Characterizing the Relation between Lightning and Wildfires in the Western United States

SCOTT D. RUDLOSKY,[a,b,d] JOSEPH PATTON,[b] DAILE ZHANG,[c] THANYAPORN NOIPLAB,[d] RUOMING JIN,[d] LENA HEUSCHER,[e] AND JASON A. OTKIN[e]

[a] *NOAA/NESDIS/GEO, Greenbelt, Maryland*
[b] *UMD/ESSIC/CISESS, College Park, Maryland*
[c] *University of North Dakota, Grand Forks, North Dakota*
[d] *Kent State University, Kent, Ohio*
[e] *UW-Madison/SSEC/CIMSS, Madison, Wisconsin*

(Manuscript received 4 March 2025, in final form 19 September 2025, accepted 21 October 2025)

ABSTRACT: Exploring lightning patterns alongside meteorological variables and fuel information helps diagnose the conditions under which lightning ignites fires. This study describes distributions of lightning, land surface, and meteorological conditions associated with known lightning ignitions in the western United States during 2020–22. Although most fire ignition lightning is classified as cloud-to-ground (89%), lightning misclassified as intracloud ignites 11% of fires. Findings indicate clear differences between the ignition and nonignition (null) lightning, and the fires detected on the day of versus those that holdover. Both the National Lightning Detection Network (NLDN) and Geostationary Lightning Mapper (GLM) observations help identify lightning more likely to ignite wildfires. On average, positive and negative polarity ignition flashes are 3.9 and 4.8 kA stronger than null flashes. Both GLMs indicate that ignition flashes are larger and 3–4 times brighter than the domainwide values. Many additional variables reveal a clear distinction between the ignition and null storm environments. The average hourly (daily) quantitative precipitation estimate is 7.08 (12.69) mm for the null flashes and only 2.76 (5.16) mm for ignitions. The Multi-Radar Multi-Sensor variables indicate weaker storms surrounding the ignitions, and several different relative humidity measures indicate ~10% drier environments. Each variable shown indicates that fires reported on the day of ignition occur in the most fire-prone environments, and null cases represent the least conducive environments, with holdover fires somewhere in between. Our analysis shows that distributions of lightning, land surface, and meteorological conditions can help identify environments most vulnerable to lightning ignitions.

SIGNIFICANCE STATEMENT: This study describes the development and application of a database of lightning known to have ignited fires in the western United States during 2020–22. Our findings indicate clear differences between the ignition and nonignition (null) lightning, and the fires detected on the day of versus those that holdover. Each of the variables shown indicates that fires detected on the day of ignition occur in the most fire-prone environments, the null cases represent the least conducive environments, and the fires that holdover fall somewhere in between. Our analysis indicates that distributions of lightning, land surface, and meteorological conditions can help identify environments most vulnerable to lightning ignitions. This knowledge will help guide the development of automated applications for identifying the most impactful lightning collocated with conditions conducive to wildfire ignition.

KEYWORDS: Lightning; Wildfires; Atmospheric electricity; Storm environments; Satellite observations; Forest fires

1. Introduction

Characterizing relationships between lightning and wildfires in the western United States helps identify regions most vulnerable to lightning ignitions. While human-caused fires are more frequent (84%), lightning-initiated wildfires account for 56% of the total acreage burned in the contiguous United States (Balch et al. 2017). Many of the lightning-ignited fires occur in less populated areas, resulting in both delayed detection and the need to bring in resources from further away. Schultz et al. (2024) reported that nearly 1/3 of lightning-ignited fires are first observed one or more days after the lightning ignites the fuels (i.e., holdover fires). Mapping potential lightning ignition locations may provide a means to identify smaller fires before they grow, mitigating the threat to life, property loss, and firefighting costs.

Exploring lightning patterns alongside other meteorological variables and fuel information helps better diagnose the conditions under which lightning ignites fires. This study first describes the development of a database of lightning known to have ignited fires in the western United States during 2020–22. Although many studies have developed automated procedures for linking lightning and wildfires, we chose to manually link the lightning and fires to develop as clean a training dataset as possible. We produce web-accessible tables that link to detailed lightning maps, fire spread animations, user input surveys, and lightning lists to help manually determine exactly which lightning stroke ignited each wildfire. These tools facilitate

Corresponding author: Scott D. Rudlosky, scott.rudlosky@noaa.gov

DOI: 10.1175/JAMC-D-25-0032.1

the development of a training dataset of lightning strokes known to have ignited wildfires.

Many different datasets help characterize the lightning ignition environments. Lightning characteristics are derived from both ground- and space-based sensors. Lightning point variables from the National Lightning Detection Network (NLDN) and Geostationary Lightning Mapper (GLM) are accompanied by GLM gridded product variables. Multi-Radar Multi-Sensor (MRMS) system data are used to characterize both the areal precipitation and storm environment. Near-storm environment variables are provided by the High-Resolution Rapid Refresh (HRRR) numerical weather prediction model. The land surface is characterized by the Landscape Fire and Resource Management Planning Tools (LANDFIRE) Project dataset (Long and Hatten 2023). These data also are combined to compute several fire weather and drought indices.

This study describes the distributions of lightning, land surface, and meteorological conditions associated with known lightning ignitions in the western United States during 2020–22. We explore lightning patterns alongside weather and fuel information to better understand conditions under which lightning ignites fires to help identify environments most vulnerable to lightning ignitions. These distributions aim to advise the development of automated applications for identifying the most impactful lightning collocated with conditions conducive to wildfire ignition.

2. Background

Taylor (1969) provided an early review of lightning effects on forests. They noted that individual thunderstorms can start tens of fires in minutes over miles of virtually inaccessible terrain. In August 1961, lightning started nearly 600 fires in one 3-day period in the National Forests in Montana, Idaho, Oregon, and Washington. This characteristic consequence of lightning, many fires in a short time, taxes fire suppression agencies beyond reasonable limits of manpower and equipment, and suppression costs soar (Taylor 1969). Their study also discussed latent (holdover) lightning fires that go undetected for several hours or days, smoldering in or surrounded by fuels too wet or otherwise unable to support flaming combustion. When fuel and weather conditions are right, the fire spreads, produces smoke and flames, and may be detected. Taylor (1969) estimated that 40% of all lightning-caused fires in the far western United States are undiscovered 12 h after estimated origin, and about 25% (10%) are still undiscovered 24 (72) h after origin.

Modern technology permits improved detection and characterization of wildfires. Over the last four decades, wildfires in the western United States have significantly increased in both size and severity after more than a century of fire suppression and exclusion (North et al. 2024). Approximately 15.5 million ha burned between 2019 and 2023, and Washington, Colorado, New Mexico, and California all experienced their largest ever recorded wildfires during that time (North et al. 2024). Comparing 1985–2022, the mean annual area burned by wildfire increased by 257% (1.89 million ha) and mean annual federal suppression costs increased by 332% [$2.01 billion U.S. dollars (USD); North et al. 2024]. Prestemon et al. (2013) reported that on lands managed by the Forest Service (Department of the Interior), the most common wildfire cause is lightning, representing 55.8% (37%) of wildfires and 74.3% (81.6%) of area burned between January 2000 and December 2008.

Ground-based lightning detection networks that precisely locate and characterize lightning are critical to understanding lightning ignitions. These networks provide the stroke type [i.e., cloud to ground (CG) or intracloud (IC)], polarity, and estimated peak current (I_p). The ground-based networks employ flash-grouping algorithms with certain temporal and spatial constraints to produce flash-level data. Previous studies have suggested that positive CG (+CG) flashes are most likely to ignite wildfires (Latham and Schlieter 1989; Dowdy and Mills 2009), but Schultz et al. (2019) found that 90% (613/681) of lightning flashes nearest to the fire starts were negative CG (−CG) and that 46% and 80% of the −CG and +CG ignition flashes were single stroke, respectively. MacNamara et al. (2020) found that 89% of wildfires in the western United States were initiated by −CG flashes, and 66% of those fire starts were due to single-stroke flashes. Schultz et al. (2024) reported that 92% (8%) of the matched lightning-ignited wildfire reports were −CG (+CG), and 54% (82%) of these were single stroke. Nearly 65% of the +CG ignition flashes were classified as holdover events (Schultz et al. 2024). Schultz et al. (2024) found nothing significantly different that could be used to separate a −CG fire starter from a nonfire starter based on peak amplitude alone but did find greater +CG peak current in ignition flashes.

Continuing current (CC) is thought to be the most critical lightning parameter for igniting wildfires (Latham and Williams 2001). McEachron and Hagenguth (1942) hypothesized that ignition is most likely to occur when a discharge contains, in addition to one or more fast return strokes, a long-continuing current phase in which current flows continuously in the lightning channel for a relatively long duration. While wide-baseline ground-based lightning detection networks cannot observe CC, the GLM can identify CC flashes (e.g., Bitzer 2017; Fairman and Bitzer 2022; Ding et al. 2024). Ding et al. (2024) noted limitations in the GLM's ability to detect the complete continuing current process and suggested that advancement in detection technologies and methodologies is needed to detect continuing current at large spatial scales. Pérez-Invernón et al. (2022) found that the lightning ignition efficiency (LIE) of flashes with CC lasting more than 10 ms is higher than the averaged LIE of CG flashes over CONUS. This indicates that models of lightning-induced wildfire risk could benefit from including information on flashes with CC observed from space (Pérez-Invernón et al. 2022). They also found that the LIE varies with the forest type and the ecological region, which suggests that other factors, such as wildland fuel characteristics and fire-related meteorological conditions, exert a substantial influence. Pérez-Invernón et al. (2022) found that both the holdover time and spatial distance between wildfires and GLM groups show remarkable similarities to the ground-based stroke and flash data.

