation. The second term is the weighted average effect of the connected airports on the delays at the particular airport i. The resultant equations can be written in a simplified matrix form as $x(t+1) = A(t)x(t)$, where $x(t)$ consists of total outbound and inbound delays of all airports and $A(t) = [\alpha] + [\beta]\bar{W}(t)$. In the model, $W(t)$ is chosen to be one of the eight characteristic network delay states (computed in the previous section). Each of the different states indicates a different pattern of interaction between the airports that result in different evolution of the delay state. Thus, the system is described as

$$x(t+1) = A_{m(t)}x(t) \quad (3)$$

where $A_{m(t)}$ is the system matrix for the delay mode $m(t)$ at time t. $m(t)$ belongs to the set $\{1, \ldots, M\}$. Since the delay modes are changing, depending on the weather phenomenon, random disruption, or just temporal traffic flow effects, the delay modes can transition from time to time. These transitions are modeled as a time-dependent Markovian process, with the transition probability given by:

$$\mathbb{P}[m(t+1) = j \mid m(t) = i] = \pi_{i,j}(t) \quad (4)$$

The delay modes, transition matrices, and coefficient α and β are learnt from data. The MJLS model is thus characterized by the finite set of system matrices $A_{m(t)}$, a time-dependant transition probability given by $\pi_{ij}(t)$, and an initial condition. The equations are non-deterministic, meaning that they describe the various possible evolutions of the airport delays in the system due to randomly varying network topologies. It has been shown in prior works that these models are representative of the air traffic delay dynamics and have a reasonable prediction performance too [21], [30]. Thus, their use for comparing the resilience of the networks using operation delay information is justified.

The last methodological detail we wish to highlight is the control of such systems. The study [31] has proposed an integrated control structure for the MJLS model to manage the system via node and topology control. But, our intention is not to develop an advanced control structure for the MJLS model. We aim to present how a control-centric method can be used to compare the delay spreading characteristics of different networks and obtain the operational insights by comparing the US and European networks. In this study, we use an LQR algorithm to implement the optimal control strategy, but the comparison strategy is not limited to the presented algorithm. Operationally, controlling the state of the MJLS model, $x(t)$ through an input vector $u(t)$ means that we can change the inbound or outbound delays at the airports by $u(t)$. In such a case, the system would evolve as

$$x(t+1) = A_{m(t)}x(t) + Bu(t) \quad (5)$$

where $u(t)$ is the control input, B is a 0-1 matrix to determine which airports can be controlled, the initial conditions $x(0)$ and $m(0)$ are given, and the probability of transitions are governed by Equation (4). Operationally, we could control delays at particular airports at such tactical time scales by a combination of preferential routing, airspace flow control programs that can re-distribute delays, or ground delay programs (in the US). While the exact mechanics of how such a control action can be achieved is not a focus of this paper, we wish to study the potential effect of one airport, or a group of airports on the entire system.

A metric of resillience for an airport would be its ability to reduce delays in the entire system. Formally, we define it in this setup. Suppose we wish to minimize the delay-cost, i.e the sum of delays during the course of a day for a given initial delay state. In other words, we want to minimize the following quadratic cost function:

$$\mathbb{E}[\sum_{t=1}^{T}(x(t)^T x(t) + \eta u(t)^T u(t))] \quad (6)$$

where $\eta \geq 1$ is a positive scalar. Naturally, to minimize delays, we would have to take a control action $u(t)$. This control action comes with a penalty, that is scaled by η. Thus, the ability of control actions at specific airports to reduce this objective function, for a given η is defined as the controllability of the airport. Lower the total cost, greater is the airports ability to influence the delay dynamics and help improve system performance. A specific way to allow for control actions only at specific airports would be to set the corresponding columns of B to zero.

The optimal control law that minimize this cost function is a state feedback law [32] and is given by:

$$u^*(t) = F_{m(t)}(t)x(t) \tag{7}$$

Analytical, closed form expressions for the gain matrix F, and the optimal cost are presented in [32], and is a standard LQR control problem.

