



Dissecting lightning strike hazard impact patterns to National Airspace System facilities in the contiguous United States

Yiyi He^{a,*}, Xiangyu Yue^b, Sarah Lindbergh^a, Jianxi Gao^c, Chuck Graves^d, Jasenka Rakas^e

^a College of Environmental Design, University of California, Berkeley, CA, United States

^b Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA, United States

^c Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, United States

^d The Federal Aviation Administration, Washington DC, United States

^e Department of Civil and Environmental Engineering, University of California, Berkeley, CA, United States

ARTICLE INFO

Keywords:

Lightning strike hazard
Pattern recognition
National Airspace System
Facility management

ABSTRACT

Lightning strikes pose a severe threat to the United States (US) National Airspace System (NAS). Although the US Federal Aviation Administration (FAA) implements lightning protection practices and procedures to protect personnel, electronic equipment, and structures within the NAS, many lightning-induced outages still occur. To date we found that most research on lightning-induced facility outages has focused on understanding the physical processes of lightning strike effects on aircraft and airport ramp operations. Very little research has been done on examining the overall patterns and characteristics of such hazards to aviation from a geo-spatial standpoint. To bridge this gap, we analyze nationwide lightning strike spatiotemporal data and FAA airport facility outage records from 2009 through 2020 and apply innovative pattern recognition methods to identify key characteristics of lightning strike hazards. Our results uncover the complexities of lightning strike hazard impact patterns to NAS facilities, identifying five distinct typologies with climatological signatures critical to creating better hazard mitigation strategies.

1. Introduction

As a global leader in the air transportation industry, the United States (U.S.) has carried the highest number of passengers for many decades (The World Bank Group, 2021). In addition, the U.S. operates one of the most complex air transportation systems world-wide, and the safety and efficiency of its air traffic operations depend on the reliability of equipment and facilities that compose the National Airspace System (NAS). The Federal Aviation Administration (FAA) is responsible for maintaining and operating over 40,000 NAS facilities and equipment, installed at approximately 6,000 locations across the U.S. (FAA, 2016). These NAS facilities and equipment are critical enablers of airspace surveillance, weather monitoring, aircraft navigation, and airport ground operations. A major impediment to air traffic reliability is the frequency at which the communication, navigation and surveillance (CNS) systems, air traffic control towers, and other infrastructure systems experience failures due to electrical power interruptions (Bates, Seliga, & Weyrauch, 2001). These failures are most often associated with power system transients that result from the occurrence of lightning

strikes in the vicinity of airport facilities. Based on the FAA's system outage database, lightning strike hazard is one of the top three environmental causes for power interruptions of airport facilities, responsible for over 14% of natural hazard-triggered outages, making it one of the most frequent threats to critical equipment and operations at airports (ICF International, 2016). Lightning strikes can cause severe damage to airport equipment and facilities such as localizers, airport surveillance radars, runway lights, air traffic control towers, and communication lines. They can also injure ground personnel on aircraft parking aprons and ramps, air traffic controllers and technicians, and boarding or deplaning staff and passengers (Tarimer, Kuca, & Kisielewicz, 2012).

Lightning strikes are initiated by large electric fields either between cloud and earth, between clouds (inter-cloud) or within clouds (intra-cloud). A high electrically-conductive plasma channel is generated when this electric field reaches 25–30 kilovolt per meter (kV/m) (Ma, Wang, Chen, Wang, & Xu, 2020; Wang et al., 2020), which could lead to a discharge current of in excess of 200 kiloampere (kA) (Uman, 2001). One entire discharge, i.e., a lightning flash, is usually made up of several

* Corresponding author at: 412 Wurster Hall, Berkeley, CA 94720, United States.

E-mail address: yiyi_he@berkeley.edu (Y. He).

<https://doi.org/10.1016/j.compenvurbsys.2021.101735>

Received 29 April 2021; Received in revised form 6 October 2021; Accepted 9 November 2021

Available online 19 November 2021

0198-9715/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

shorter discharges, i.e., single lightning strokes, which last less than a millisecond and repeat rapidly. Lightning flashes possess strong destructive capacity and high hazard potential that may cause serious damage to aircraft (Abdelal & Murphy, 2017; Fisher & Plumer, 1977; Gagné & Theriault, 2014), airport infrastructure (Bejleri et al., 2004) and airspace facilities (Ding & Rakas, 2015; He, Lindbergh, Rakas, & Graves, 2019), either directly through thermal ablation or indirectly through electromagnetic interference (Ma et al., 2020). Direct thermal ablation originates from the interaction of the non-steady lightning plasma channel discharging an electric current, inducing a severe heat flux at the surface of facilities and results in pitting, melt-through and erosion of equipment materials (Wang & Zhupanska, 2015). The indirect effects of lightning are caused primarily by earth-voltage increase, which occur during and after lightning discharge and by the intense electromagnetic field associated with the flash. These fields and earth-voltage rises have enough energy to cause damage a kilometer or more outwards through air or ground, where electromagnetic fields can penetrate significant depths downwards from the actual strike (Rabbani & Oo, 2019). Strikes, on or near overhead conductors, can also cause current and voltage surges, which, independent of distance from the strike, may be transferred along circuit lines and damage connected equipment in the subsystem (Panteli & Mancarella, 2015).

Current climate model projections suggest an intensification of precipitation and convective storm activities in the United States (Lenderink & van Meijgaard, 2008; Liu et al., 2017; Prein et al., 2017a; Prein et al., 2017b) and the temporal frequency, spatial density and severity of lightning strikes are likely to increase in the future (Romps, Seeley, Vollaro, & Molinari, 2014). To climate-proof aviation infrastructures from environmental hazards, especially lightning strikes, the FAA has allocated its resources to improve lightning protection systems. The FAA puts forth standards for Lightning and Surge Protection, Grounding, Bonding, and Shielding (LPGBS) for NAS facilities. Aligned with industry standards for lightning hardening from the Electronic Industries Alliance (EIA), the National Fire Protection Association (NFPA), the Institute of Electrical and Electronic Engineers (IEEE) and others, the FAA standards provide minimum requirements for procurement, design, installation, implementation, operation, and maintenance of LPGBS systems. These industries, led by the FAA, also assign responsibilities for the protection of people, sensitive electrical and electronic equipment, and structures of the NAS operational facilities (FAA, 2018). Although standards and procedures for hardening NAS facilities and infrastructure have significantly improved over time, lightning-induced outages still occur across the NAS. Research on lightning threats to structures on the ground, traditionally developed within electrical and electronic engineering communities, follows national and international standards on lightning protection. However, climatological lightning hazard parameters are restricted to two basic parameters: occurrence and density.

Past approaches to the assessment of climatological patterns associated with lightning hazard and the damage caused by lightning strikes are mainly qualitative or semi-quantitative, relying mostly on expert judgment (Necci et al., 2014). Recent studies on lightning hazard and risk assessment show lightning intensity (i.e. peak current) (He, Lindbergh, Graves, & Rakas, 2020; Shindo & Suda, 2008) along with its occurrence pattern (Tovar, Aranguren, López, Inampué, & Torres, 2014) as predominant climatological hazard characteristics. Few studies incorporated other lightning climatological hazard characteristics such as polarity, average lightning intensity, average steepness of the front impulse current (Yu & Ren, 2014), cumulative value of lightning current (Ishimoto, Asakawa, & Shindo, 2017), multiplicity (He et al., 2019; He et al., 2020), and action integral (i.e. specific energy) (Rakov, 2010). Even fewer studies examined in detail how major climatological lightning hazard parameters might explain lightning-induced outages at large spatial and temporal scales. A comprehensive understanding of the complex lightning hazard impact patterns to NAS facilities is still absent in current aviation research. We bridge this gap through ‘dissecting’ lightning strike hazard impact patterns to NAS facilities in the

contiguous United States (CONUS) by combining spatiotemporal data of lightning strike records from the National Lightning Detection Network (NLDN) and facility outage data created and maintained by the FAA. Our study assessed 5,565 lightning-induced facility outages and their respective lightning strike occurrences impacting 111 types of NAS facilities in the last 11 years. Our results specify, for the first time, spatiotemporal characteristics of NAS-outage related lightning intensity at the CONUS scale which can be generalized to airports suffering from lightning impacts worldwide. In addition, the results show that five typologies, or specific combinations of lightning intensity (i.e. maximum peak current), multiplicity¹, spatial proximity to the impacted facility, and temporal proximity to outage occurrence, better depict climatological hazard parameters of lightning strikes to aviation infrastructure.