Kalashnikov et al. (2023) reviewed lightning-ignited wildfires (LIWs) in the western United States during May–September (2015–20), focusing on the relative occurrence of

dry thunderstorms and the influence of the environment on holdover wildfires (i.e., those not detected until 2–5 days following ignition). The median holdover precipitation was more than double that of the promptly detected LIWs (5.1 vs 2.5 mm; $P < 0.10$). Of the 3726 identified LIWs, 3157 (~84.7%) were promptly detected, while 569 (~15.3%) were holdovers (Kalashnikov et al. 2023). The desert and semidesert environments of the Great Basin and interior Southwest have the largest proportion of promptly detected LIWs (>90%). Conversely, the highest proportion of holdovers (~20%–30%) is found in the largely mountainous, forested terrain of the Arizona/New Mexico mountains, the middle and southern Rockies, and the Nevada/Utah mountains. Kalashnikov et al. (2023) also found that most LIWs are smaller and associated with higher ignition precipitation (2.8–3.0 mm), whereas the largest LIWs occur with lower ignition precipitation (~2.5 mm), likely reflecting increased flammability due to less precipitation. Vapor pressure deficit (VPD; from −9.3 to −3.8 hPa) and maximum temperatures (Tmax; from −5.6° to −2.0°C) are significantly lower for holdovers compared to promptly detected LIWs across all ecoprovinces, and 100-h fuel moisture (FM100; from +0.4% to +2.6%) is significantly higher for holdovers. Kalashnikov et al. (2023) conclude by suggesting that spatially varying, vegetation-specific precipitation thresholds would more accurately characterize the risk of LIW ignition and holdover potential in different ecoprovinces of the western United States.

3. Data and methods

a. *Identifying and characterizing ignition strokes and flashes*

Automated identification of lightning that ignites wildfires leaves more uncertainty than desired (i.e., requires too many assumptions), so tools are developed to bring human observers into the loop when determining which lightning ignited each wildfire. Automated identification assumes the accuracy of the initial fire report locations, requires firm space and time criteria, and precludes using the resulting dataset to train artificial intelligence models. To address these limitations, we developed a semiautomated, human-in-the-loop methodology that leverages a more accurate, but less comprehensive, fire perimeter dataset to pinpoint lightning ignition events. The goal of this process is to create a robust, high-quality dataset that can be used for future research, including training artificial intelligence (AI) models, by transparently addressing the biases inherent in human-driven analysis.

Rather than linking lightning with initial wildfire report locations, wildfire perimeters form the basis of our analysis. Although initial report locations include many more fires, the perimeter dataset provides the most accurate permanent record (i.e., both location and fire type classification). The Wildland Fire Interagency Geospatial Service group provides a national historical fire perimeter data layer of conglomerated agency perimeters [The Wildland Fire Interagency Geospatial Service (WFIGS) 2024]. This dataset has evolved over the years, and the version used in this study included perimeters for the 2020–23 fire seasons. The most recent version includes fires from 2021 to the present. Requirements for fire perimeter inclusion are set by the contributing agencies (e.g., minimum acreage requirements). To limit the project scope, the present study includes wildfire perimeters from the western U.S. states of Washington, Oregon, California, Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, and New Mexico.

The core of our methodology is the manual determination of the ignition stroke by a human observer (i.e., the study authors). This process is guided by a suite of interactive visualization tools to ensure a systematic and repeatable analysis. The primary biases introduced by this approach are the reliance on human judgment to visually interpret complex spatiotemporal data and the potential for subjective interpretation of ambiguous cases. Our process mitigates these biases by standardizing the analysis with a precise, multistep workflow. Figure 1 illustrates our analysis steps and their relationships, and Fig. 2 provides an example of the detailed lightning maps and fire spread animations.

Detailed lightning maps, fire spread animations, and lightning lists help determine exactly which lightning ignited each wildfire (Fig. 2). We first search broadly for all lightning and fire hotspot detections within 0.5° of the final fire perimeter polygons to create the lightning and fire maps for each incident. Lightning must occur within 10 days before the fire begins through the day after ignition. Schultz et al. (2019) found that the longest duration holdover event was 12 days, indicating we may have missed a small number of matches (~1.5%). Fifty-four fires holdover for 9 or 10 days, so we may miss up to 50 fires at 11 or 12 days, or ~1.5% of all ignitions. Fire hotspots are plotted from 10 days prior to ignition through the date of fire suppression.

Lightning maps (Fig. 2a) are created using the Plotly Graph Objects Python package. These interactive Hypertext Markup Language (HTML) maps are customized to enhance their usability. The lightning maps display data from the NLDN, Earth Networks Total Lightning Network (ENTLN), and the GOES-East and GOES-West GLMs. GLM flashes are displayed as circles scaled by their area at their radiance-weighted centroid (locations include parallax offset). Note that for proximity analyses, GLM flashes are represented as polygons that characterize the full spatial footprint covered by the constituent GLM groups. All lightning symbols are color coded by time to fire start (darker shades are nearer 0000 UTC on the official start date). The lightning maps also display the initial fire hotspot detections from each satellite sensor.

Fire spread animations (Fig. 2b) are created using the Plotly Express Scatter Mapbox Python tool, permitting play, pause, zoom, and pan functionality. The maps display fire hotspot detections from both geostationary and low-Earth-orbiting sensors to help determine where the fire originated. Fire hotspots from both the East and West Geostationary Operational Environmental Satellite (GOES) Advanced Baseline Imager (ABI) are combined into 15-min bins with hotspot detections from the Moderate Resolution Imaging Spectroradiometer (MODIS) and Visible Infrared Imaging Radiometer Suite (VIIRS) sensors. The GOES ABI Fire Detection and Characterization (FDC) product is based on the heritage

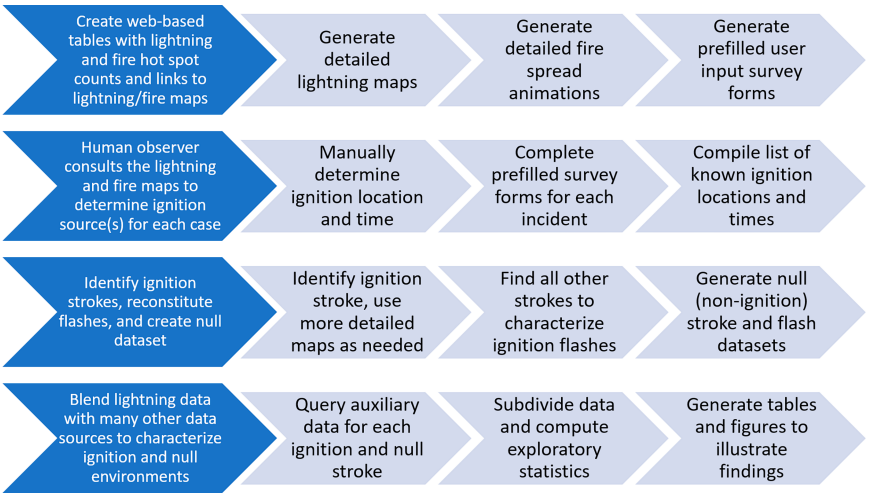

FIG. 1. Diagram illustrating the analysis steps and their relationships.

Wildfire Automated Biomass Burning Algorithm (Schmidt et al. 2012). The MODIS and VIIRS hotspots were obtained from the National Aeronautics and Space Administration Fire Information for Resource Management System (FIRMS). Symbol color and size both correspond to the estimated fire radiative power [megawatts (MW)].

The observer meticulously reviews the lightning maps and fire spread animations to determine where the fire originated. The timeliness and accuracy of the fire, lightning, and hotspot observations introduce variability that the human observer attempts to address. For natural (unknown) fires with hotspot detections, lightning strokes are linked to the fires if they are within ~0.01° (~0.005°) of the initial hotspot locations (i.e., ~1 km and ~500 m). Smaller fires sometimes have no hotspot detections; in these cases, lightning strokes are linked to the fires if they are within ~0.01° (~0.005°) of the final fire perimeter. The acceptable offset distance for fires with only GEO hotspots is closer to 0.1° (~10 km). The acceptable offset distance also increases to ~0.05° (~5 km) for natural fires with only IC strokes nearby (especially weaker IC pulses, which are typically more challenging for the NLDN to geolocate). For unknown fires, weaker peak current IC pulses are less likely to be linked to fires regardless of distance. Figure 2 illustrates a case where lightning ignited two fires within the same fire perimeter. Our methodology allows for the inclusion of all such ignition flashes in the final database.

Individualized prefilled survey forms for each incident require minimal input from the human making the determination (i.e., study authors). Once the lightning stroke that ignited the fire is identified, the user selects the time and day of the fire ignition (1-min accuracy) and indicates which networks observed the lightning. The user also indicates the potential for multiple start locations and adds any notes of interest. The user clicks submit and then manually copies and pastes the results of the HTML survey form onto a new line in an annual spreadsheet. For each incident (spreadsheet entry), we next query the NLDN stroke/pulse archive for only the minute of ignition and a closer proximity to the fire ignition (0.1°).

The results of this final query are manually trimmed to only the strokes that ignited the fires (i.e., the stroke/pulse nearest the hotspot or fire perimeter). To aid our final determination, new lightning maps are created for each incident with only the potential ignition strokes/pulses, the fire perimeter, and the initial hotspot locations. We characterize this full list of lightning ignition strokes and also trim the list down to remove duplicate strokes (i.e., remove subsequent strokes that strike the same ignition location). Although these subsequent strokes are not considered ignition strokes, their characteristics are captured in the flash-level NLDN variables. The trimmed list of fire ignition strokes/pulses forms the basis of our dataset development, with many other fields calculated at the locations and times of these ignitions. Examining hundreds of individual cases revealed that the NLDN provided the most accurate lightning locations relative to the initial hotspots observed from low-Earth orbit. Although ENTLN observations are included in our analysis, only NLDN stroke locations are used to designate fire ignition locations. The ignition strokes most often fall within the fire footprint, with strokes outside the footprint relocated to the nearest fire perimeter boundary.

b. *Characterizing the ignition environment*

We first produce both stroke- and flash-level NLDN variables, along with GLM point and gridded product characteristics. Stroke-level variables include type (IC and CG), polarity (positive or negative), and I_p (kA). NLDN flash-level variables also include flash type (negative, positive, and bipolar), maximum/minimum I_p, and the duration. The GLM flash-level characteristics include flash area (km^2), flash energy (fJ), maximum consecutive frames, maximum number of events per group, maximum group area (km^2), maximum group energy (fJ), duration (s), and distance between flash and fire polygons. We also calculate the GLM flash extent density (FED), minimum flash area (MFA), and total optical energy (TOE) for the grid cell nearest the ignitions (i.e., nearest within a 1° × 1° box surrounding the ignition location). We