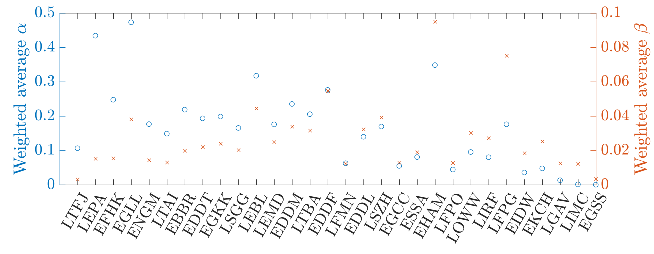
B. Comparing the model parameters in European and US Networks

First, we identify these model parameters (i.e the α, β, π, A) for the Europe and US models. The discrete modes, i.e the weight influence matrix W are chosen based on the 8 characteristic delay states. Additionally, since each corresponding delay state may represent either an increasing or decreasing delay trend, they are further split into two, based on the trend for the total system delay. For instance, delay state 1 in the US, corresponding to Chicago ORD high delays is split into two modes, one corresponding to Chicago increasing delays, and other corresponding to Chicago decreasing delays. Thus, there are 16 discrete modes for each of the networks. This allows for the empirical evaluation of the transition matrices and the average weights W for each of the modes. The parameters α and β are evaluated through linear regression.

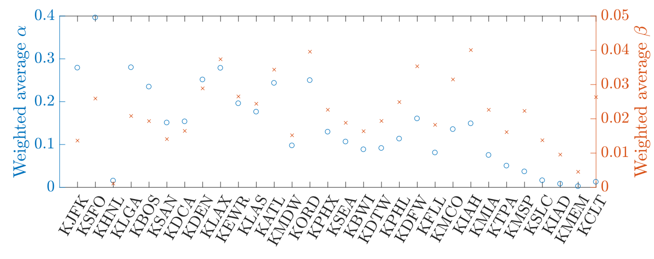
TABLE VII: Weighted Average Values of α and β Parameters for the EU and US Networks.

	Weighted Average	
	α	β
Europe	0.164	0.027
US	0.142	0.022

The model parameters α and β are interesting novel measures of delay resilience. Unlike conventional network theoretic measures that just depend on the structure and connectivity of the airports, these measures explicitly take into account the delay process that is evolving on the network. From data, we learn an α (and β) for each airport (or airport pair) and mode. Thus, when we refer to the alpha-value of an airport, it is the mode-frequency weighted average value. Formally, if f_m denotes the empirically observed frequency of occurrence of mode $m \in \{1, \ldots, , M\}$, then the average value of alpha, for an airport is given by $\sum_{m=1}^{M} f_m a_m$. Consequently, the average alpha for the whole network is the average over all the airports in the network. A similar analogy exists for the betas. It is worth reminding that the α parameter represents the persistence term in the delay dynamics. Higher the α, higher is the persistence of the delay level. This not only means that high delays at an airport tend to take longer to die down, but



(a) Airports in the European Network



(b) Airports in the US Network

Fig. 11: Weighted average value of the model parameters for the 30 busiest airports. Airports are sorted in the descending order of α/β.

also that an increase in delay also is a slow process. Overall, it simply means that the inertia, or resistance to change in delay levels is high. Airports where demand is very close to capacity, and with very little buffer in terms of operations tend to have highly disciplined schedules to avoid delays. This explains the high resistance to change, when the delays are low. On the other hand, when delays are inevitably encountered, the limited buffer also provides very little relief to aid in the recovery, thus making it slow. The beta term refers to the airports ability to influence, or get influenced by others. When β is high, the in-bound and out-bound delays at an airport depend strongly on the outbound and inbound delays at other airports respectively. Airports with high values of β are typically well connected in the system, both in terms of traffic, and also the nature of airports it is connected to.

The mode and airport averaged values of alpha and beta for the US and Europe network is presented in Table VII. On the whole, both networks have a similar value for the α and β parameter. Going one step further, the ratio α/β is used to identify which network tends to have more inter-connectivity effects dominate the delay dynamics. Interestingly, this parameter is also similar for both Europe and the US network.