2. Methods

2.1. Lightning flash time series clustering

Past research in time series clustering falls into 3 main categories: shape-based (raw-data-based) (Faloutsos, Ranganathan, & Manolopoulos, 1994; Golay et al., 1998; Liu, Maharaj, & Inder, 2014; Sakoe & Chiba, 1978; Vlachos, Kollios, & Gunopulos, 2002; Zhang, Wu, Yang, Ou, & Lv, 2009), model-based (Abfalq et al., 2006; Chen, Nascimento, Ooi, & Tung, 2007; Corduas & Piccolo, 2008; Minnen, Isbell, Essa, & Starner, 2007; Panuccio, Bicego, & Murino, 2002; Smyth, 1996; Xiong & Yeung, 2002) and feature-based (Chan & Fu, 1999; Guo, Jia, & Zhang, 2008; Keogh, Chakrabarti, Pazzani, & Mehrotra, 2001; Keogh & Pazzani, 1998; Möller-Levet, Klawonn, Cho, & Wolkenhauer, 2003; Popivanov & Miller, 2002; Zhang, Ho, Zhang, & Lin, 2006). Shape-based approaches cluster time series by applying modified similarity/distance measure on raw time series that are stretched or contracted along the time axis. Model-based approaches fit time series with parametric models which then translate into model parameters. These parameters are then fed into a clustering algorithm with suitable model distance which yields final partitioning results (Warren Liao, 2005). Feature-based approaches extract a low-dimensional feature vector from each time-series, followed by a conventional clustering algorithm usually with Euclidean distance measurement (Aghabozorgi, Seyed Shirkhorshidi, & Ying Wah, 2015; Hautamaki, Nykanen, & Franti, 2008). Unlike previous approaches, which are either sensitive to noise in data or subject to heavy model assumptions, feature-based methods are more robust to noise, require no user assumptions and can be easily tailored to incorporate specific research objectives. We therefore adopt a unique feature-based clustering approach to identify distinct typologies in lightning flash time series.

In this study, we extract 1,509 lightning flash time series directly associated with NAS facility outages (outage code: 85-3²) for 279 airports over the past 11 years (January 1st, 2009 to June 30th, 2020) from publicly available records provided in the NLDN database based on spatial and temporal criteria. Spatially, lightning flashes within the 15 nautical mile³ (NM) radius area around the Air Traffic Control Towers (ATCTs) of airports are retrieved. According to the statistical model developed by Vidal and Rakas (Vidal and Rakas, 2010), the 15 NM radius is a spatial window that represents the maximum area in which a strike can be considered as a potential trigger to an NAS outage. Temporally, lightning flashes that occurred <1 h prior to the time of an outage (code 85-3) are retrieved. Each lightning flash time series *X*

¹ Number of lightning strokes within a lightning flash.

² The FAA outage data used in the analysis were manually gathered on-site by NAS maintenance technicians. A coded system is used to identify the airport to which each NAS facility belongs and to classify the outages based on their cause. Code 85-3 refers to lightning strike induced facility outages.

³ A nautical mile is a unit of length used in air, marine, and space navigation. One nautical mile is defined as exactly 1,852 meters (6,076 ft; 1.151mi).

contains multiple lightning flash records x_i , each with 5 raw attributes: time of occurrence, polarity, multiplicity, peak current, and geographic location. Leveraging information about the locations of the airports' ATCTs, we are able to calculate the distance of each flash to the ATCT and add it to the existing list of attributes. Here, the distance to ATCT is used as an approximation to the distance between a lightning flash and an individual facility. We also estimate the action integral (Gamerota, Elismé, Uman, & Rakov, 2012; Hirano, Katsumata, Iwahori, & Todoroki, 2010), i.e. energy per unit resistance at the strike point. The equation used to calculate action integral is $\int i^2 dt$ where i is the lightning current. We assume a standardized lightning waveform that is compliant with the U.S. military standard (MIL-STD-464C (DOD, 2010)) based on which the duration of one strike is approximately 500 microseconds. For lightning flashes with a multiplicity greater than one, we assume that the peak current of the first lightning stroke equals the value of the peak current attribute of the entire flash P_{max} , and the peak current of the subsequent strokes is $\frac{P_{max}}{2}$ due to attenuation in the electrical current of lightning waveforms. Moreover, we measure the time gap (in seconds) from the time of each flash x_i to the time of the NAS outage. In total, we summarized 6 unique attributes characterizing each lightning flash time series: polarity, peak current, multiplicity, distance to ATCT, action integral, and time to the outage; i.e. $X = [x_1, x_2, \dots, x_n]$, $x_i \in \mathbb{R}^6$.

Each lightning flash x_i can be mapped into a 6-dimensional space as a point based on its attributes. All lightning flash time series (containing multiple lightning flashes) can therefore be represented as a set of points in the same 6-dimensional space. To identify interpretable attribute combinations in the multidimensional feature space and create a uniform analytical framework for time series with variable lengths, we propose an innovative "grid-based" approach to transform multivariate time series data into N-dimensional vectors. This is achieved through dividing each of the dimensions/attributes into multiple bins recording the counts of lightning flashes under a multidimensional unit (see also Supplementary Table 2). Formally, let us denote A^i the i^{th} attribute, D^i the number of bins for attribute i , and A_{ij} the j^{th} bin of attribute A^i . The resulting feature space would be a cross product of all attributes, and the new feature space has a total of $\prod_{i=1}^6 D_i = 5400$ dimensions. This process transforms each time series into a 5400-D vector, $X \in \mathbb{R}^{5400}$, which is then fed into the pattern recognition stage.

2.2. Pattern recognition

We perform unsupervised clustering on the 5,400-D vectors to identify different lightning hazard impact typologies. It should be noted that most, if not all, unsupervised clustering methods (e.g. k -means) do not directly yield interpretable results. This occurs partially because the clustering outputs are usually determined by all features, making it difficult to discern commonalities between points of the same cluster (Moshkovitz, Dasgupta, Rashtchian, & Frost, 2020). In the field of machine learning, the decision tree is recognized as a canonical example for best model interpretability. Similar to previous works (Frost, Moshkovitz, & Rashtchian, 2020; Moshkovitz et al., 2020), we perform explainable unsupervised clustering in order to provide explicable patterns in lightning time series that are directly associated with NAS outages of the lightning strikes that result in outages.

First, we perform vanilla k -means (MacQueen, 1967) clustering on the processed dataset. Given a set of observations $(X_1, X_2, \dots, X_n)$, where each observation is a D -dimensional vector, vanilla k -means aims to partition the observations into k ($\leq n$) sets $S = \{S_1, S_2, \dots, S_k\}$ so as to minimize the within-cluster sum of squares. Formally, the objective is to find:

$$\arg \min_S \sum_{i=1}^k \sum_{X \in S_i} \|X - \mu^i\|_2^2 \quad (1)$$

where μ^i is the mean of points (i.e. centroid/cluster center) in S_i , and $\|\cdot\|$ represents the ℓ^2 norm of the vector. To decide the optimal value for parameter k , we calculate the Inertia value and the Silhouette Score. Inertia evaluates the internal coherence of clusters by calculating the sum of squared distances between each instance and its assigned centroid: $\sum_{j=1}^n X_j - \mu^{y^j}$, where y^j is the index of the closest cluster center to X_j , i.e. $y^j = \operatorname{argmin}_{1 \leq j \leq k} \|X^j - \mu^j\|_2^2$. Smaller inertia values suggest better

clustering outcomes. Intuitively, larger k yields smaller inertia yet the resulting clusters shrink in size as well as in between-cluster dissimilarity. The elbow method is often applied in practice to determine the smallest k with diminishing returns in terms of inertia decrease. Silhouette Coefficient (Rousseeuw, 1987) is defined as $\frac{p-q}{\max(p,q)}$, where q is the mean distance to the other instances in the same cluster and p is the mean distance to the instances of the next closest cluster. After inertia and Silhouette Score calculation, $k = 5$ is chosen as the optimal number of clusters and $\mu^1, \mu^2, \dots, \mu^k$ represent cluster centers after applying vanilla k -means.