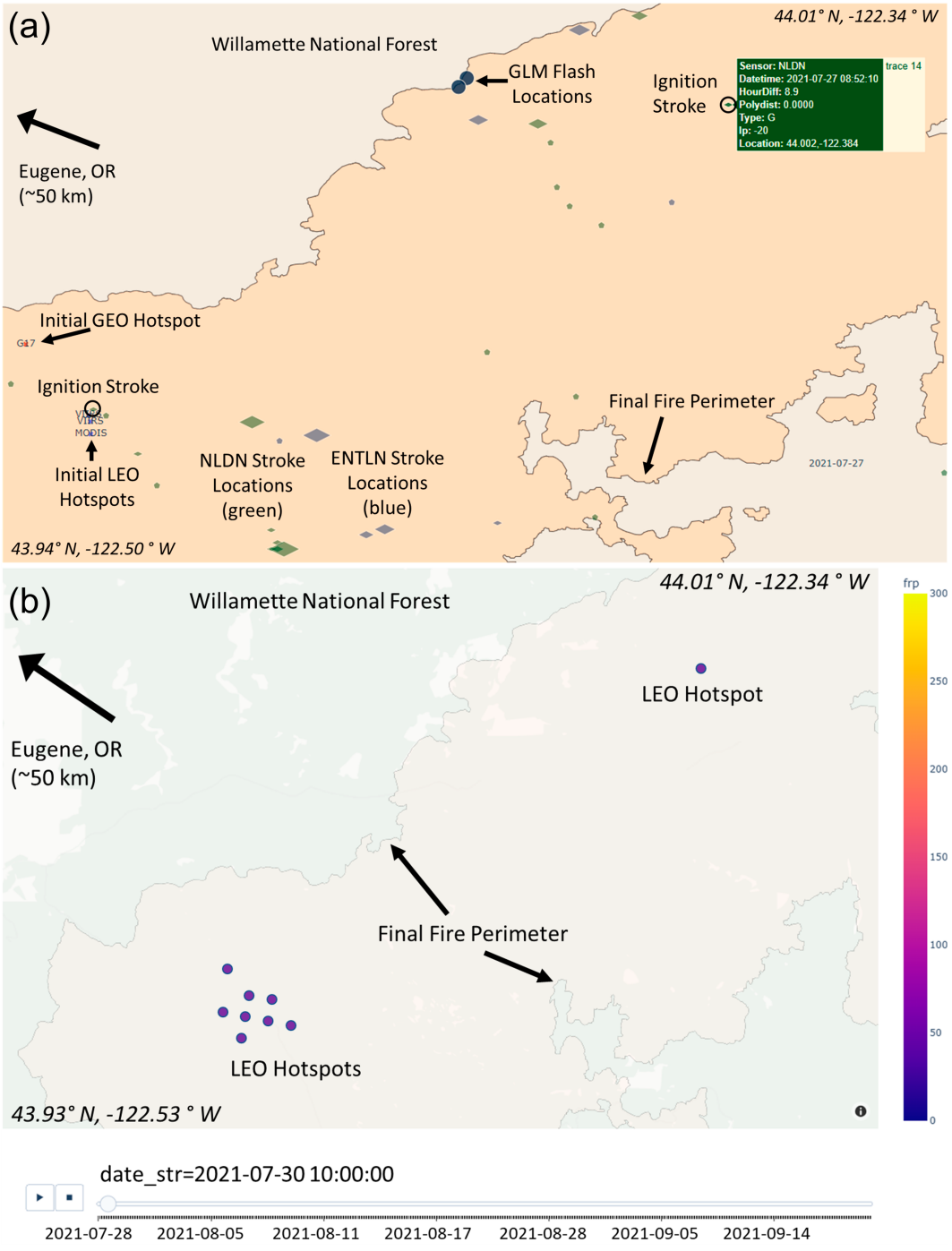

FIG. 2. (a) Examples of the detailed lightning maps and (b) fire spread animations used to help manually determine which lightning strokes ignited each fire. In (a), diamonds indicate CG strokes observed by the NLDN (green) and ENTLN (blue) and hexagons indicate IC pulses, with the size of both scaled by the estimated peak current. (a) The initial GOES, MODIS, and VIIRS hotspot detections, and black circles indicate two separate lightning ignition locations. (b) A screenshot of a movie that illustrates the locations and temporal evolution of all hotspots detected by geostationary and low-Earth-orbiting satellites (i.e., at 15-min intervals), with size and color corresponding to the fire radiative power (MW). This is the Gales Fire in the Middle Fork Complex, which ignited during July 2021.

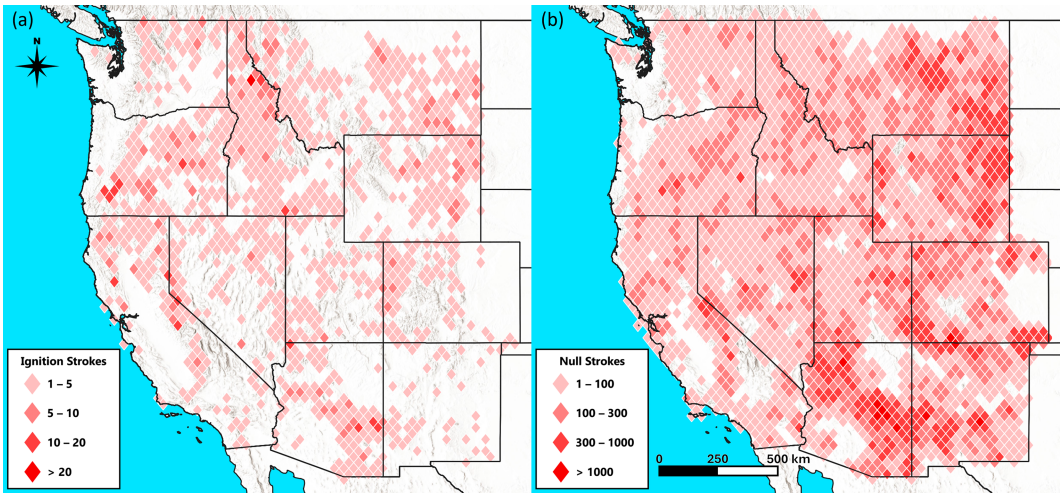

FIG. 3. Spatial distribution of (a) confirmed lightning ignitions and (b) null (nonignition) strokes during 2020–22. The top-left (bottom-right) corner of these frames is at 49.5°N/128°W (31°N/102°W).

compute averages, maxima, and minima for each of these gridded variables to characterize the GLM observations within the surrounding 1° × 1° box (like many other spatial fields described below).

In addition to the ignition locations, we gathered 100 times as many nonignition (null) stroke locations. The null strokes are all selected from the original candidate strokes (i.e., those within the broad initial time and space criteria of any fire perimeter) to ensure they sampled fire environments. The null strokes were trimmed to be at least 0.1° outside any fire perimeter and outside the ignition times of any fires to ensure they did not ignite fires. Of this subset between 0.1° and 0.5°, we retained as nulls all CG strokes and every 35th IC pulse to more closely match the CG/IC ratio of the ignition dataset. Although the ignition and null datasets only include first strokes, their flash-level variables characterize the subsequent strokes. Figure 3 illustrates the spatial distribution of all ignition and null stroke locations. Ignition stroke locations spread broadly across the western United States, with null strokes sampling an even larger portion of the study domain.

The locations, dates, and times of the lightning ignition and null strokes are used to query several databases and compute many variables. We gathered information on precipitation from the blended MRMS radar/gauge dataset, used HRRR data to characterize the near-storm environment, and gathered additional MRMS data to characterize the storms surrounding the ignitions. Various Python packages were employed to compute many meteorological and land surface variables (e.g., *xarray*, *geopandas*, *rioxarray*, and *rasterio*).

To characterize the areal precipitation, we gathered the blended MRMS radar/gauge quantitative precipitation estimate (QPE) product when available (radar only otherwise; <1% of cases) from the Iowa State University archive (http://mtarchive.geol.iastate.edu). Zhang et al. (2016) described challenges facing radar QPE in complex terrain, noting that the MRMS radar QPE still suffers from blockages in complex terrain and may underestimate orographically enhanced precipitation. Osborne et al. (2023) summarized the radar QPE limitations in the western United States and described efforts to develop a statistically based, deep learning approach to address these challenges. The present study uses the MRMS data as provided while noting these limitations. For each ignition and null stroke, we identify the nearest 1-h QPE at the top of the hour and the 24-h QPE 12 h later. We also cut out a 1° × 1° box centered on the ignition from the 1- and 24-h raster images to characterize the broader precipitation distribution. Resulting parameters include the 1- and 24-h local QPE, nearest QPE, nearest QPE distance, maximum QPE, maximum QPE distance, mean QPE, median QPE, and QPE coverage (i.e., fraction of 1° × 1° box with nonzero values).

Archived MRMS data also were used to characterize the types of storms surrounding the fire ignitions (similar to QPE analysis). We computed statistics to describe the spatial distribution of echo-top height (18 dBZ), composite reflectivity, and vertically integrated liquid (VIL). For each of these fields, we computed the nearest value, nearest value distance, maximum value, maximum value distance, mean value, median value, and value coverage (fraction of 1° × 1° box with nonzero values).

We gathered HRRR analysis data to characterize the near-storm environment. We computed local values and regional means at ignition time and tracked how some variables changed over time (e.g., 24-h maximum/minimum, 12-h delta). In addition to temperature, humidity, and wind variables, the HRRR data also provide soil moisture estimates and contribute to many of the fire weather parameters described below. Researchers have documented biases in HRRR wind speed and soil moisture parameters, specifically in mountainous and arid regions. HRRR wind speed forecasts have an overall tendency to overpredict at lower observed wind speeds and underpredict at higher wind speeds (Seto et al. 2025). Marinescu et al. (2024) found that the HRRR soil moisture estimates produce wetter soils in dry regions and drier soils in wet regions relative to in situ observations. These soil moisture differences occur at most

soil depths but are larger at deeper depths below the surface (100 cm).

These datasets were combined to compute several fire weather indices that indicate the potential for fire ignitions and/or the intensity of existing fire activity. Indices typically combine basic meteorological data to create a weighted index that considers the atmospheric conditions conducive to fire behavior. We compute the Fosberg fire weather index (includes both fuel moisture and rate of spread component; Fosberg 1978), Munger index (indicates number of last daily rainfall greater than 12.6 mm; Munger 1916), Sharples index (assesses moisture content and fire danger in eucalypt forests in a simple way; Sharples et al. 2009a,b), Canadian fire weather index (FWI; composed of weather data only; consists of three primary fuel moisture indices, two intermediate fire behavior indices, and a final fire intensity index; Van Wagner and Pickett 1985; Van Wagner 1987), Angstrom index (simple index utilizing temperature and humidity; Chandler et al. 1983), Nesterov index (daily fire danger rating index requiring daily air temperature, dewpoint temperature, and precipitation data; Nesterov 1949), and Baumgartner index (based on the concept that fire danger is mainly conditioned by fuel dryness; Baumgartner et al. 1967).