A more nuanced view emerges when we look at α, β, and their ratio at an airport level (Fig. 11). In Europe, the three airports that have highest average α values are London (EGLL), Palma Spain (LEPA) and Amsterdam (EHAM), and the airports that have highest average β values are Amsterdam (EHAM), Paris (LFPG), and Frankfurt (EDDF). All these airports that have high α or β, except Palma, are large hub airports. However, the characteristics of these hub airports are not all the same. The airports in Fig. 11 are sorted in descending order of the ratio α/β. This means that London (EGLL) is more persistence to delays rather than network interactions when compared with the Paris (LFPG). This fidelity of characterization of airport hubs is only possible through our analysis using the dynamics of delays on networks. The other airports in Europe having high α/β ratios are Istanbul Sabiha (LTFJ), Palma (LEPA), and Helsinki (EFHK). In US, the top three airports that have highest average α values are San Francisco (KSFO), New York Laguardia (KLGA), and New York Kennedy (KJFK), and the airports that have highest average β values are Houston (KIAH), Chicago (KORD), and Los Angeles (KLAX). All these airports that have high α or β are also large hubs. The airports that have highest α/β values are New York Kennedy (KJFK), San Francisco (KSFO), and Honolulu (KHNL). These airports are more sensitive to the persistence of delays than the network interactions when compared with the airports that have relatively small α/β values such as Houston (KIAH) and Dallas/Fort Wroth (KDFW).

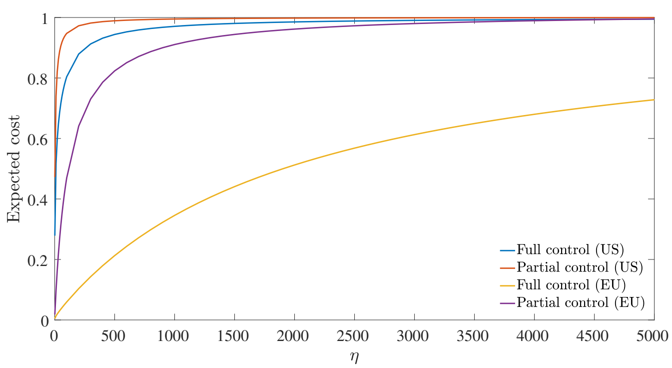
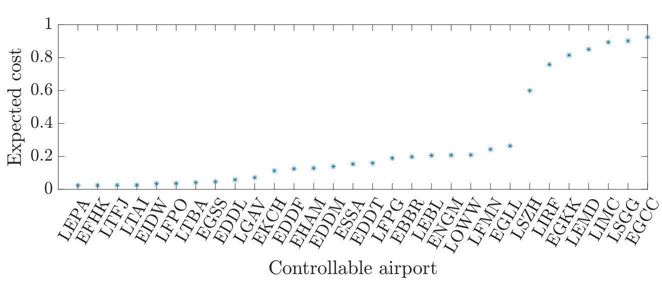
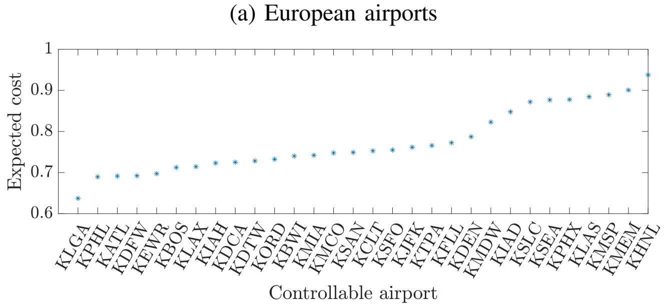


Fig. 12: Comparison of the full and partial control strategies



(a) European airports



(b) US airports

Fig. 13: Expected cost when the control action is restricted to one airport in the network

C. *Comparing the control effectiveness of US and European airports*

We also analyse the European and US networks in terms of the ability of airports to reduce the spread of delays. We utilize the aforementioned LQR-based controller to reduce the delays as described by the MJLS models for both networks. In the first analysis, we compare the performance of the full control and partial control strategies. The full control corresponds to the control of all airports in the network, while the partial control refers to the control of a subset of the airports in the network. This controllable set of airports is specified by setting the B matrix. In the first analysis, the controllable set contains the top five airports that have highest degree centralities in the network. The aim is to analyse the impact of the highest degree airports on the control performance. The expected costs in the full control and partial control strategies are presented for both networks in Fig. 12 as a function of the penalty parameter η. Recall that the total cost is the sum of the expected delay cost over the 24 hour period and the penalty for the control action. The cost is normalized such that when no action is taken, the total is 1, and any control action would lead to a total cost strictly less than 1. For a fixed η, lower the cost, more effective is the chosen subset of airports in controlling the spread of delays. As shown in Fig. 12, in US, the performance of the partial control setting is closer to the performance of the full control. However, this is not the case for Europe. The performance difference between the full and partial control settings in EU are higher than the difference in US. Therefore, the high-degree nodes have higher impact on the control of the network in US when compared with the European network.