Second, we build a decision tree top-down with binary splits. A tree structure is particularly useful in high-dimensional space since the number of clusters is much smaller than the input dimension ($5 \ll 5,400$). The tree consists of a root node, k leaves and multiple internal nodes, each containing a single decision feature and an associated threshold value. Each decision node (the root node and all internal nodes) partitions the input space into 2 hyper-rectangular cells with no overlaps. Starting at the root node, an initial feature i with a threshold is selected such that it produces the least mistakes when parsing input data through the node. A mistake occurs when a data point X^j and its corresponding cluster center μ^{y^j} are partitioned into different cells at the decision node. With dynamic programming, we search for the optimal splitting rule (i, τ) efficiently (Eq. (2)). Next, if 2 or more cluster centers fall within the same cell, another split is legitimized. All data points whose cluster centers lie outside of the cell are intentionally left out since they are considered as mistakes in the previous splits. This process is performed recursively on the left and right children until each leaf node is pure (i.e., contains only one cluster center).

$$i, \tau = \operatorname{argmin}_{i, i_L \leq \tau \leq i_R} \sum_{j=1}^m \operatorname{mistake}(X^j, \mu^{y^j}, i, \tau),$$

$$\operatorname{mistake}(X^j, \mu^{y^j}, i, \tau) = \begin{cases} 1 & (X_i^j \leq \tau) \neq (\mu_i^{y^j} \leq \tau) \\ 0 & \text{else} \end{cases} \quad (2)$$

It should be noted that the true representation space of individual clusters are not necessarily hyper-rectangular. The approach to creating a binary tree above is likely to introduce large bias in classification results. In other words, data points are systematically misclassified due to the rigid decision boundary of the binary tree classifier that has few leaf nodes. To balance the bias-variance tradeoff in this approach, we use ExKMC (Frost et al., 2020) to allow more flexibility to the previous decision tree. We expand the original binary tree with k leaf nodes to one with k' nodes ($k' \geq k$) based on a surrogate cost and associate/classify each of the k' leaves with one of the cluster centers i . The surrogate cost is defined as follows: given k cluster centers $M = [\mu^1, \mu^2, \dots, \mu^k]$, and a binary tree T which defines clustering $(\hat{C}^1, \hat{C}^2, \dots, \hat{C}^{k'})$, the surrogate

$$\text{cost } \widetilde{\text{cost}}^M(T) = \sum_{j=1}^{k'} \sum_{X \in \hat{C}^j} \|X - \mu^{c^j}\|_2^2, \text{ where } c^j = \operatorname{argmin}_{i \in [1,k]} \sum_{X \in \hat{C}^j} \|X - \mu^i\|_2^2.$$

For each leaf node n_{leaf} with associated samples X_{leaf} , the gain by splitting n_{leaf} is calculated as (Eq. (3)), where $X_L = \{x \in X | x_i \leq \tau\}$ and $X_R = \{x \in X | x_i > \tau\}$. Finally, we split the leaf node with the largest gain into two child nodes n_{new_L} and n_{new_R} and compute $\text{gain}(n_{new_L})$ and $\text{gain}(n_{new_R})$ respectively. This expansion process is repeated recursively until the tree has exactly k' leaf nodes. In our analysis, $k' = 2k = 10$ is chosen as the optimal value for all experiments in this study based on sensitivity

testing.

$$\text{gain}(n_{\text{leaf}}) = \widehat{\text{cost}}^M(X_{\text{leaf}}) - \widehat{\text{cost}}^M(X_L) - \widehat{\text{cost}}^M(X_R), \quad (3)$$

2.3. Critical feature identification

Due to the non-deterministic characteristic of the above unsupervised clustering method, the resulting clusters may not accurately describe the semantic structure of the input lightning flash time series. To mitigate the randomness and capture the semantic structure more robustly, we propose to repeat the clustering process on the lightning dataset m times (m is set to 10,000 in this study), and assemble the clustering results into a fully connected graph.

Formally, let's denote the set of lightning flash time series as $\mathcal{X} = \{X^1, X^2, \dots, X^n\}$, and m clustering models as $M = \{M_1, M_2, \dots, M_m\}$, in which $M_i(X^j)$ returns the cluster label of time series sample X^j in the i^{th} cluster model. Combining $\mathcal{X}$ and M , we build a weighted complete graph $G = (V, E)$, where each node n_j represents a lightning time series X^j and an edge (n_i, n_j) denotes that n_i and n_j are in the same cluster. Theoretically, this complete graph has n nodes and $\frac{n(n-1)}{2}$ edges, with edge weight w_{ij} which corresponds to the number of times (out of m) that $M_i(X^i)$ and $M_i(X^j)$ returns the same cluster label i.e. $w_{ij} = |\{m : m \in M, m(X^i) = m(X^j)\}|$. An edge (n_i, n_j) with large weight suggests a closer relationship between X^i and X^j since they are often grouped into the same cluster. After the graph is created, we prune edges with smaller weights with a specific threshold, i.e., (n_i, n_j) is removed from G if $w_{ij} < \tau$, where τ is the threshold hyper-parameter. We test the effect of different values

(from small to large) on pruning results and discover that the resulting graph G' is relatively stable even when taking on larger values such as 5,000. After pruning, we identify k cliques/connected components within G' to arrive at the final clustering result.

To pinpoint a subset of features (among 5,400 total) that best explains the key characteristics of time series that share the same cluster label, a critical feature identification process is developed to evaluate the importance of features. Leveraging information of the decision nodes in binary trees created beforehand, we record the number of occurrences C_i when a feature is identified to be associated with a decision node $C_i = [C_{i1}, C_{i2}, \dots, C_{iD}]$, where $D = 5,400$. Then we compute the cumulative count over all decision trees: $C_{\text{acc}} = \sum_i C_i = [C_1, C_2, \dots, C_D]$ and sort out top 10 features with higher C_{acc} (see Supplementary Table 3 for details).

3. Results

3.1. Beyond the peak current

Peak current, or lightning intensity, is shown to be highly correlated with thermal and mechanical damage typologies to infrastructure that are exposed to lightning hazards in empirical studies (Ma et al., 2020; Millen, Murphy, Catalanotti, & Abdelal, 2019; Muñoz et al., 2014). We first explore spatiotemporal characteristics of NAS outage-related lightning strike peak current and go beyond to show how specific combinations of different climatological lightning hazards parameters can become better indicators of lightning damage potential to NAS

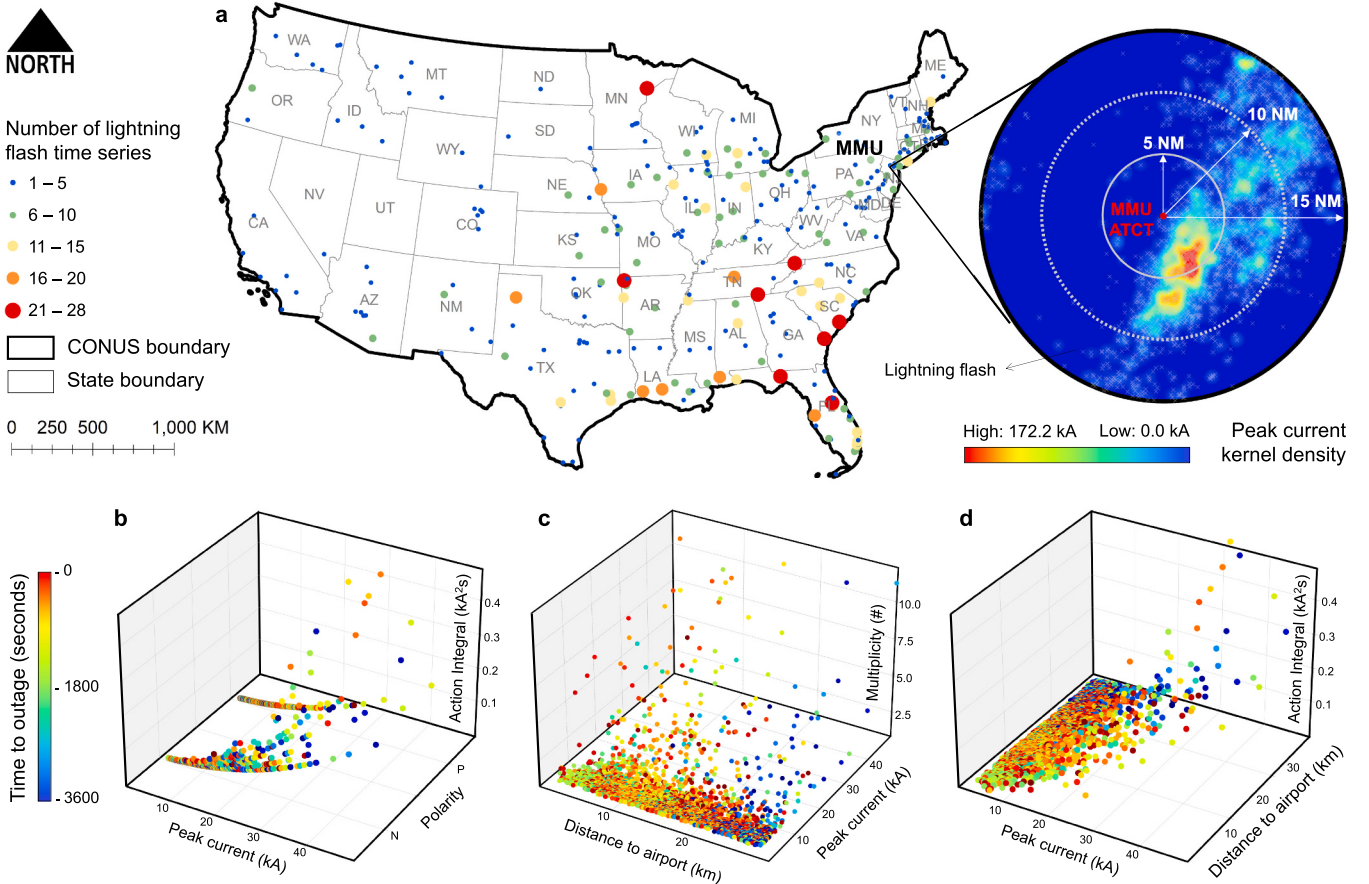


Fig. 1. Lightning strike time series extracted from NLDN. a. Distribution of the total number of lightning flash time series extracted for airports in CONUS and spatial criterion for lightning flash time series extraction using Morristown Airport (MMU) as an example. All flashes within 15 NM radius from the location of MMU's ATCT are included in initial consideration. b–d. 3D scatter plots of combinations of lightning flash attributes (polarity, peak current, multiplicity, distance to airport, action integral, and time to the outage) using one lightning flash time series at MMU airport.