We obtained land surface and drought variables from a number of sources. Several fire-related variables were obtained from gridded surface meteorological (gridMET) (Abatzoglou 2013), including burning index, energy release component, and the 1000- and 100-h fuel moisture. LANDFIRE (Long and Hatten 2023) data are used to obtain the existing vegetation cover (crops, developed, vegetation, and water/nonflammable) and the 13 Anderson fire behavior fuel classifications. Fuel classifications were grouped into five different classes: grass dominated, chaparral/shrub, timber litter, slash/logging, and nonflammable/urban. For each strike location, percentages of each vegetation cover and fuel classification within 250 m are returned. Drought variables were obtained from an internal University of Wisconsin–Madison (UW) Cooperative Institute for Meteorological Satellite Studies (CIMSS) archive. Variables include the evaporative drought demand index (EDDI; Hobbins et al. 2016) at 7-, 14-, 56-, and 84-day intervals, the evaporative stress index (ESI; Anderson et al. 2007) at 7-, 56-, and 84-day intervals, and the standardized precipitation index (SPI; McKee et al. 1993) at 30- and 60-day intervals. These data were regridded to the HRRR resolution and are available at a daily temporal resolution.

c. *Dataset analysis*

We compute monthly counts of lightning and wildfires (i.e., only natural and unknown causes) in the western United States during 2020–22. The spatial and temporal domains were selected to manage the project scope. The fire counts only reflect fire perimeters that met inclusion criteria for the annual perimeter dataset (i.e., set by the contributing agencies) and likely omit an unknown number of fires that ignited, required no response, and were never mapped or entered into the fire perimeter archive. The WFIGS fire perimeter dataset has evolved over the years (WFIGS 2024) and underwent a major overhaul in 2023. The version used in this study included perimeters for the 2020–23 fire seasons, while the most recent version includes fires from 2021 to the present. This transition appears to have resulted in many fewer of the 2022 fires being designated as natural (rather than unknown) than in 2020–21. Since our manual analysis techniques required more certainty for changing the fire cause from unknown to natural, the omission of more marginal cases resulted in fewer matches during 2022. This should have a minimal impact on the study conclusions because the null dataset excludes lightning near both the natural and unknown fire perimeters.

We summarize the distributions of strokes and flashes observed by the ground-based NLDN. Known misclassification issues (e.g., Rudlosky and Fuelberg 2010; Mallick et al. 2015) were confirmed during manual inspection of the lightning ignitions, with lightning reported as IC confirmed to have ignited fires. In many of these cases, the ENTLN characterized the flash as CG and/or of the opposite polarity. The reverse condition also was observed, with NLDN-reported CG strokes reported as IC pulses by the ENTLN (~20% of the time; see section 4a). We chose to analyze the data as reported by the NLDN because it is the best characterized ground-based network (i.e., more than 5 times more references than the ENTLN; per Google Scholar). The ignition strokes/pulses were trimmed to remove subsequent strokes in the same location, and these ignition strokes were then joined back to the full NLDN database to reconstruct the ignition flashes. We used relatively broad space (0.15°) and time (1 s) criteria for reconstructing the ignition flashes.

GLM flash polygons (i.e., representing all constituent GLM groups) were matched to the NLDN-reported ignition locations within 0.5° and 0.5 s. The GLM boasts broad coverage but reduced performance at steeper viewing angles. Some limitations are anticipated because our study region lies outside the optimal GLM coverage area. Matched GLM flash characteristics are compared with annual average values and previously reported values (rather than matched null strokes). Manual inspection revealed that in nearly all cases with no direct GLM match, there are GLM flashes observed in close proximity. Despite missing ignition flashes, the types of lightning surrounding the ignitions can be suggestive of increased risk of fire ignitions. This motivates the analysis of the gridded GLM products. We compute averages, maxima, and minima of the GLM gridded variables to characterize the GLM observations nearest the fire ignition and within the surrounding $1° \times 1°$ scene.

We employ simple statistics to differentiate the ignition and null samples. We first use a Kolmogorov–Smirnov (KS) test (Smirnov 1939; Massey 1952) to determine which variables differ most between the ignition and null datasets. The resulting p values are very low for most variables (i.e., much less than 10^{-10}), so the discussion focuses on the distributions of the KS test statistic (larger values indicate more difference). We then examine the variables appearing most different within each of the variable groups. Summary tables are prepared for variables indicated as most different. Lightning type and polarity distributions are examined for fire ignition strokes/flashes and null strokes/flashes along with the distributions of select environmental variables. We also separate fires

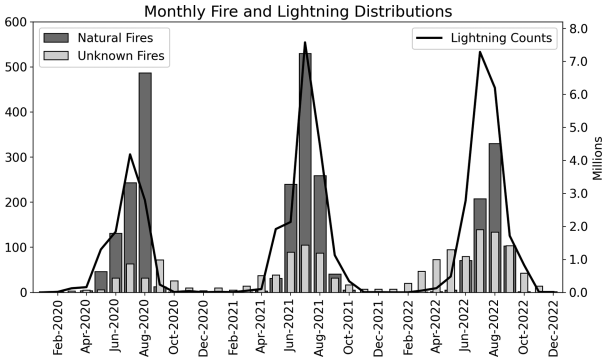

FIG. 4. Monthly counts of wildfire ignitions from the WFIGS national historical fire perimeter dataset and lightning strokes observed by the ground-based NLDN within the study domain.

that dwell for one or more days before being reported (i.e., holdover fires) from those reported on the day of ignition. Our fire perimeter dataset only reports the ignition date (i.e., not time), so a lag time of <1 h could be counted as a 1-day delay if the fire ignited just before 0000 UTC and was reported shortly after. Distributions of many different variables are shown for ignition flashes, those detected on the day of ignition, holdovers, and null (nonignition) flashes.

4. Results and discussion

a. *Lightning distributions*

The annual lightning and wildfire climatology follow a very similar pattern. Figure 4 illustrates the monthly counts of lightning and wildfires in the western United States during 2020–22 (i.e., includes only natural and unknown causes). Lightning counts peak during July each year with more than 7 million strokes/pulses during July 2021 and 2022 (over 4 million in July 2020). Lightning counts also exceeded 4 million strokes per month during August 2021 (4.4 million) and August 2022 (6.2 million). The wildfire counts peak during August each year with 519, 346, and 464 during 2020, 2021, and 2022, respectively. These fire counts only reflect fire perimeters that met the inclusion criteria for the annual perimeter dataset. Fires are undercounted because many fires ignited, required no response, and were never mapped and entered into the fire perimeter archive.

TABLE 1. Wildfire distributions from the WFIGS national historical fire perimeter data layer. The total column represents all fires labeled as natural or unknown/undetermined, the natural column represents all fires labeled as natural, and the matched columns represent the fire perimeters where lightning was confirmed to have ignited the fire.

	Total	Natural	Matched	Matched (%)
2020	1185	917	929	78.4
2021	1559	1122	1110	71.2
2022	1481	133	722	48.8
2020–22	4225	2172	2761	65.3

During the study years, 65.3% of all fires initially characterized as natural or unknown are confirmed to be natural (Table 1). The matched fraction is greatest during 2020 (78.4%) and lowest during 2022 (48.8%). The WFIGS fire perimeter dataset overhaul in 2023 resulted in many fewer of the 2022 fires being designated as natural (rather than unknown) than in 2020–21. Since our manual analysis techniques required more certainty for changing the fire cause from unknown to natural, omission of more marginal cases resulted in fewer matches during 2022. Seasonally, this matched fraction peaks during July (76%) and August (80.1%) when the likelihood of natural ignition is greatest (not shown).

Table 2 summarizes the distributions of strokes and flashes observed by the ground-based NLDN. Overall, 4400 strokes/pulses are linked to lightning ignition locations. Of the 3911 CG strokes (88.9%), only 216 are +CG (5.5%). Alternatively, of the 488 IC pulses (11.1%), 301 (61.7%) are characterized as +IC. Manual inspection of the lightning ignitions confirms known misclassification issues, with lightning reported as IC shown to have ignited fires. The ENTLN characterizes some of these flashes as CG and/or of the opposite polarity (not shown). Analysis reveals that 17.1% of ignition flashes with only IC components are classified as CG by the ENTLN and that 20.5% (27.8%) of the single-stroke CG (IC) ignitions are classified as IC (CG) by the ENTLN (not shown). We choose to analyze the data as reported by the NLDN.

The ignition strokes/pulses are trimmed to remove subsequent strokes in the same location, resulting in 2761 unique ignition locations (Table 2). These ignition strokes are then joined back to the full NLDN database to reconstruct the ignition flashes. The fraction of single-stroke flashes (21.4%) is lower than the 46% value reported by Schultz et al. (2019). This likely reflects our inclusion of IC pulses and the relatively broad space and time criteria used for reconstructing the ignition flashes (i.e., 0.15° and 1 s). The proportion of CG (88.5%) and IC (11.5%) flashes closely resembles the stroke distribution, although the +CG flash percentage is double the +CG stroke/pulse percentage (10.9% vs 5.5%). Another 24% of the flashes have some +CG or +IC component (i.e., bipolar flashes), while 65.1% are purely negative.

The flash polarity distributions provide additional insights (Table 3). On average, negative ignition flashes consist of 3.5 strokes/pulses, with 2.6 CG strokes and 1 IC pulse. The negative ignition flashes last 0.163 s on average with an estimated peak current of −26.2 kA. Positive ignition flashes average 2.1 strokes/pulses with more IC (1.60) than CG (0.5) components. The average positive flash estimated peak current is 28.9 kA, and the average duration is 0.059 s. Bipolar ignition flashes have the greatest average stroke count (5.9), estimated peak current (−30.9 kA), and duration (0.294 s).

The GLM distributions provide an alternate perspective on the lightning that ignites wildfires. The GLM boasts very broad coverage but reduced performance at steeper viewing angles. The GOES-East and GOES-West GLMs detect 37.8% and 51.9% of the ignition flashes, respectively (Table 4). The GOES-West GLM instrument performance and point of view help explain the greater detection efficiency relative to the GOES-East GLM (Rudlosky and Virts 2021). The matched fraction varies

TABLE 2. Distributions of fire ignition strokes and flashes observed by the ground-based NLDN.