In the second analysis, we focus on the individual airports to quantify their influence on the system. In this case, only one airport in the network is controllable. The parameter η is set to be 1. For both networks, the individual control performance of the busiest 30 airports are presented in Fig. 13. In Europe, the performances of the first 23 airports are closer to each other, while the rest of the airports have worse performances than these airports. For the first 23 airports, the consecutive airports have approximately same effectiveness, but there is a steady increase in the expected cost. So, there is a performance difference between the 1^{st} and 23^{rd} airports. The top three airports that have highest impact on the control performance are LEPA, EFHK, and LTFJ. It is interesting to see that these three airports have also highest α/β ratios. However, it is not always the case. For example, LTBA is one of the airports that have high impact on the control performance, whereas it has an average α/β ratio. The control performance is affected from the model's parameters, connectivity structure and delay levels, so it is more complicated than the model's parameters. However, when compared with the US network, the airports that have high α/β ratio in Europe have a tendency to have high impact on the control performance. In US, the top three airports that have highest impact on the control performance are KLGA, KPHL, and KATL. In these three airports, only KLGA is one of the airports that have highest α/β ratio. Besides, KATL and KDFW are in the third and fourth places in terms of control effectiveness, and these two airports are 2 of 3 airports that have highest degree centralities in US. When compared with Europe in terms of the control effectiveness, there is a tendency in US to prefer the high degree nodes to the nodes that have high α/β ratio. Furthermore, as is the case in Europe, the performances of the first 21 airports in US are closer to each other, while the rest of the airports have worse performances than these airports. However, the performance difference between the first and second groups of airports in US is not as large as the difference in Europe. The performance difference between the best and worst cases in US is smaller than the difference in Europe. Therefore, the US network has a more homogeneous control effectiveness than the European network.

D. Discussions

On the whole, the US and Europe networks have similar delay persistence and delay-spreading tendencies. However, we do identify specific airports that have high persistence (α) and delay-interconnections (β). The control effectiveness of an airport in reducing the delays in the entire system is a novel, and practical measure of the role of the airport in system performance. In fact, as seen with the example of the Europe network, it is clear that this metric is different from the traditional notion of degree centrality. Thus, a more operationally driven measure of resilience, that incorporates information about the dynamics of delays is more valuable for researchers in identifying critical airports. Lastly, we re emphasize that the stronger inter-connectivity in terms of traffic in the US is a double edged sword- it not only increased the delays and contributes to greater spreading, but also provides greater flexibility and controllable in recovering from disruptions.

V. Conclusions

This work compared the resilience of aviation networks in the US and Europe in terms of flight connectivity and delay propagation. The main objective was to compare the resilience of the two systems, and identify critical airports for these two major regions of aviation traffic in the world.

Through our analysis, we identify that the US airport network is more dense in terms of the traffic, has stronger connections between major hubs, has higher delays on an average, and can lead to greater spreading of flight delays. On the other hand, such high inter-connectivity also enables better disruption management and recover, as highlighted from the control theoretic analysis and the MJLS network models. A key contribution of our work is the identification of airports that are critical for the resilience and robustness of the system both in Europe and the US.

A comparison of different aviation infrastructures is useful for understanding the geographical and historical context for many of the observed effects on system resilience. Using models and control theoretic measures to quantify the role of airports is a new approach towards quantifying resilience not only in the aviation context, but also other networked systems ranging from transportation, communication, and critical infrastructures. The use of our simple, data-driven MJLS model to analyze other systems from a network resilience perspective is an area of ongoing work.

Appendix A

Consider a weighted directed graph with N nodes and the edge weights represented using the weight matrix $\mathcal{W} \in \Re^{N\times N}$. The underlying connectivity is represented using the adjacency matrix $\mathcal{A} \in \Re^{N\times N}$, where $a_{ij} = 1$ if $w_{ij} > 0$, and $a_{ij} = 0$ if $w_{ij} = 0$. And, the symmetrized weight matrix is defined as follows: $W^{sym} = (W + W^T)/2$.