facilities. In this study, a total of 1,509 lightning strike time series directly associated with NAS facility outages (cause code 85-3) were extracted from the NLDN database over the past 11 years for a total of 284 distinct airports in the CONUS (Fig. 1a). Each time series contains multiple lightning flash records with attributes indicating the location where the flash occurred, peak current of the first stroke, polarity, and multiplicity (Fig. 1b–d). A spatiotemporal screening criterion was applied to include lightning flashes that occurred within 15 NM radius of the airports' air traffic control tower (ATCT) and 1 h time window to the NAS outages (see Methods for details). The spatial distribution of these time series reveal a heterogeneous lightning hazard impact pattern with more lightning-induced outage occurrences in Northeastern, Southeastern and Midwestern states than Western and Southwestern states. Airports, namely Savannah/Hilton Head International Airport (SAV), Duluth International Airport (DLH), Tri-Cities Airport (TRI), Chattanooga Airport (CHA), and Charleston International Airport (CHS) have the highest outage occurrences of all-28, 28, 27, 26 and 25 respectively (see also Supplementary Table 1 for more details).

Our spatiotemporal assessment of the maximum peak current distribution of NAS outage-related lightning strikes in the CONUS (Fig. 2) shows that the kernel density curve of the maximum peak current values of all time series peaks around 100 kA with a minimum of 20.2 kA and a maximum of 427.6 kA (Fig. 2a). To investigate the spatial distribution of maximum peak current values, we grouped the lightning strikes maximum peak current values by airport and extracted the maximum and minimum values of the maximum peak current. The 75th percentile of the maximum values is 187.85 kA, which suggests that the majority of the maximum peak current observed at airports is below this threshold. The minimum values exhibit a unique distribution with 66.25 kA as the 50th percentile and 47.77 kA as the 25th percentile (Fig. 2b). This suggests that for some airports, lightning strikes with relatively lower maximum peak current could still induce NAS outages. Fig. 2c shows the spatial distribution of the maximum and minimum values of the maximum peak current at each airport overlaid on a kernel density estimation map of NAS facility outages in the past 11 years. We can observe that the variation in minimum values across the CONUS is

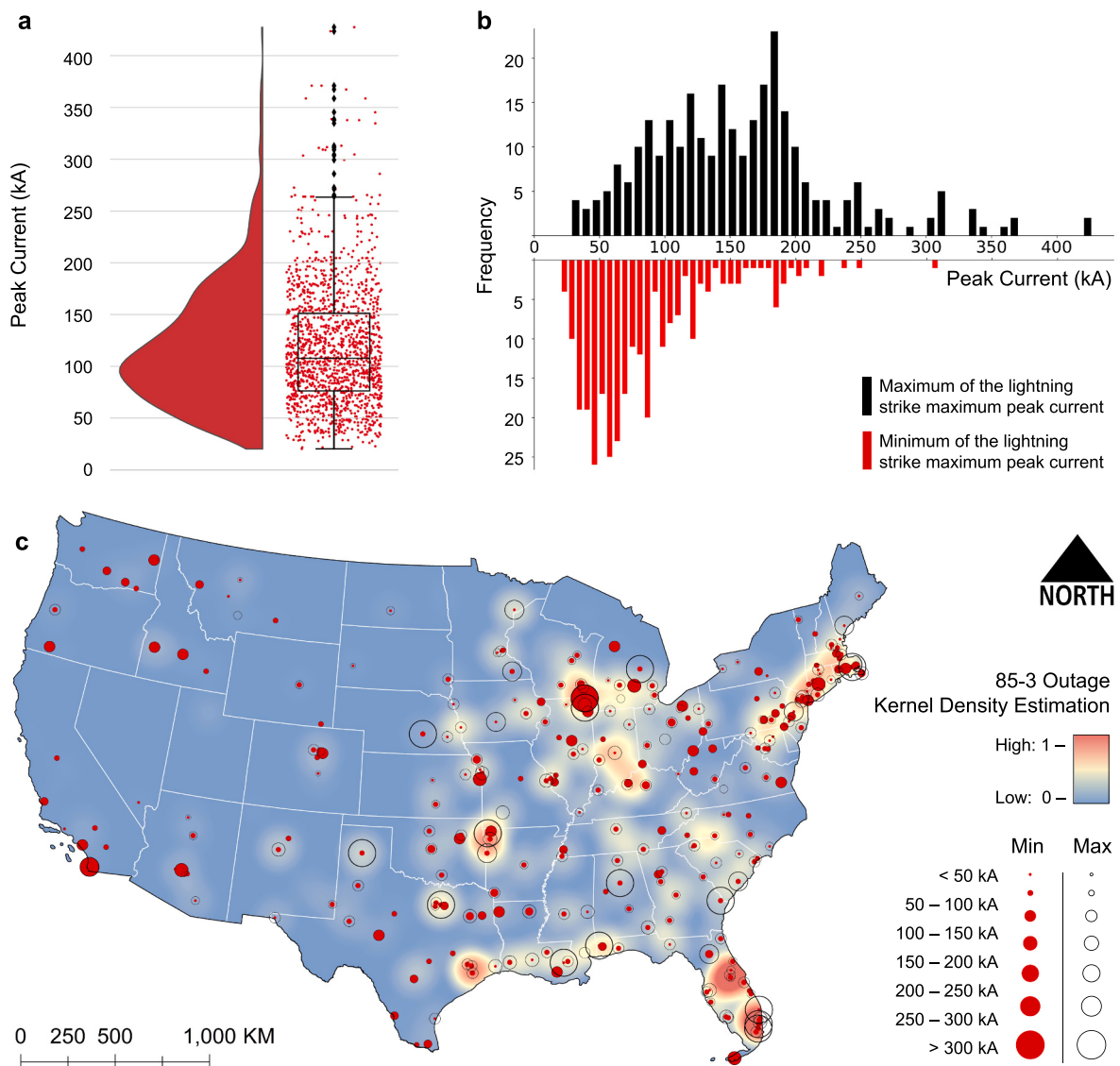


Fig. 2. Maximum peak current distribution across all airports in the CONUS. a. Density, box and scatter plots of the maximum peak current values of all lightning strike time series. b. Histograms of maximum and minimum values of the maximum peak current of the lightning time series grouped by airport. c. Maximum and minimum values of the maximum peak current grouped by airport overlaid on kernel density estimation plot based on the number of 85-3 outages from 2009 through 2020.

smaller than that of the maximum values. The disparity between minimum and maximum values are more obvious in the Eastern states and spatially coincides with areas where more NAS facility outages are observed. In general, lightning time series with a higher maximum peak current is more likely to induce facility outages. This pattern can be observed in areas near Chicago, Boston, and southern Florida. However, areas near southern Indiana, Huston and Philadelphia display contrasting patterns which indicate that the underlying lightning hazard impact pattern is complex and spatially heterogeneous. The subsequent sections will discuss distinct patterns in lightning time series that are associated with NAS facility outages.