	Strokes	CG strokes	−CG strokes	+CG strokes	IC pulses	−IC pulses	+IC pulses
2020–22	4400	3911 (88.9%)	3696 (94.5%)	216 (5.5%)	488 (11.1%)	187 (38.3%)	301 (61.7%)
	Flashes	CG flashes	IC flashes	Negative flashes	Positive flashes	Bipolar flashes	Single stroke
2020–22	2761	2443 (88.5%)	318 (11.5%)	1798 (65.1%)	300 (10.9%)	663 (24.0%)	591 (21.4%)

year to year with the largest fraction during 2022 (42.2% and 58.0%). Manual inspection reveals that in nearly all cases with no direct GLM match, there are GLM flashes observed in close proximity. Despite missing ignition flashes, the types of lightning surrounding the ignitions can be suggestive of increased risk of fire ignitions.

Table 5 summarizes the GLM flashes that match NLDN-confirmed ignitions. The ignition flashes observed by the GOES-West GLM are larger but less energetic than the GOES-East GLM flashes. Both the GOES-West GLM flash area and maximum group area are greater than the GOES-East GLM. Observations suggest that the ignition flashes are larger and brighter than the average GLM flash. The average ignition flash areas are near the 99th percentile of domainwide values reported by Rudlosky and Virts (2021). Both GLMs indicate that the ignition flashes are larger than the domainwide values, with greater flash area and maximum group area (mga) values indicative of stronger CG flashes (Koshak 2010; Rudlosky 2015; Ringhausen et al. 2021). The flash energy and maximum group energy reveal the greatest difference between the GLM ignition flashes and the domainwide sample. Both GOES-East and GOES-West suggest that ignition flashes are 3–4 times brighter than the domainwide sample.

The other GLM level 2 (L2) variables provide more mixed signals. The GOES-East and GOES-West ignition flashes have nearly identical average durations (0.338 vs 0.335 s; Table 5), both exceeding the domainwide averages (0.270 and 0.232). Ignition flashes observed by the GOES-West GLM have more consecutive frames (maxcon) and a maximum number of events per group (mneg) on average than both the GOES-East GLM ignitions and the domainwide averages. The GOES-West maxcon and mneg values are nearly double the domainwide averages and may be indicative of continuing current flashes. The GOES-East GLM patterns are less clear, likely relating to instrument performance and viewing angle. Our findings indicate the GLM observations can help identify lightning more likely to ignite wildfires.

b. *Characterizing the ignition environments*

Both the satellite- and ground-based lightning observations suggest that ignition flashes differ from average lightning.

TABLE 3. Lightning flash type and polarity distributions of fire ignition flashes observed by the ground-based NLDN. The table reports the average values for all ignition flashes during 2020–22.

	Stroke count	CG count	IC count	Max Ip	Min Ip	Duration
Bipolar	5.9	2.8	3.1	14.5	−30.9	0.294
Negative	3.5	2.6	1.0		−26.2	0.163
Positive	2.1	0.5	1.6	28.9		0.059

Observations suggest that the ignition flashes are larger, brighter, and last longer than the average GLM flash. Table 6 indicates similar proportions of CG (~88%) and IC (~12%) flashes between ignition and null flashes (by design), similar average durations, and comparable polarity distributions. Ignition flashes are more likely to be negative (65.1% vs 61.5%), while null flashes are more likely to be bipolar (i.e., contain both positive and negative strokes/pulses; 15.6% vs 10.9%). Both categories average ~4 strokes/pulses per flash, with ~2.5 CG strokes and ~1.5 IC pulses. The estimated peak current presents the greatest difference between the ignition and null flashes. On average, positive and negative polarity ignition flashes are 3.87 and 4.75 kA stronger than null flashes.

We next use a KS test to determine which variables differ most between the ignition and null datasets. Table 7 presents the KS statistics for select variables from each of the variable groups (see Table 8 for variable descriptions). Each of the variable groups shows differences between the ignition and null samples. The 1- and 24-h local/nearest QPEs differ most between samples, with the 24-h values indicating the greatest difference (i.e., largest KS statistic, 0.328 and 0.325). The fifth most different QPE variable is the mean QPE for all nonzero QPE in the surrounding 1° × 1° scene (0.244). This suggests that characterizing nearby storms also can provide important insights. Of the model-derived variables, the KS statistics are greatest for the average 12-h, maximum 24-h, and minimum 12-h relative humidity (RH). The regional mean (RM) equilibrium height (elhRM) and convective available potential energy (capeRM) also differ between the ignition and null samples, again suggesting the importance of considering the surrounding environment.

Fire weather indices, MRMS variables, and lightning characteristics reveal additional differences between the ignition and null datasets. The Nesterov and Munger fire weather indices seem most indicative of differences between the ignition and null datasets. The fuel moisture for 100-h time lag (FM100) and the FWI also indicate differences. Of the MRMS variables, the nearest VIL, reflectivity (reflqc), and 18-dBZ echo-top height (echotop18) all have KS statistics > 0.2. Both the maximum echotop18 and VIL in the surrounding 1° × 1° scene differ between the ignition and null samples. As suggested earlier, the estimated peak current I_p differs most between the ignition and null samples. GLM variables with relatively large KS values include the regional mean and regional maximum flash extent density along with the minimum flash area of the nearest GLM observations.

The environmental variables and storm characteristics reveal a clear distinction between the ignition and null storm environments (Table 9). The average hourly (daily) QPE is 7.08

TABLE 4. The number and percentage of ignition flashes observed by the GOES-East (G16) and GOES-West (G17) GLMs.

Year	Total flashes	G16 matches	G16 (%)	G17 matches	G17 (%)
2020	929	337	36.3	468	50.4
2021	1110	403	36.3	547	49.3
2022	722	305	42.2	419	58.0
2020–22	2761	1045	37.8	1434	51.9

(12.69) mm for the null flashes and only 2.76 mm (5.16 mm) for ignitions. The regional mean QPE is 2.35 mm less for ignitions with less of the surrounding 1° × 1° scene observing nonzero QPE24 (62% vs 75%). The MRMS variables indicate that the storms surrounding the ignitions are weaker than storms associated with null flashes. The nearest (regional maximum) reflectivity is 6.73 (3.15) dBZ lower for the ignitions than for the nulls, and the nearest (regional maximum) echotop18 is 1.5 (1.08) km lower. The nearest (regional maximum) VIL is 2.29 (4.87) kg m^{-2} less for the ignitions than for the nulls. Each of the aforementioned comparisons was statistically significant (i.e., $P < 0.1$). Moisture variables dominate the selection of model-derived fields. Several different measures of RH indicate ~10% drier environments during ignitions. Assuming similar temperature profiles, less moisture results in lower equilibrium heights and less CAPE for the ignition environments, which may help explain the tendency for weaker storms surrounding the ignitions. The Nesterov ignition, Munger, and FWI are all much larger for the ignitions than for the nulls. The FM100 indicates drier fuels, and the buildup index (BI) indicates more available fuel for combustion for the ignition cases than for the nulls.

Ignition samples are subdivided further into those observed on the day of lightning ignition and holdover fires (i.e., those that dwell for 1+ days). Fifty-six percent of ignitions are reported that calendar day, while 44% are holdovers (not shown). Figure 5 clearly indicates differences between the ignition and null samples and also highlights important differences for holdover fires. Ignitions reported on the day of ignition differ most from the null samples for each of the variables shown. Day of detections has ~33% of the 1- and 24-h precipitation and 15% less RH than the null cases. Alternatively, the 1-h (24-h) QPE is 1.26 (2.51) mm greater for the holdover fires versus those detected on the day of ignition, and the RH is ~7% greater for the holdovers (Fig. 5).

Many lightning variables follow this same pattern, only with less pronounced differences between the day of and holdover cases than the environmental variables (Fig. 5 and Table 10). Null environments are characterized by more frequent smaller GLM flashes, while the day of detections has the least frequent and largest GLM flashes (Fig. 5). The NLDN indicates the strongest average I_p for day of ignitions and the weakest average I_p for null samples. Table 10 shows greater flash extent density and minimum flash area values for the ignition flashes than for the nulls. The NLDN indicates the day of detections is marginally more energetic with slightly longer durations than the holdover ignitions. The I_p shows the greatest difference between the ignition and null datasets.

Our findings indicate clear differences between both the ignitions and the nulls, and the fires detected on the day of versus holdover fires. Each of the variables indicates that fires detected on the day of ignition occur in the most fire-prone environments, and null cases represent the least conducive environments, with holdover fires somewhere in between. This suggests that the holdover fires will continue to be the most difficult to identify and should continue to be a focus of future research. Our analysis suggests that distributions of lightning, land surface, and meteorological conditions can be used to help identify locations most vulnerable to lightning ignitions. This knowledge will help guide the development of automated applications for identifying the most impactful lightning collocated with conditions conducive to wildfire ignition.

c. *Limitations*

The wildfire perimeter dataset is both a limitation and a strength. The perimeter dataset contains fewer fires than the daily new start reports. This decreases the sample size but increases certainty in our fire ignition determinations. Manual

TABLE 5. GLM L2 statistics for ignition flashes (top) and all GLM flashes within our study domain (bottom) during 2020–22. The "All" category includes ~14.3 million and ~17.3 million G16 and G17 GLM flashes, respectively. Both median (underlined) and mean values are reported for both the GOES-East (G16) and GOES-West (G17) GLMs. Flash area and maximum group area are reported in square kilometers (km^2), flash energy and maximum group energy are reported in femtojoules (1×10^{-15} J), the duration is reported in seconds, and the maximum number of consecutive frames and maximum number of events per group are counts.

	Flash area (km^2)	Maximum group area (km^2)	Flash energy (fJ)	Maximum group energy (fJ)	Maximum consecutive frames (count)	Maximum events per group (count)	Duration (s)
G16 Ignit	1341.8	388.8	345.0	129.3	1.0	5.0	0.171
	2344.6	486.5	1024.4	413.6	2.7	6.6	0.338
G16 All	1174.7	264.8	88.3	27.0	1.0	4.0	0.22
	2849.3	383.9	275.8	65.5	3.1	5.4	0.27
G17 Ignit	1738.0	460.6	285.6	99.1	1.0	6.0	0.209
	3446.1	579.0	861.9	308.5	3.9	7.8	0.335
G17 All	802.5	217.0	91.8	32.0	1.0	3.0	0.183
	1886.9	312.7	304.2	86.4	2.2	4.0	0.232

TABLE 6. Lightning flash type and polarity distributions for fire ignition flashes and null flashes observed by the ground-based NLDN. Note that the nulls include all CG strokes and every 35th IC pulse within 0.1° and 0.5° of the fire polygons to more closely match the CG/IC ratio of the ignition dataset.