A. Node properties

a) Degree: The degree of a node captures the number of connections to or from other nodes. The in-degree of a node i is defined as $d_i^{in} = \sum_j^N a_{ji}$, and the out-degree as $d_i^{out} = \sum_j^N a_{ij}$. The total degree is the sum of its in-degree and out-degree: $d_i^{tot} = d_i^{in} + d_i^{out}$.

b) Strength: The strength, in contrast to the degree, captures the intensity of the connection, and not just the number of such connections [9]. The in-strength, out-strength, and total strength of node i are defined as $s_i^{in} = \sum_j^N w_{ji}$, $s_i^{out} = \sum_j^N w_{ij}$, and $s_i^{tot} = s_i^{in} + s_i^{out}$.

c) Betweenness Centrality: Bridging nodes that connect disparate parts of the network often have a high betweenness centrality. It is defined using the shortest path concept for every node i as $B_i = \sum_{s\neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$, where σ_{st} is the number of shortest paths going from node s to node t, and $\sigma_{st}(i)$ is the number of shortest paths going from s to t passing through i.

d) Eigenvector Centrality: The influence of a node in the network can be measured by the eigenvector centrality as nodes with a high eigenvector centrality have strong connections with other nodes that have a high eigenvector centrality. It's value at each node is the eigenvector corresponding to the largest eigenvalue of W^{sym}.

e) Expansion (Rate of Spreading): The expansion metric for each node is the number of nodes that can be reached within h hops. This can be computed for the whole network [33] by averaging over all the nodes,and normalizing by N, to quantify the faction of the network that can be covered by the whole graph (the reachable set).

f) Clustering Coefficient: The clustering coefficient for a node i is the ratio of the number of triangles with node i as the vertex to the maximum possible triangles that could have existed for another node of the same degree. This metric captures the tendency of the graph to form small tight communities, and can be extended for weighted directed graphs too as shown in [34].

B. Classification of nodes

a) Core and Periphery Nodes: The core/periphery subdivision is a partition of the network into two non-overlapping groups of nodes based on whether a node is central in the network or not [35].

b) Modularity: The modularity is used to subdivide the network into groups of nodes in a way that maximizes the number of within-group edges, and minimizes the number of between-group edges [36].

c) Gateway Coefficient: The gateway coefficient identifies nodes which have both high inter-community, and intra-community connections [37]. It is thus used to identify gateways, or critical nodes in a community that serve as links to other communities.

C. Graph properties

a) Size of giant component: The giant component refers to the largest connected nodes in the network. The size of the giant component denotes the number of nodes in this set. This size could also be normalized for comparison purposes.

b) Bow-Tie Structure: The bow-tie structure (adapted from [38]) is obtained by dividing the giant component into three regions. The center region corresponds to the set of connected nodes such that for any pair of node i and j in the set there is an edge from i to j. The left region contains the nodes that can reach the center region, while the nodes in the left region can't be visited from the nodes in the center region. The right region consists of the nodes that can be reached from the center region, whereas the nodes in the right region can't reach the nodes in the center region. The identification of this structure highlights the directional influence of nodes in the network.

c) Rich-Club Parameter: This measures the tendency of prominent elements to form clubs with exclusive control over the majority of a system's resources [39]. It refers to the fraction of weights shared by the rich nodes compared with the total amount they could share if they were connected through the strongest links of the network. As a richness parameter, any network metric can be used. In case the degree is used as the richness variable, the rich-club parameter evaluates the tendency for high-degree nodes to be more densely connected among themselves than nodes of a lower degree [40].

d) Robustness to targeted node removal: : In this metric, the size of the giant component is computed for repeated removals and isolation of nodes from the network. This measure of resilience quantifies the ability of the node to retain connectivity and function even when some parts are compromised. Several approaches to removing the nodes can be used, ranging from centrality measures like eigencentrality, degree, strength etc., or even a simple greedy approach (e.g. damage metric [41]).