3.2. New climatological lightning hazard typologies

Our analysis identifies five distinct typologies of lightning strike hazard patterns that are directly associated with NAS outages. The first typology (T1) involves occurrence(s) of one or more lightning flashes with high peak current within 5 kilometers (km) from the ATCT of the airport (Fig. 3 first column). The peak current values for the majority of

lightning flash time series in this typology concentrates in the 70–100 kA range with a minimum of 34 kA (Palm Beach International Airport (PBI)) and a maximum of 371 kA (Northwest Arkansas National Airport (XNA)). In general, lightning flash time series in the T1 category can be characterized as high lightning intensity with close distance to airports (Fig. 3b and c). The flashes in these time series carry a large amount of electromagnetic energy and usually occur right before the time of the outage. This typology aligns well with general assumptions about lightning hazard impact patterns reflected in dominant lightning protection and hardening standards. According to the FAA's standard for lightning and surge protection, grounding, bonding, and shielding requirements for facilities and electronic equipment, NAS facilities are required to install lightning protection systems to provide preferred paths for lightning discharges to enter or leave the earth without causing damage to facility or equipment or injury to personnel. Our results show that in areas with frequent lightning activity and high lightning flash intensity, such hazard typology still plays a dominant role in inducing facility outages. At Northeast, Mid-Atlantic and Southeastern state airports and airports in cities such as the New York City, T1 is the dominant

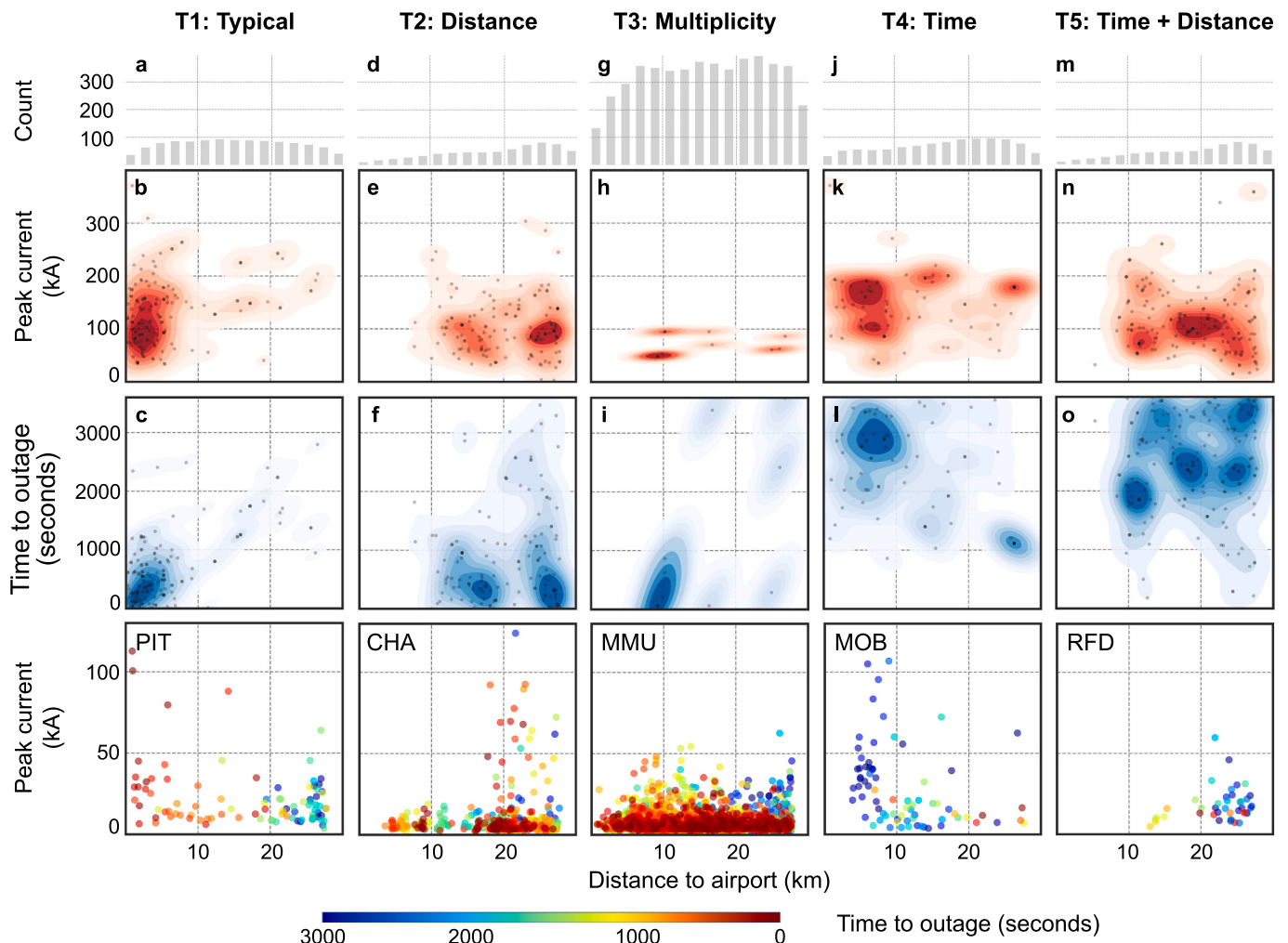


Fig. 3. Distribution of sum of flash multiplicity for five lightning flash time series typologies, respective bivariate kernel density estimate of lightning flash attributes and examples. Each column (e.g. a, b, c) consists of three visualizations of lightning flash time series attributes that belong to a specific typology (e.g. T1) and an specific airport example. The first row (a, d, g, j, m) are histograms of the number of the lightning flash time series that fall into 10 distance ranges (minimum: 0, maximum: 30, step: 3). The second and third rows are 2-D kernel density contour plots of 2 combinations of three key lightning flash attributes: peak current, time to outage and distance to airport. For each typology, we extract the lightning flash with maximum peak current value for all lightning flash time series and find corresponding distance to airport and time to outage attributes. Each extracted lightning flash is represented as a black dot in the contour plots based on which the bivariate kernel density estimations are generated. The last row contains examples of each typology at airports. The three letter identification code of the airport is labeled on the top left corner of the subplots. Each dot in the subplot represents the occurrence of a lightning flash color-coded by the time of its occurrence. Dark red indicates that the flash occurred right before the outage whereas dark blue indicates otherwise. For more examples please see Supplementary Fig. 1.

lightning hazard pattern that causes outages (Fig. 4a).

The second typology (T2) can be characterized by higher occurrences of lightning flashes at further distances (10–20 km or, in some cases, nearly 30 km) from the ATCT with high peak current. Adjacent to the airport T2 typology presents absence or low lightning activity (Fig. 3 second column). Our pattern signature identification also shows comparable results indicating that longer distances to ATCT (e.g. 10, 20, and 30 km thresholds) are strong features in characterizing such lightning strike hazard patterns. Bivariate kernel density estimate results show that the time of high intensity lightning flash occurrence from these time series is generally closer to the time of outage (Fig. 3f). Current LPGBS requirements, as well as the latest revisions of the National Fire Protection Association Standard for the Installation of Lightning Protection Systems (NFPA 780), put much focus on protecting facilities that are located within the airport boundary focusing on ATCT, air terminals, and conductors. Facilities beyond the airport boundary and/or infrastructure (that are not owned by the FAA) are often neglected for lightning hazard assessment. The International Air Transport Association (IATA) recommends airports to issue lightning alerts when lightning strikes are detected at 8 km from the airport boundary, and stop high risk operations when they are detected at 5 km. They suggest using a 5 km critical radius and resuming operations once the lightning activity has moved beyond this radius (IATA, 2021; *Earth Networks*, 2021). Our results suggest that lightning flash occurrences further away from airports (20–30 km) could also induce NAS outages. This typology is more evident in Midwestern and Southeastern states where it would be desirable to review the critical radius for certain high risk operations (Fig. 4). The physical and cyber interconnections and interdependencies between facilities within the airport boundary and infrastructures in its vicinity and/or beyond could offer one explanation for this phenomenon.

The third typology (T3) can be summarized by large quantities of lightning flash occurrences both adjacent and far away from the ATCT with relatively low peak current (Fig. 3 third column). This typology is relatively rare compared with other typologies with a total of 11 lightning flash time series samples covering 8 airports: Jack Brooks Regional Airport (BPT), Fort Smith Regional Airport (FSM), Groton-New London Airport (GON), Hagerstown Regional Airport (HGR), Morristown Airport (MMU), Dane County Regional Airport (MSN), South Bend International Airport (SBN), and Toledo Express Airport (TOL). Spatially, these airports are located in Northeastern states as well as Southwestern states, namely Texas and Oklahoma. Aside from peak current intensity

and distance, this lightning hazard typology points to the importance of considering the chronic impact of lightning flashes with high multiplicity and lower intensity strikes. Recognizing and acknowledging that low intensity and high multiplicity lightning flashes can help better characterize lightning hazard to NAS facilities in the context of airport facility reliability, is essential for a successful LPGBS program. Lightning hazard protection of buildings and other safety recommendations are well documented and have precise technical standards such as those referred by the NFPA 780. Such protection measures have proven to reduce the probability of outages from lightning strike occurrence, but residual risk remains (He et al., 2019).