	Count	CG (%)	IC (%)	Neg (%)	Pos (%)	Bipolar (%)
Ignits	2761	88.5	11.5	65.1	24.0	10.9
Nulls	289 060	87.8	12.2	61.5	22.9	15.6
Diff	−286 299	0.7	−0.7	3.6	1.1	−4.7
	Stroke count	CG count	IC count	maxIp	minIp	Duration
Ignits	3.97	2.49	1.48	19.72	−28.19	0.172
Nulls	4.04	2.38	1.66	15.85	−23.44	0.165
Diff	−0.07	0.11	−0.18	3.87	−4.75	0.007

determination of ignition strokes introduces subjectivity in our results. The WFIGS fire perimeter data layer has evolved over the years. The version used in this study included perimeters for the 2020–23 fire seasons, while the most recent version includes fires from 2021 to the present. This transition appears to have resulted in many fewer of the 2022 fires being designated as natural (rather than unknown) than in 2020–21.

TABLE 7. KS test statistic for select variables from each of the variable groups (see Table 8 for variable descriptions). Greater KS values indicate more difference between the ignition and null samples. Note that the p values are very low for most variables, and the largest p value for the variables shown is 10^{-4} (i.e., maxqpedist).

QPE	KSstat	MRMS	KSstat	HRRR	KSstat
localqpe24	0.33	nearestvil	0.27	RHavg	0.31
nearestqpe24	0.32	nearestref	0.26	elhRM	0.25
Localqpe	0.30	maxet18	0.22	capeRM	0.24
nearestqpe	0.29	nearestet18	0.21	RHmin	0.24
meanqpe24	0.24	maxvil	0.20	RHLV	0.23
medianqpe24	0.24	meanvil	0.19	dewptLV	0.23
maxqpe24	0.24	maxref	0.18	elhLV	0.22
coverage24	0.21	meanref	0.18	pwRM	0.21
Maxqpe	0.21	meanet18	0.15	pwLV	0.21
meanqpe	0.20	medianvil	0.14	capeLV	0.20
medianqpe	0.19	medianref	0.14	umotRM	0.17
coverage	0.16	coverageet18	0.13	dewptRM	0.17
nearestdist	0.14	medianet18	0.13	h0LV	0.17
nearestdist24	0.11	coveragevil	0.11	tavg	0.16
maxqpedist	0.04	coverageref	0.11	h0RM	0.16
Lightning	KSstat	Fire weather	KSstat		
Ip	0.17	NI_Index	0.30		
maxfed	0.16	Munger_Index	0.30		
avgfed	0.16	Nesterov_Class	0.29		
minIp	0.14	Max_24h_RH	0.29		
nearestMFA	0.12	FM100	0.26		
minmfa	0.12	FWI	0.25		
medmfa	0.10	BI	0.24		
medfed	0.10	ERC	0.24		
nearestFED	0.10	Max_24h_Dpt	0.22		
maxIp	0.08	2m_RH	0.22		
maxtoe	0.07	Min_24h_RH	0.20		
maxmfa	0.07	Angstrom_Index	0.19		
distance	0.06	FM1000	0.17		
avgmfa	0.06	%TimberLitter	0.17		
mintoe	0.05	VPD	0.16		

This increased uncertainty translated to fewer matches during 2022 and a larger proportion that remained unknown. This should have a minimal impact on the study conclusions because the null dataset excludes lightning near both the natural and unknown fire perimeters.

Reduced GLM performance at steeper viewing angles contributes to the GLMs only observing roughly half of the ignition flashes. The GLMs are more likely to observe larger and brighter flashes (Zhang and Cummins 2020), especially in regions of reduced performance. This likely skews the GLM characteristics toward the strongest flashes. It is likely that some of the unmatched ignition flashes are in the GLM level 0 data but are lost during filtering. Future work will search for these flashes and also examine the influence of the GLM performance on the ability to resolve continuing current.

Many studies have documented misclassification between CG strokes and IC pulses as well as polarity by the ground-based networks (e.g., Rudlosky and Fuelberg 2010; Mallick et al. 2015). The data vendors continuously work to improve the accuracy of these classifications (e.g., Murphy et al. 2021), but the variety of lightning waveforms and their relation to lightning physics make some misclassification inevitable. Our results indicate that lightning reported as IC pulses does start fires. In many of these cases, the alternate ground-based network classified the same lightning as CG strokes. This illustrates the importance of considering all available data sources whenever possible.

5. Summary

Exploring lightning patterns alongside other meteorological variables and fuel information helps to better diagnose the conditions under which lightning ignites fires. The study first describes a manual analysis method to develop a database of lightning strokes known to have ignited fires in the western United States during 2020–22. Distributions of lightning, land surface, and meteorological conditions associated with known lightning ignitions are described. Lightning counts peak during July each year with the counts of natural and unknown wildfires peaking during August each year. During the study years, 65% of all fires initially characterized as natural or unknown were confirmed to be natural (i.e., ~35% were human caused or had no matching lightning). This matched fraction peaks during July (76%) and August (80.1%) when the likelihood of natural ignition is greatest.

TABLE 8. Abbreviations and descriptions of select variables. An asterisk (*) indicates surrounding 1° × 1° scene.

Variable	Description	Variable	Description
LV	Local value	capeRM	Convective available potential energy*
RM	Regional mean*	RHLV	RH at flash location
QPE	Quantitative precipitation estimate	RHavg	RH 12-h average
QPE24	QPE (24 h)	RHmin	RH 12-h minimum
RH	Relative humidity	pwRM	Precipitable water regional mean*
Reflqc	Composite reflectivity	elhLV	Local/nearest ELH
echotop18	18-dBZ echo-top height	elhRM	ELH regional mean*
VIL	Vertically integrated liquid	Nesterov_Index	Nesterov ignition index
ELH	Height of the equilibrium level	Munger_Index	Munger index
H0	Melting level height	Nesterov_Class	Nesterov ignition index class
Localqpe	Local QPE at ignition location	Max24hRH	Maximum 24-h RH
Nearestqpe	Local/nearest nonzero QPE	FM100	100-h fuel moisture
Nearestdist	Distance to nearest nonzero QPE	FM1000	1000-h fuel moisture
localqpe24	Local QPE24 at ignition location	ERC	Energy release component
nearestqpe24	Local/nearest nonzero QPE24	FWI	Canadian fire weather index
meanqpe24	Mean QPE of all nonzero QPE*	BI	Burning index
coverage24	Fraction of nonzero QPE24 pixels*	VPD	Vapor pressure deficit
Nearestref	Local/nearest reflqc	Ip	Estimated peak current (stroke)
Maxref	Maximum reflqc*	maxfed	Maximum flash extent density*
nearestet18	Local/nearest echotop18	avgfed	Average flash extent density*
maxet18	Maximum echotop18*	mfaLV	Local minimum flash area
Nearestvil	Local/nearest VIL	mintoe	Minimum total optical energy*
Meanvil	Mean nonzero VIL*	distance	Distance to nearest GLM
Maxvil	Maximum VIL*	maxIp	Maximum Ip (+CG flash)
Dewpt	Dewpoint	minIp	Minimum Ip (−CG flash)
umotRM	U component of storm motion		

Compiling this fire-centric lightning climatology provides great insights into the types of lightning that ignite fires. Of the 4400 strokes/pulses linked to lightning ignition locations, 88.9% are CG, with 5.5% of those reported as +CG, while 67.1% of the reported IC pulses are characterized as +IC. Lightning misclassified as IC ignites fires, with roughly 11% of ignition strokes and flashes containing only IC components. The ENTLN classifies 17.1% of ignition flashes with

TABLE 9. Distributions of select environmental variables for the ignition and null strokes. Tables present the mean values for ignitions and nulls along with the difference (ignitions minus nulls). Diff % represents Diff/Nulls. See Table 8 for a description of variables.

	Local QPE	Nearest QPE	Local QPE24	Nearest QPE24	Mean QPE24	Coverage QPE24	
Ignits	2.76	2.93	5.16	5.25	4.01	0.62	
Nulls	7.08	7.20	12.69	12.73	6.36	0.75	
Diff	−4.32	−4.27	−7.53	−7.48	−2.35	−0.13	
Diff (%)	−61.2	−59.3	−59.3	−58.8	−36.9	−17.7	
	Nearest REF	Maximum REF	Nearest ET18	Maximum ET18	Nearest VIL	Mean VIL	Maximum VIL
Ignits	30.5	50.9	10.6	14.6	1.69	0.44	11.11
Nulls	37.3	54.0	12.1	15.7	3.98	0.63	15.97
Diff	−6.80	−3.10	−1.50	−1.10	−2.29	−0.19	−4.86
Diff (%)	−18.0	−5.8	−12.4	−6.9	−57.5	−28.6	−30.5
	RH LV	RH AVG	RH MIN	Pw RM	ELH LV	ELH RM	CAPE RM
Ignits	33.9	45.6	28.8	20.0	9802.3	8000.2	329.8
Nulls	44.3	58.6	37.7	22.4	10981.2	9296.8	485.4
Diff	−10.4	−13.0	−8.9	−2.4	−1178.9	−1296.6	−155.6
Diff (%)	−23.5	−22.2	−23.6	−10.7	−10.7	−13.9	−32.1
	Nesterov ignition index	Munger index	Nesterov index class	Maximum 24-h RH	100-h fuel moisture	FWI	Buildup index
Ignits	15 897.8	278 924.6	3.61	52.8	8.6	80.5	45.8
Nulls	8784.8	117 242.8	2.55	66.4	10.8	51.4	31.9
Diff	7113.0	161 681.8	1.06	−13.6	−2.2	29.1	13.9
Diff (%)	81.0	137.9	41.2	−20.6	−20.4	56.7	43.6