ICAO	Airport Name	ICAO	Airport Name
EBBR	Brussels	KATL	Atlanta
EDDF	Frankfurt	KBOS	Boston
EDDL	Dusseldorf	KBWI	Baltimore
EDDM	Munich	KCLT	Charlotte
EDDT	Berlin-Tegel	KDCA	Washington-National
EFHK	Helsinki-Vantaa	KDEN	Denver
EGCC	Manchester	KDFW	Dallas-Fort Worth
EGKK	London-Gatwick	KDTW	Detroit
EGLL	London-Heathrow	KEWR	Newark
EGSS	London-Stansted	KFLL	Fort Lauderdale
EHAM	Amsterdam-Schiphol	KHNL	Honolulu
EIDW	Dublin	KIAD	Washington-Dulles
EKCH	Copenhagen-Kastrup	KIAH	Houston-Intercontinental
ENGM	Oslo-Gardermoen	KJFK	New York-J. F. Kennedy
ESSA	Stockholm-Arlanda	KLAS	Las Vegas
LEBL	Barcelona-El Prat	KLAX	Los Angeles
LEMD	Madrid-Barajas	KLGA	New York-LaGuardia
LEPA	Palma De Mallorco	KMCO	Orlando
LFMN	Nice-Cote d'Azur	KMDW	Chicago-Midway
LFPG	Paris Charles De Gaulle	KMIA	Miami
LFPO	Paris-Orly	KMSP	Minneapolis
LGAV	Athens International	KORD	Chicago-O'Hare
LIMC	Milan-Malpensa	KPDX	Portland
LIRF	Rome-Fiumicino	KPHL	Philadelphia
LOWW	Vienna International	KPHX	Phoenix
LSGG	Genava	KSAN	San Diego
LSZH	Zurich	KSEA	Seattle
LTAI	Antalya	KSFO	San Francisco
LTBA	Istanbul-Ataturk	KSLC	Salt Lake City
LTFJ	Sabiha Gokcen	KTPA	Tampa

TABLE VIII: ICAO code and corresponding full airport name of the airports within in Europe (left) and the US (right)

REFERENCES

[1] S. Paleari, R. Redondi, and P. Malighetti, "A comparative study of airport connectivity in china, europe and us: which network provides the best service to passengers?" *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 2, pp. 198–210, 2010.

[2] *Comparison of Air Traffic Management Related Operational Performance US/Europe*, 2017. [Online]. Available: https://www.eurocontrol.int/sites/default/files/2019-05/us-europe-comparison-operational-performance-2017.pdf

[3] M. Zanin and F. Lillo, "Modelling the air transport with complex networks: A short review," *The European Physical Journal Special Topics*, vol. 215, no. 1, pp. 5–21, 2013.

[4] B. Baspinar, N. K. Ure, E. Koyuncu, and G. Inalhan, "Analysis of delay characteristics of european air traffic through a data-driven airport-centric queuing network model," *IFAC-PapersOnLine*, vol. 49, no. 3, pp. 359–364, 2016.

[5] P. Fleurquin, J. J. Ramasco, and V. M. Eguiluz, "Systemic delay propagation in the us airport network," *Scientific reports*, vol. 3, p. 1159, 2013.

[6] B. Baspinar, E. Koyuncu, and G. Inalhan, "Large scale data-driven delay distribution models of european air traffic flow network," *Transportation research procedia*, vol. 22, pp. 499–508, 2017.

[7] S. Dunn and S. M. Wilkinson, "Increasing the resilience of air traffic networks using a network graph theory approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 90, pp. 39–50, 2016.

[8] R. Guimera, S. Mossa, A. Turtschi, and L. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 22, pp. 7794–7799, 2005.

[9] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, "The architecture of complex weighted networks," *Proceedings of the national academy of sciences*, vol. 101, no. 11, pp. 3747–3752, 2004.

[10] D. DeLaurentis, E.-P. Han, and T. Kotegawa, "Network-theoretic approach for analyzing connectivity in air transportation networks," *Journal of aircraft*, vol. 45, no. 5, pp. 1669–1679, 2008.

[11] G. Bagler, "Analysis of the airport network of india as a complex weighted network," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 12, pp. 2972–2980, 2008.

[12] D.-D. Han, J.-H. Qian, and J.-G. Liu, "Network topology and correlation features affiliated with european airline companies," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 1, pp. 71–81, 2009.

[13] W. Li and X. Cai, "Statistical analysis of airport network of china," *Physical Review E*, vol. 69, no. 4, p. 046106, 2004.

[14] A. Cook, S. Belkoura, and M. Zanin, "Atm performance measurement in europe, the us and china," *Chinese Journal of Aeronautics*, vol. 30, no. 2, pp. 479–490, 2017.

[15] J. J. Rebollo and H. Balakrishnan, "Characterization and prediction of air traffic delays," *Transportation research part C: Emerging technologies*, vol. 44, pp. 231–241, 2014.

[16] K. Gopalakrishnan, M. Z. Li, and H. Balakrishnan, "Network-centric benchmarking of operational performance in aviation," *Transportation Research Part C: Emerging Technologies*, vol. 126, p. 103041, 2021.