The fourth (T4) and fifth (T5) typologies are both related to the temporal pattern of the lightning flash occurrences. The key feature that differentiates T4 from the previous three typologies is early occurrence (Fig. 3 fourth column). Lightning flashes that took place as early as 1 h prior to the time of the reported outage could induce NAS facility failures. This result aligns well with our critical feature identification results which highlights time of occurrence as a key feature in differentiating lightning hazard patterns. In terms of spatial distribution, T4 is dominant in states such as Michigan, Oklahoma, Texas, New Jersey, and Florida. T5 can be considered as the multiplication between T2 and T4. Lightning flash time series that belong to T5 are often distant from the airport ATCT both spatially and temporally (Fig. 3 fifth column). This typology is shown to be most notable in Southern states such as Tennessee, Oklahoma, and Northern states such as Massachusetts. The wide geographical distribution and common occurrence of T4 and T5 calls for a review of temporal thresholds of early warning systems at airports to improve organizational preparedness to potential operational impacts. The current lightning warning systems provide two levels of warning: an alert, indicating that lightning may develop or move into the area of protection in the near future, and an alarm, indicating that lightning has been detected in the immediate vicinity or is expected to develop at any moment. While there are no universally recognized standards for issuing alerts or alarms for airport operations, the American Meteorological Society and the National Oceanic and Atmospheric Administration (NOAA) have endorsed the “30–30 rule” which states that operations should be limited or curtailed whenever there has been a lightning strike detected within 6 miles (approximately 9.6 km) (based on 30 s between an observed flash and the sound of the thunder) and within 30 min (Heitkemper & Johnson, 2008). The fourth and fifth typologies identified in our analysis, however, in part contradict this recommendation and call for lightning warning thresholds for longer time periods and

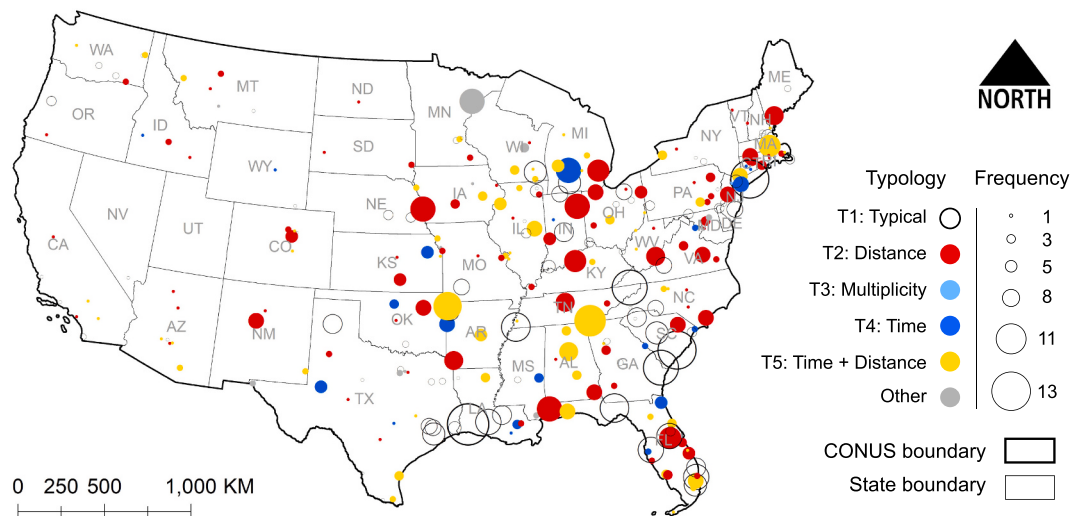


Fig. 4. Spatial distribution of typologies in the CONUS. The spatial distribution of the mode of lightning strike hazard pattern at the national scale. The size of the circle/dot represents the frequency of the corresponding typology. Refer to Supplementary Fig. 2 for individual distribution plots for each typology.

greater spatial distances. This could potentially induce organizational shifts in early warning systems with the creation of a “pre-alert phase” where operational procedures seek to mitigate the risk of facility outages due to lightning strikes.

4. Conclusion

The impact of lightning strike hazard in the vicinity of, and on, airport operating areas has long been recognized by airport and airline operators as both a safety and an operational challenge. The cost to the flying public of lightning induced outages (which in turn cause delays and cancellations) can be almost five times the cost of retrofitting an existing facility (Ding & Rakas, 2015; He et al., 2020). Our analysis identifies new lightning strike hazard impact patterns on NAS facilities based on historical facility outage records maintained by the FAA and a nationwide spatiotemporal lightning strike dataset from NLDN over the past 11 years. Our findings shed new light into the design of lightning detection and warning systems at airports and LPGBS at NAS facilities.

Airports are high reliability organizations (Roberts, 1989) where safety goals override performance goals, but it is difficult to incorporate higher restrictions on operations due to lightning activity. For airport business continuity, there is a delicate balance between preventive lightning procedures and operational reliability, but our results show how airport operators can be better prepared to deal with potential lightning-induced outages. Four of the five distinct lightning flash time series typologies (T2-5) help identify airports with atypical hazard impact patterns (Fig. 5) and unique risk profiles that may benefit from distinct lightning hazard mitigation strategies. Airport operators can benefit from new insights into how lightning strikes damage their facilities and employ more accurate lightning observation and detection technology to develop robust operational management procedures. Airports presenting more atypical typologies have a higher motivation to revisit their operational management systems including thresholds for early warnings based on the onset and duration of lightning. Current aviation lightning protection and prevention policies employed at U.S.

and other exposed airports worldwide can be improved using spatial and temporal lightning detection thresholds uncovered in this study. New lightning impact patterns detected here help prioritize investment decisions and present new and low cost alternatives for airport safety managers to improve current procedures and policies for lightning strike resilience beyond traditional infrastructure hardening.

5. Discussion

It should be noted that our results are more robust for T1, T2, T4 and T5 where samples are relatively abundant. T3 typology is rarer and therefore we have less information on the distribution and occurrence frequencies locally at airports and throughout the United States. In addition, airports categorized with the same hazard typology do not necessarily translate to the same lightning hazard impact mechanism. Other important attributes of NAS facilities that should be considered are: NAS facility type, location, age, fault tolerance and infrastructure connectivity, along with existing lightning protection requirements and operational procedures at place. Since all lightning-induced outages are recorded manually by FAA field technicians, it is likely that the total number of outages (and associated lightning strike hazard typologies) are underestimated (He et al., 2019). In other words, there might exist additional climatological lightning hazard typologies beyond the five typologies identified in this article.

Lightning strike hazard not only impacts the NAS facilities but also interconnected critical infrastructure systems such as the electricity grid (Maruyama Rentschler, Obolensky, & Kornejew, 2019). The impact of lightning strikes on humans in airports usually involves personnel operating sensitive equipment in buildings without lightning protection, grounding, bonding, and shielding (Ding & Rakas, 2015; World News, 2014) and especially personnel working outdoors (Steiner, Deierling, Ikeda, Nelson, & Bass, 2014; Steiner, Deierling, & Johnson, 2012). In contrast to the greatly improved knowledge of lightning occurrences, the distribution of lightning-related human fatalities and injuries is not well characterized in many regions, especially in lesser-developed