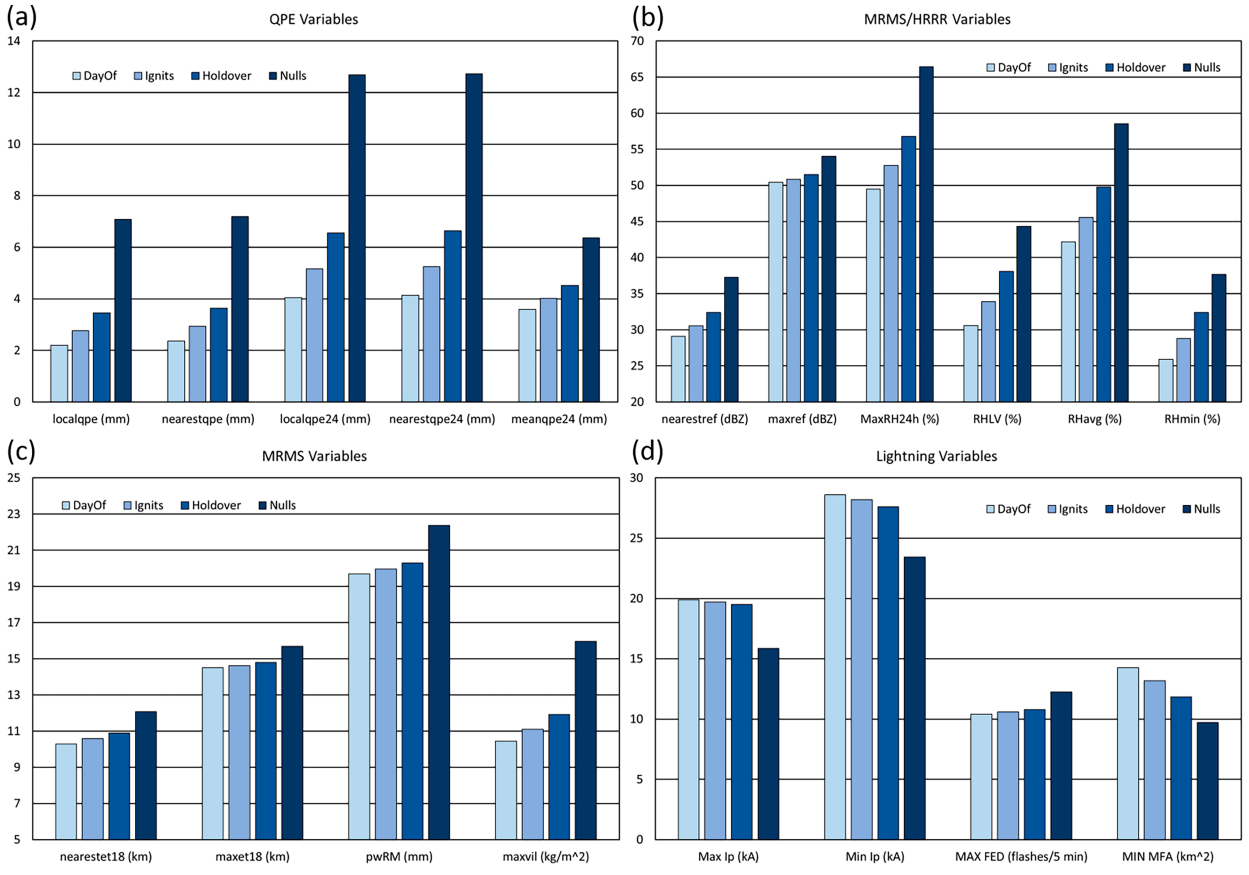

FIG. 5. Distributions of (a) QPE, (b) MRMS/HRRR, (c) MRMS, and (d) lightning variables for fires detected on the day of ignition (DayOf), all ignitions (Ignits), holdover fires (Holdover), and null samples (Nulls). Units are indicated by each parameter label, including QPE (mm), RH (%), reflectivity (dBZ), VIL (kg m^2), echo top (km), precipitable water (mm), I_p (kA), FED (count per 5 min), and MFA (km^2). Note that the MinIp is multiplied by −1 for display purposes and that the MFA was divided by 10 (i.e., the minimum DayOf I_p is −26.8 kA and the minimum DayOf MFA is 142.5 km^2).

only IC components reported by the NLDN as CG, and 20.5% (27.8%) of the single-stroke CG (IC) ignitions are classified as IC (CG) by the ENTLN. This illustrates the complexity of characterizing flash type and reveals the importance of considering all potential ignition sources (i.e., CG and IC). The GOES-East and GOES-West GLMs detect 37.8% and 51.9% of the ignition flashes, respectively. Manual inspection reveals GLM flashes in close proximity in nearly all cases with no direct GLM match. Despite missing ignition flashes, the types of lightning surrounding the ignitions can be suggestive of increased risk of fire ignitions. Our findings illustrate the importance of analyzing all available lightning data.

TABLE 10. Distributions of GLM and NLDN lightning variables for ignition flashes, those detected on the day of ignition, holdover fires, and null flashes.

	Nearest FED	Nearest MFA	AVG FED	MAX FED	MIN MFA	
DayOf	3.2	381.8	2.9	10.4	142.5	
Ignits	3.3	354.3	2.9	10.6	131.8	
Holdover	3.4	320.1	3	10.8	118.4	
Nulls	2.0	278.9	1.2	12.3	97.2	
	Stroke count	**CG count**	**IC count**	**MaxIp**	**MinIp**	**Duration**
DayOf	3.9	2.6	1.3	19.9	−28.6	0.177
Ignits	4.0	2.5	1.5	19.7	−28.2	0.172
Holdover	4.0	2.4	1.7	19.5	−27.6	0.165
Nulls	4.0	2.4	1.7	15.9	−23.4	0.165

Lightning variables differ between the ignition and null samples. The ignition and null flashes have similar proportions of CG (~88%) and IC (~12%) flashes (by design), similar average durations, and comparable polarity distributions. The estimated peak current differs most between the ignition and null flashes. On average, positive and negative polarity ignition flashes are 3.87 and 4.75 kA stronger than null flashes. Both GLMs indicate that ignition flashes are larger and 3–4 times brighter than the domainwide sample. The GOES-West instrument performs best in our study domain, revealing clearer differences between the ignition and null samples. Our findings indicate that both the NLDN and GLM observations can help identify lightning more likely to ignite wildfires.

The environmental variables reveal a clear distinction between the ignition and null storm environments. The average hourly (daily) QPE is 7.08 (12.69) mm for the null flashes and only 2.76 mm (5.16 mm) for ignitions. The MRMS variables indicate that storms surrounding the ignitions are weaker than storms surrounding null flashes. Moisture variables dominate the selection of model-derived fields. Several different measures of RH indicate ~10% drier environments during ignitions. The Nesterov ignition, Munger, and FWI are all much larger for the ignitions than for the nulls.

Ignition samples are subdivided further into those observed on the day of lightning ignition (56%) and holdover fires (44%). Ignitions detected on the day of differ most from the null samples for each of the variables shown. Day of detections has ~33% of the 1- and 24-h precipitation and 15% less RH than the null cases. Alternatively, the 1-h (24-h) QPE is 1.26 (2.51) mm greater for the holdover fires versus those detected on the day of ignition, and the RH is ~7% greater for the holdover fires. The NLDN indicates the strongest average I_p for day of ignitions and the weakest average I_p for null samples. More frequent smaller GLM flashes occur in null environments, while day of detections has less frequent larger GLM flashes.

Our findings indicate clear differences between both the ignitions and nulls, and the fires detected on the day of versus those that holdover. Each of the variables indicates that fires detected on the day of ignition occur in the most fire-prone environments, null cases represent the least conducive environments, and holdover fires fall somewhere in between. Our analysis suggests that distributions of lightning, land surface, and meteorological conditions can be used to help identify environments most vulnerable to lightning ignitions. This knowledge will help guide the development of automated applications for identifying the most impactful lightning collocated with conditions conducive to wildfire ignition.

Acknowledgments. Funding for this project was provided by NOAA/NESDIS through the Cooperative Institute for Satellite Earth System Studies (CISESS; NA19NES4320002) via Award NA22NES4050010I and the Cooperative Institute for Meteorological Satellite Studies via Award NA22NES405006I. The contents of this paper are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U.S. government.

Data availability statement. The high-resolution NLDN and ENTLN lightning data are not publicly available but can be purchased directly from Vaisala, Inc. (https://www.vaisala.com/en/products/national-lightning-detection-network-nldn), or Earth Networks (https://www.earthnetworks.com/product/lightning-data/). GOES-East and GOES-West GLM level 2 data and ABI fire hotspots are available via Amazon Web Services (AWS) at https://registry.opendata.aws/noaa-goes/. Gridded GLM products are available at https://search.earthdata.nasa.gov/search/granules/collection-details?portal=ghrc&p=C2278812167-GHRC_DAAC&pg[0][v]=f&q=glm. The Wildland Fire Interagency Geospatial Services (WFIGS) wildfire perimeters are available at https://services3.arcgis.com/T4QMspbfLg3qTGWY/arcgis/rest/services/WFIGS_Interagency_Perimeters/FeatureServer. The MODIS and VIIRS fire hotspot data are available at https://firms.modaps.eosdis.nasa.gov/download/. MRMS data are available at the Iowa State University's Environmental Mesonet website (https://mtarchive.geol.iastate.edu/). HRRR model data are available via AWS at https://registry.opendata.aws/noaa-hrrr-pds/. LANDFIRE data are available at https://landfire.gov/getdata.php. Drought variables were obtained from an internal University of Wisconsin–Madison (UW) Cooperative Institute for Meteorological Satellite Studies (CIMSS) archive. Combinations of these datasets were used to compute fire weather indices. Various Python packages were employed to compute many meteorological and land surface variables (e.g., xarray, geopandas, rioxarray, and rasterio).

REFERENCES

Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, **33**, 121–131, https://doi.org/10.1002/joc.3413.

Anderson, M. C., J. M. Norman, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas, 2007: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *J. Geophys. Res.*, **112**, D10117, https://doi.org/10.1029/2006JD007506.

Balch, J. K., B. A. Bradley, J. T. Abatzoglou, R. C. Nagy, E. J. Fusco, and A. L. Mahood, 2017: Human-started wildfires expand the fire niche across the United States. *Proc. Natl. Acad. Sci. USA*, **114**, 2946–2951, https://doi.org/10.1073/pnas.1617394114.