[17] K. Gopalakrishnan, H. Balakrishnan, and R. Jordan, "Deconstructing delay dynamics: An air traffic network example," in *International Conference on Research in Air Transportation (ICRAT)*, 2016.

[18] N. Pyrgiotis, K. M. Malone, and A. Odoni, “Modelling delay propagation within an airport network,” *Transportation Research Part C: Emerging Technologies*, vol. 27, pp. 60–75, 2013.

[19] O. Lordan, J. M. Sallan, P. Simo, and D. Gonzalez-Prieto, “Robustness of the air transport network,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 68, pp. 155–163, 2014.

[20] S. Wandelt, X. Sun, and X. Cao, “Computationally efficient attack design for robustness analysis of air transportation networks,” *Transportmetrica A: Transport Science*, vol. 11, no. 10, pp. 939–966, 2015.

[21] K. Gopalakrishnan, H. Balakrishnan, and R. Jordan, “Stability of networked systems with switching topologies,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 2601–2608.

[22] B. Baspinar and E. Koyuncu, “A data-driven air transportation delay propagation model using epidemic process models,” *International Journal of Aerospace Engineering*, vol. 2016, 2016.

[23] D. Lu, S. Yang, J. Zhang, H. Wang, and D. Li, “Resilience of epidemics for sis model on networks,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 8, p. 083105, 2017.

[24] J. P. Sterbenz, D. Hutchison, E. K. Çetinkaya, A. Jabbar, J. P. Rohrer, M. Schöller, and P. Smith, “Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines,” *Computer Networks*, vol. 54, no. 8, pp. 1245–1265, 2010.

[25] H. Amini, R. Cont, and A. Minca, “Resilience to contagion in financial networks,” *Mathematical finance*, vol. 26, no. 2, pp. 329–365, 2016.

[26] J. Ash and D. Newth, “Optimizing complex networks for resilience against cascading failure,” *Physica A: Statistical Mechanics and its Applications*, vol. 380, pp. 673–683, 2007.

[27] *EUROCONTROL, OneSky Online, Demand Data Repository (DDR2)*, 2019. [Online]. Available: https://ext.eurocontrol.int/

[28] *Bureau of Transportation Statistics, Airline On-Time Statistics and Delay Causes*, 2019. [Online]. Available: http://www.transtats.bts.gov/

[29] X. Sun, V. Gollnick, and S. Wandelt, “Robustness analysis metrics for worldwide airport network: A comprehensive study,” *Chinese Journal of Aeronautics*, vol. 30, no. 2, pp. 500–512, 2017.

[30] K. Gopalakrishnan and H. Balakrishnan, “A comparative analysis of models for predicting delays in air traffic networks.” ATM Seminar, 2017.

[31] ——, “Control and optimization of air traffic networks,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, 2021.

[32] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov jump linear systems*. Springer Science & Business Media, 2006.

[33] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, “Network topology generators: Degree-based vs. structural,” in *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 4. ACM, 2002, pp. 147–159.

[34] G. Fagiolo, “Clustering in complex directed networks,” *Physical Review E*, vol. 76, no. 2, p. 026107, 2007.

[35] S. P. Borgatti and M. G. Everett, “Models of core/periphery structures,” *Social networks*, vol. 21, no. 4, pp. 375–395, 2000.

[36] M. E. Newman, “Fast algorithm for detecting community structure in networks,” *Physical review E*, vol. 69, no. 6, p. 066133, 2004.

[37] E. R. Vargas and L. M. Wahl, “The gateway coefficient: a novel metric for identifying critical connections in modular networks,” *The European Physical Journal B*, vol. 87, no. 7, p. 161, 2014.

[38] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, “Graph structure in the web,” *Computer networks*, vol. 33, no. 1-6, pp. 309–320, 2000.

[39] T. Opsahl, V. Colizza, P. Panzarasa, and J. J. Ramasco, “Prominence and control: the weighted rich-club effect,” *Physical review letters*, vol. 101, no. 16, p. 168702, 2008.

[40] M. P. Van Den Heuvel and O. Sporns, “Rich-club organization of the human connectome,” *Journal of Neuroscience*, vol. 31, no. 44, pp. 15 775–15 786, 2011.

[41] V. Latora and M. Marchiori, “Vulnerability and protection of infrastructure networks,” *Physical Review E*, vol. 71, no. 1, p. 015103, 2005.