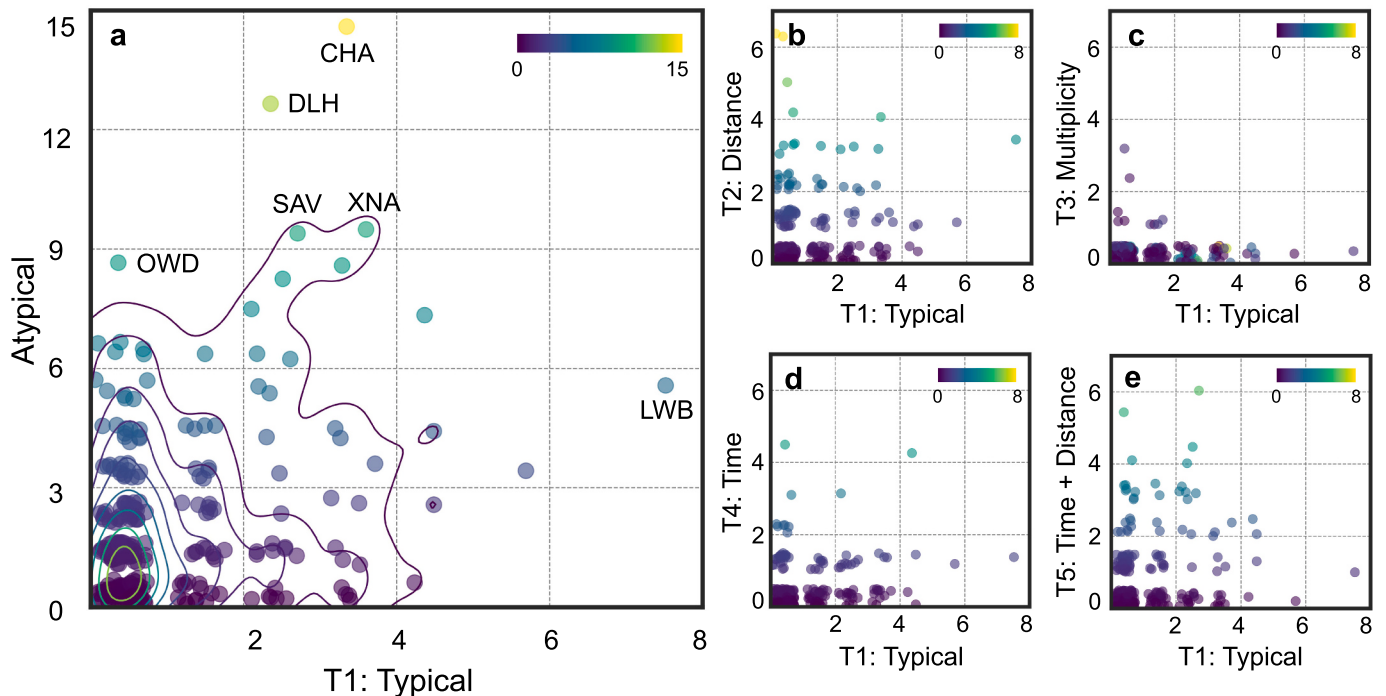


Fig. 5. Distribution of climatological lightning hazard typologies at airports. a. The distribution of typical (T1) and atypical (T2-5 and others) typology counts at airports overlaid on kernel density contours. Airports in the top right corner of the graph have more atypical than typical hazard typologies. b-e. The distribution T1 and T2, T3, T4 and T5 typology counts at airports respectively. The location of points in these scatter plots are offset by adding a random number from 0 to 0.5 to avoid overlaps. The color of the points corresponds to the value of the variable plotted as the vertical axis.

nations (Dewan, Hossain, Rahman, Yamane, & Holle, 2017) where outdoors workers are highly impacted (Adhikari, 2021; Holle et al., 2019). In developed countries, lightning fatalities have been greatly reduced during the last century thanks to economic advances that provide lightning-safe structures and dwellings as well as transitions in industrial structure (Holle, 2016). In developing countries, however, the impact of lightning strike hazard is heavily underestimated and the majority of the population continues to be engaged in subsistence agriculture for long periods, live in lightning-unsafe dwellings, and work in lightning-unsafe structures (Dlamini, 2009). The pattern recognition methods proposed in this study can be adapted to facilitate the identification of other impact typologies involving injuries and fatalities, especially in developing countries where such research is scarce. Similar spatiotemporal thresholds applied to human impact patterns could hence be applied to create or update early warning systems which directly mitigate the risk of injury or fatalities.

Data availability

Historical lightning strike data from NLDN can be accessed from https://ghrc.nsstc.nasa.gov/home/lightning/index/data_nldn

Code availability

The Python code used to complete the analysis and produce the figures in this study will be available in the following online repository [https://github.com/jesuslovesyiyi/Lightning_outage_pattern.git].

Ethics declarations

The authors declare no competing interests.

Author contributions

Y.H. and J.R. conceived the study. Y.H., X.Y. performed analytical analysis and interpreted the results. S.L., C.G., J.R., and J.G. discussed the results and provided useful edits to the manuscript.

Acknowledgements

This research was funded by the Federal Aviation Administration (FAA) through Contract 693KA9-18-D-00015 Delivery Order 693KA9-19-F-00124, The Impact of Lightning Strikes on The National Airspace System (NAS) Critical Infrastructure and NAS Performance for ASPM Airports. J.G. acknowledges the support of National Science Foundation under Grant No. 2047488, and the Rensselaer-IBM AI Research Collaboration.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compenvurbsys.2021.101735>.

References

- Abdelal, G. F., & Murphy, A. (2017). A multiphysics simulation approach for efficient modeling of lightning strike tests on aircraft structures. *IEEE Transactions on Plasma Science*, 45, 725–735.
- Adhikari, B. R. (2021). Lightning fatalities and injuries in Nepal. *Weather, Climate, and Society*, 13, 449–458.
- Aghabozorgi, S., Seyed Shirkhorshidi, A., & Ying Wah, T. (2015). Time-series clustering – A decade review. *Information Systems*, 53, 16–38.
- Abfal, J., et al. (2006). Similarity search on time series based on threshold queries. In Y. Ioannidis, et al. (Eds.), *Advances in database technology - EDBT 2006* (pp. 276–294). Springer. https://doi.org/10.1007/11687238_19.
- Bates, E. G., Seliga, T. A., & Weyrauch, T. (2001). Implementation of grounding, bonding, shielding and power system improvements in FAA systems: Reliability assessments of the Terminal Doppler Weather Radar (TDWR). In, vol. 1 2A5/7. *AIAA/IEEE Digital avionics systems conference – Proceedings*.
- Bejleri, M., et al. (2004). Triggered lightning testing of an airport runway lighting system. *IEEE Transactions on Electromagnetic Compatibility*, 46, 96–101.
- Chan, K.-P., & Fu, A. W.-C. (1999). Efficient time series matching by wavelets. In *Proceedings 15th international conference on data engineering (Cat. No.99CB36337)* (pp. 126–133). <https://doi.org/10.1109/ICDE.1999.754915>
- Chen, Y., Nascimento, M. A., Ooi, B. C., & Tung, A. K. H. (2007). SpAdE: On shape-based pattern detection in streaming time series. In *2007 IEEE 23rd international conference on data engineering* (pp. 786–795). <https://doi.org/10.1109/ICDE.2007.367924>
- Corduas, M., & Piccolo, D. (2008). Time series clustering and classification by the autoregressive metric. *Computational Statistics & Data Analysis*, 52, 1860–1872.
- Dewan, A., Hossain, M. F., Rahman, M. M., Yamane, Y., & Holle, R. L. (2017). Recent lightning-related fatalities and injuries in Bangladesh. *Weather, Climate, and Society*, 9, 575–589.
- Ding, W., & Rakas, J. (2015). Economic impact of a lightning strike-induced outage of air traffic control tower: Case study of Baltimore–Washington International Airport. *Transportation Research Record: Journal of the Transportation Research Board*, 2501, 76–84.
- Ding, W., & Rakas, J. (2015). Economic impact of a lightning strike-induced outage of air traffic control tower case study of Baltimore–Washington International Airport. *Transportation Research Record*, 76–84. <https://doi.org/10.3141/2501-10>
- Dlamini, W. M. (2009). Lightning fatalities in Swaziland: 2000–2007. *Natural Hazards*, 50, 179–191.
- DOD. (2010). *Electromagnetic environmental effects requirements for systems*.
- FAA. (2016). *The future of the NAS* (vol. 44).
- FAA. (2018). *JO 6900.25 – Lightning and surge protection, grounding, bonding and shielding (LPGBS) policy*.
- Faloutsos, C., Ranganathan, M., & Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. *ACM SIGMOD Record*, 23, 419–429.
- Fisher, F. A., & Plumer, J. A. (1977). *Lightning protection of aircraft*. National Aeronautics and Space Administration, Scientific and Technical Information Office.
- Frost, N., Moshkovitz, M., & Rashtchian, C. (2020). ExKMC: expanding explainable k-means clustering. *arXiv:2006.02399 [cs, stat]*.
- Gagné, M., & Theriault, D. (2014). Lightning strike protection of composites. *Progress in Aerospace Sciences*, 64, 1–16.
- Gamerota, W. R., Elismé, J. O., Uman, M. A., & Rakov, V. A. (2012). Current waveforms for lightning simulation. *IEEE Transactions on Electromagnetic Compatibility*, 54, 880–888.
- Golay, X., et al. (1998). A new correlation-based fuzzy logic clustering algorithm for FMRI. *Magnetic Resonance in Medicine*, 40, 249–260.
- Guo, C., Jia, H., & Zhang, N. (2008). Time series clustering based on ICA for stock data analysis. In *2008 4th international conference on wireless communications, networking and mobile computing* (pp. 1–4). <https://doi.org/10.1109/WiCom.2008.2534>
- Hautamaki, V., Nykanen, P., & Franti, P. (2008). Time-series clustering by approximate prototypes. In *2008 19th international conference on pattern recognition* (pp. 1–4). <https://doi.org/10.1109/ICPR.2008.4761105>
- He, Y., Lindbergh, S., Graves, C., & Rakas, J. (2020). Airport exposure to lightning strike hazard in the contiguous United States. *Risk Analysis*. <https://doi.org/10.1111/risa.13630>
- He, Y., Lindbergh, S., Rakas, J., & Graves, C. M. C. (2019). Characterizing lightning-strike hazard to airport facilities: A case study of baltimore washington international airport. In *2019 integrated communications, navigation and surveillance conference (ICNS)* (pp. 1–9). <https://doi.org/10.1109/ICNSURV.2019.8735415>
- Heitkemper, Lawrence, Johnson, David, & Transportation Research Board. (2008). *Lightning-warning systems for use by airports*. Washington D.C.: National Academies Press.
- Hirano, Y., Katsumata, S., Iwahori, Y., & Todoroki, A. (2010). Artificial lightning testing on graphite/epoxy composite laminate. *Composites Part A: Applied Science and Manufacturing*, 41, 1461–1470.
- Holle, R. L. (2016). A summary of recent national-scale lightning fatality studies. *Weather, Climate, and Society*, 8, 35–42.
- Holle, R. L., et al. (2019). Fatalities related to lightning occurrence and agriculture in Bangladesh. *International Journal of Disaster Risk Reduction*, 41, Article 101264.
- IATA. (2021). *ISAGO standards manuals*.
- ICF International, et al. (2016). *Addressing significant weather impacts on airports: Quick start guide and toolkit*. Transportation Research Board. <https://doi.org/10.17226/23629>
- Ishimoto, K., Asakawa, A., & Shindo, T. (2017). Study of lightning risk evaluation of distribution systems. *Electrical Engineering in Japan*, 198, 34–45.
- Keogh, E., Chakrabarti, K., Pazzani, M., & Mehrotra, S. (2001). Locally adaptive dimensionality reduction for indexing large time series databases. *SIGMOD Record*, 30, 151–162.
- Keogh, E. J., & Pazzani, M. J. (1998). An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. In *Proceedings of the fourth international conference on knowledge discovery and data mining* (pp. 239–243). AAAI Press.
- Lenderink, G., & van Meijgaard, E. (2008). Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, 1, 511–514.
- Liu, C., et al. (2017). Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, 49, 71–95.
- Liu, S., Maharaj, E. A., & Inder, B. (2014). Polarization of forecast densities: A new approach to time series classification. *Computational Statistics & Data Analysis*, 70, 345–361.
- Ma, X., Wang, F., Chen, H., Wang, D., & Xu, B. (2020). Thermal damage analysis of aircraft composite laminate suffered from lightning swept stroke and arc propagation. *Chinese Journal of Aeronautics*, 33, 1242–1251.

- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In , vol. 1. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*.
- Maruyama Rentschler, J. E., Obolensky, M. A. B., & Kornejew, M. G. M. (2019). Candle in the wind? Energy system resilience to natural shocks. <https://papers.ssrn.com/abstract=3430507>.
- Millen, S. L. J., Murphy, A., Catalanotti, G., & Abdelal, G. (2019). Coupled thermal-mechanical progressive damage model with strain and heating rate effects for lightning strike damage assessment. *Applied Composite Materials*, 26, 1437–1459.
- Minnen, D., Isbell, C. L., Essa, I., & Starner, T. (2007). *Discovering multivariate motifs using subsequence density estimation and greedy mixture learning* (vol. 1, pp. 615–620).
- Möller-Levet, C. S., Klawonn, F., Cho, K.-H., & Wolkenhauer, O. (2003). Fuzzy clustering of short time-series and unevenly distributed sampling points. In M. Berthold, H.-J. Lenz, E. Bradley, R. Kruse, & C. Borgelt (Eds.), *Advances in intelligent data analysis V* (pp. 330–340). Springer. https://doi.org/10.1007/978-3-540-45231-7_31.
- Moshkovitz, M., Dasgupta, S., Rashtchian, C., & Frost, N. (2020). Explainable k-means and k-medians clustering. In *International conference on machine learning* (pp. 7055–7065). PMLR.
- Muñoz, R., et al. (2014). Modeling lightning impact thermo-mechanical damage on composite materials. *Applied Composite Materials*, 21, 149–164.
- Necci, A., et al. (2014). Assessment of lightning impact frequency for process equipment. *Reliability Engineering & System Safety*, 130, 95–105.
- Panteli, M., & Mancarella, P. (2015). Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. *Electric Power Systems Research*, 127, 259–270.
- Panuccio, A., Bicego, M., & Murino, V. (2002). A hidden markov model-based approach to sequential data clustering. In T. Caelli, A. Amin, R. P. W. Duin, D. de Ridder, & M. Kamel (Eds.), *Structural, syntactic, and statistical pattern recognition* (pp. 734–743). Springer. https://doi.org/10.1007/3-540-70659-3_77.
- Popivanov, I., & Miller, R. J. (2002). Similarity search over time-series data using wavelets. In *Proceedings 18th international conference on data engineering* (pp. 212–221). <https://doi.org/10.1109/ICDE.2002.994711>
- Prein, A. F., et al. (2017a). Increased rainfall volume from future convective storms in the US. *Nature Climate Change*, 7, 880–884.
- Prein, A. F., et al. (2017b). The future intensification of hourly precipitation extremes. *Nature Climate Change*, 7, 48–52.
- Rabbani, M., & Oo, A. M. T. (2019). Electromagnetic effect of lightning return stroke to buried energy pipelines. *Energy Procedia*, 160, 467–474.
- Rakov, V. A. (2010). Lightning parameters for engineering applications (keynote speech). In *2010 Asia-Pacific international symposium on electromagnetic compatibility* (pp. 1120–1123). <https://doi.org/10.1109/APEMC.2010.5475697>
- Roberts, K. H. (1989). New challenges in organizational research: High reliability organizations. *Industrial Crisis Quarterly*, 3, 111–125.
- Romps, D. M., Seeley, J. T., Vollaro, D., & Molinari, J. (2014). Projected increase in lightning strikes in the United States due to global warming. *Science*, 346, 851–854.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65.
- Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26, 43–49.
- Shindo, T., & Suda, T. (2008). A study of lightning risk. *IEEE Transactions on Electrical and Electronic Engineering*, 3, 583–589.
- Smyth, P. (1996). Clustering sequences with hidden Markov models. In *Proceedings of the 9th international conference on neural information processing systems* (pp. 648–654). MIT Press.
- Steiner, M., Deierling, W., Ikeda, K., Nelson, E., & Bass, R. (2014). Airline and airport operations under lightning threats – Safety risks, impacts, uncertainties, and how to deal with them all. In *6th AIAA atmospheric and space environments conference*. American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/6.2014-2900>.
- Steiner, M., Deierling, W., & Johnson, D. (2012). *Lightning safety at airports. Material for thunder* (Vol. 5).
- Tarimer, I., Kuca, B., & Kisielewicz, T. (2012). A case study to risk assessment for protecting airports against lightning. *Elektron Elektrotech*, 117, 49–52.
- Earth Networks. (2021). *The ultimate lightning guide for airport operations*. Earth Networks. <https://www.earthnetworks.com/the-ultimate-lightning-guide-for-airport-operations/>.
- The World Bank Group. (2021). World development indicators. <https://databank.worldbank.org/reports.aspx?source=2&series=IS.AIR.PSGR&country=#>.
- Tovar, C., Aranguren, D., López, J., Inampúes, J., & Torres, H. (2014). Lightning risk assessment and thunderstorm warning systems. In *2014 international conference on lightning protection (ICLP)* (pp. 1870–1874). <https://doi.org/10.1109/ICLP.2014.6973434>
- Uman, M. A. (2001). *The lightning discharge*. Courier Corporation.
- Vidal, A., & Rakas, J. (2010). Impact of Lightning Strikes on National Airspace System (NAS) outages: A statistical approach. *4th International Conference on Research in Air Transportation (ICRAT)*. https://www.researchgate.net/publication/265413838_Impact_of_Lightning_Strikes_on_National_Airspace_System_NAS_Outages_A_Statistical_Approach.
- Vlachos, M., Kollios, G., & Gunopulos, D. (2002). Discovering similar multidimensional trajectories. In *Proceedings 18th international conference on data engineering* (pp. 673–684). <https://doi.org/10.1109/ICDE.2002.994784>
- Wang, B., et al. (2020). Understanding lightning strike induced damage mechanism of carbon fiber reinforced polymer composites: An experimental study. *Materials & Design*, 192.
- Wang, Y., & Zhupanska, O. I. (2015). Lightning strike thermal damage model for glass fiber reinforced polymer matrix composites and its application to wind turbine blades. *Composite Structures*, 132, 1182–1191.
- Warren Liao, T. (2005). Clustering of time series data—a survey. *Pattern Recognition*, 38, 1857–1874.
- World News. (2014). FAA investigates airport towers after Baltimore lightning strike. *The Guardian*. <https://www.theguardian.com/world/2014/feb/06/faa-investigates-airport-towers-baltimore-lightning>.
- Xiong, Y., & Yeung, D.-Y. (2002). *Mixtures of ARMA models for model-based time series clustering* (pp. 717–720).
- Yu, S., & Ren, Y. (2014). Research on the lightning risk assessment method for Chongqing based on fuzzy mathematics. In *2014 international conference on lightning protection (ICLP)* (pp. 1054–1057). <https://doi.org/10.1109/ICLP.2014.6973280>
- Zhang, H., Ho, T. B., Zhang, Y., & Lin, M.-S. (2006). Unsupervised feature extraction for time series clustering using orthogonal wavelet transform. *Informatica*, 30, 305–320.
- Zhang, X., Wu, J., Yang, X., Ou, H., & Lv, T. (2009). A novel pattern extraction method for time series classification. *Optimization and Engineering*, 10, 253–271.