Baumgartner, A., L. Klemmer, E. Raschke, and G. Waldmann, 1967: Waldbrände in bayern 1950 bis 1959. *Mitteilungen Aus der Staatsforstverwaltung Bayerns*, Vol. **36**, 57–79.

Bitzer, P. M., 2017: Global distribution and properties of continuing current in lightning. *J. Geophys. Res. Atmos.*, **122**, 1033–1041, https://doi.org/10.1002/2016JD025532.

Chandler, C., P. Cheney, P. Thomas, L. Trabaud, and D. Williams, 1983: Forest fire behavior and effects. *Fire in Forestry*, Vol. **1**, John Wiley and Sons, 450 pp.

Ding, Z., and Coauthors, 2024: Continuing current seen above and below the cloud: Comparing observations from GLM and high-speed video cameras. *Geophys. Res. Lett.*, **51**, e2024GL110099, https://doi.org/10.1029/2024GL110099.

Dowdy, A. J., and G. A. Mills, 2009: Atmospheric states associated with the ignition of lightning-attributed fires. Centre for Australian Weather and Climate Research Tech. Rep. 19, 34 pp.

Fairman, S. I., and P. M. Bitzer, 2022: The detection of continuing current in lightning using the geostationary lightning mapper. *J. Geophys. Res. Atmos.*, **127**, e2020JD033451, https://doi.org/10.1029/2020JD033451.

Fosberg, M. A., 1978: Weather in wildland fire management: the fire weather index. *Proc. Conf. on Sierra Nevada Meteorology*, Lake Tahoe, CA, Amer. Meteor. Soc., 1–4.

Hobbins, M. T., A. Wood, D. J. McEvoy, J. L. Huntington, C. Morton, M. Anderson, and C. Hain, 2016: The evaporative demand drought index. Part I: Linking drought evolution to variations in evaporative demand. *J. Hydrometeor.*, **17**, 1745–1761, https://doi.org/10.1175/JHM-D-15-0121.1.

Kalashnikov, D. A., J. T. Abatzoglou, P. C. Loikith, N. J. Nauslar, Y. Bekris, and D. Singh, 2023: Lightning-ignited wildfires in the western United States: Ignition precipitation and associated environmental conditions. *Geophys. Res. Lett.*, **50**, e2023GL103785, https://doi.org/10.1029/2023GL103785.

Koshak, W. J., 2010: Optical characteristics of OTD flashes and the implications for flash-type discrimination. *J. Atmos. Oceanic Technol.*, **27**, 1822–1838, https://doi.org/10.1175/2010JTECHA1405.1.

Latham, D., and E. Williams, 2001: Lightning and forest fires. *Forest Fires*, Academic Press, 375–418.

Latham, D. J., and J. A. Schlieter, 1989: Ignition probabilities of wildland fuels based on simulated lightning discharges. U.S. Department of Agriculture, Forest Service, Intermountain Research Station Research Paper INT-411, 20 pp., https://www.frames.gov/documents/behaveplus/publications/Latham_and_Schlieter_1989_INT-RP-411_ocr.pdf.

Long, J. L., and T. D. Hatten, 2023: LANDFIRE. U.S. Geological Survey Fact Sheet 2023–3044, 4 pp., https://doi.org/10.3133/fs20233044.

MacNamara, B. R., C. J. Schultz, and H. E. Fuelberg, 2020: Flash characteristics and precipitation metrics of western U.S. lightning-initiated wildfires from 2017. *Fire*, **3**, 5, https://doi.org/10.3390/fire3010005.

Mallick, S., and Coauthors, 2015: Performance characteristics of the ENTLN evaluated using rocket-triggered lightning data. *Electr. Power Syst. Res.*, **118**, 15–28, https://doi.org/10.1016/j.epsr.2014.06.007.

Marinescu, P. J., D. Abdi, K. Hilburn, I. Jankov, and L.-F. Lin, 2024: An evaluation of NOAA modeled and in situ soil moisture values and variability across the continental United States. *Wea. Forecasting*, **39**, 523–540, https://doi.org/10.1175/WAF-D-23-0136.1.

Massey, F. J., Jr., 1952: Distribution table for the deviation between two sample cumulatives. *Ann. Math. Stat.*, **23**, 435–441, https://doi.org/10.1214/aoms/1177729388.

McEachron, K. B., and J. H. Hagenguth, 1942: Effect of lightning on thin metal surfaces. *Electr. Eng.*, **61**, 559–564, https://doi.org/10.1109/T-AIEE.1942.5058563.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. *Eighth Conf. on Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184, https://www.droughtmanagement.info/literature/AMS_Relationship_Drought_Frequency_Duration_Time_Scales_1993.pdf.

Munger, T. T., 1916: Graphic method of representing and comparing drought intensities. *Mon. Wea. Rev.*, **44**, 642–643, https://doi.org/10.1175/1520-0493(1916)44<642:GMORAC>2.0.CO;2.

Murphy, M. J., J. A. Cramer, and R. K. Said, 2021: Recent history of upgrades to the U.S. National Lightning Detection Network. *J. Atmos. Oceanic Technol.*, **38**, 573–585, https://doi.org/10.1175/JTECH-D-19-0215.1.

Nesterov, V. G., 1949: Combustibility of the forest and methods for its determination (in Russian). USSR State Industry Press, 76 pp.

North, M. P., and Coauthors, 2024: Strategic fire zones are essential to wildfire risk reduction in the western United States. *Fire Ecol.*, **20**, 50, https://doi.org/10.1186/s42408-024-00282-y.

Osborne, A. P., J. Zhang, M. J. Simpson, K. W. Howard, and S. B. Cocks, 2023: Application of machine learning techniques to improve Multi-Radar Multi-Sensor (MRMS) precipitation estimates in the western United States. *Artif. Intell. Earth Syst.*, **2**, 220053, https://doi.org/10.1175/AIES-D-22-0053.1.

Pérez-Invernón, F. J., H. Huntrieser, and J. V. Moris, 2022: Meteorological conditions associated with lightning ignited fires and long-continuing-current lightning in Arizona, New Mexico and Florida. *Fire*, **5**, 96, https://doi.org/10.3390/fire5040096.

Prestemon, J. P., T. J. Hawbaker, M. Bowden, J. Carpenter, M. T. Brooks, K. L. Abt, R. Sutphen, and S. Scranton, 2013: Wildfire ignitions: A review of the science and recommendations for empirical modeling. USDA-Forest Service, Southern Research Station General Tech. Rep. SRS-GTR-171, 20 pp., https://research.fs.usda.gov/treesearch/42766#.

Ringhausen, J., P. Bitzer, W. Koshak, and J. Mecikalski, 2021: Classification of GLM flashes using random forests. *Earth Space Sci.*, **8**, e2021EA001861, https://doi.org/10.1029/2021EA001861.

Rudlosky, S. D., 2015: Evaluating ENTLN performance relative to TRMM/LIS. *J. Oper. Meteor.*, **3**, 11–20, https://doi.org/10.15191/nwajom.2015.0302.

——, and H. E. Fuelberg, 2010: Pre- and postupgrade distributions of NLDN reported cloud-to-ground lightning characteristics in the contiguous United States. *Mon. Wea. Rev.*, **138**, 3623–3633, https://doi.org/10.1175/2010MWR3283.1.

——, and K. S. Virts, 2021: Dual geostationary lightning mapper observations. *Mon. Wea. Rev.*, **149**, 979–998, https://doi.org/10.1175/MWR-D-20-0242.1.

Schmidt, C., J. Hoffman, E. Prins, and S. Lindstrom, 2012: GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for fire/hot spot characterization. NOAA NESDIS STAR Doc., 97 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_FIRE_v2.6_Oct2013.pdf.

Schultz, C. J., N. J. Nauslar, J. B. Wachter, C. R. Hain, and J. R. Bell, 2019: Spatial, temporal and electrical characteristics of lightning in reported lightning-initiated wildfire events. *Fire*, **2**, 18, https://doi.org/10.3390/fire2020018.

——, P. M. Bitzer, M. Antia, J. L. Case, and C. R. Hain, 2024: Assessing flash characteristics in lightning-initiated wildfire events between 1995 and 2020 within the contiguous United States. *J. Appl. Meteor. Climatol.*, **63**, 543–551, https://doi.org/10.1175/JAMC-D-23-0166.1.

Seto, D., C. Jones, D. Siuta, N. Wagenbrenner, C. Thompson, and N. Quinn, 2025: Evaluation of HRRR wind speed forecast and WindNinja downscaling accuracy during Santa Ana wind events in Southern California. *Wea. Forecasting*, **40**, 525–541, https://doi.org/10.1175/WAF-D-24-0013.1.

Sharples, J. J., R. H. D. McRae, R. O. Weber, and A. M. Gill, 2009a: A simple index for assessing fuel moisture content. *Environ. Modell. Software.*, **24**, 637–646, https://doi.org/10.1016/j.envsoft.2008.10.012.

——, ——, ——, and ——, 2009b: A simple index for assessing fire danger rating. *Environ. Modell. Software.*, **24**, 764–774, https://doi.org/10.1016/j.envsoft.2008.11.004.

Smirnov, N. V., 1939: Estimate of deviation between empirical distribution functions in two independent samples (in Russian). *Bull. Moscow Univ.*, **2**, 3–16.

Taylor, A. R., 1969: Lightning effects on the forest complex. *Ninth Tall Timbers Fire Ecology Conf.*, Tallahassee, FL, Association for Fire Ecology and Tall Timbers, 127–150, https://talltimbers.org/wp-content/uploads/2018/09/127-Taylor1969_op.pdf.

Van Wagner, C. E., 1987: Development and structure of the Canadian Forest Fire Weather Index System. Canadian Forest Service, Forestry Tech. Rep. 35, 37 pp.

——, and T. L. Pickett, 1985: Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. Canadian Forest Service, Forestry Tech. Rep. 33, 18 pp.

WFIGS, 2024: The Wildland Fire Interagency Geospatial Service Interagency Fire Perimeters. National Interagency Fire Center, accessed 4 May 2023, https://data-nifc.opendata.arcgis.com/datasets/nifc::wfigs-interagency-fire-perimeters/explore?location=0.000000%2C0.000000%2C1.82.

Zhang, D., and K. L. Cummins, 2020: Time evolution of satellite-based optical properties in lightning flashes, and its impact on GLM flash detection. *J. Geophys. Res.*, **125**, e2019JD032024, https://doi.org/10.1029/2019JD032024.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.