https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

Distribution and Characteristics of Lightning-Ignited Wildfires in Boreal Forests - the BoLtFire database

Brittany Engle[1], Ivan Bratoev[2], Morgan A. Crowley[3], Yanan Zhu[4], Cornelius Senf[1]

[1] Technical University of Munich, School of Life Sciences, Earth Observation for Ecosystem Management, Freising, Germany
[2] Technical University of Munich, School of Engineering and Design, Architectural Informatics, Munich, Germany
[3] Canadian Forest Service (Great Lakes Forestry Centre), Natural Resources Canada, Sault Sainte Marie, Ontario, Canada
[4] Advanced Environmental Monitoring (AEM), Germantown, Maryland, USA

Correspondence to: Brittany Engle (Brittany.Engle@tum.de)

Abstract. The frequency and severity of fire weather have increased under climate change, particularly in high-latitude boreal forests. Lightning, a key ignition source globally, is also expected to become more frequent with climate change and could significantly increase burn area. Current research on lightning-ignited wildfires (LIW) has a long history in boreal ecosystems but has typically focused on North America due to better data availability, while the lack of publicly available data for Eurasia has hindered our comprehensive understanding of important characteristics of LIW, such as holdover time, lightning-ignition efficiency, frequency, and spatial distribution of lightning-ignited wildfires in boreal forests. This study introduces the Temporal Minimum Distance (TMin) method, a novel approach to matching lightning strikes with wildfires without requiring ignition location, that outperformed current methodologies. As a result, we developed a comprehensive dataset of lightning-ignited wildfires across the entire boreal forest from 2012 to 2022, encompassing 6,228 fires — 4,186 in Eurasia and 2,042 in North America — each over 200 hectares in size. This dataset provides new opportunities to model ignition and spread dynamics of boreal wildfires and offers deeper insights into lightning-driven fire activity globally.

1. Introduction

Forest fires are the primary disturbance agent in global boreal forests, burning an estimated 10-15 million ha each year on average across Siberia, Canada, and Alaska (Flannigan et al., 2009). Boreal fires play a crucial role in shaping boreal forest composition and structure (Podur et al., 2003), and they also have a significant impact on the global carbon cycle (Stocks et al., 2001). The global boreal forests house roughly 32 % of terrestrial carbon stocks (Pan et al., 2011) and are considered to be globally important carbon sinks. However, due to the observed and predicted increases in temperatures in high latitudes (Melillo et al., 2014; Natali et al., 2019; Post et al., 2019; Soja et al., 2007), thawing of the permafrost, increase in vegetation stress due to drought, and an increase in fire frequency and burned area (Gillett et al., 2004; Flannigan et al., 2009), there is concern that boreal forests could be pushed to be a net positive carbon source (Watts et al., 2023 and references therein). Since at least 2000, carbon dioxide emissions from boreal forest fires have been on the rise, reaching a record high in 2021, where

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

they contributed to 23 % of global wildfire carbon dioxide emissions, a significant increase from the historical average of 10 % (Zheng et al., 2023). Additionally, Black carbon from boreal forest fires contributes to increased pollution levels in the Arctic (Bond et al., 2013; Lavoué et al., 2000; Stohl, 2006) and could accelerate losses of snow and ice (Hansen and Nazarenko, 2004). Kim et al.'s (2005) findings indicate, for example, that black carbon aerosols (soot) are rapidly transported from central Alaska to the Arctic Ocean, as well as to glaciers in southern Alaska, where up to 20 % of these aerosols may be deposited, changing the albedo and possibly increase the melting of the arctic. All in all, increasing forest fires will challenge the boreal carbon sink and it is thus essential to understand global distribution and trends in boreal forest fires.

One of the major processes underlying the observed increase in forest fires in boreal regions is a notable rise in the frequency and intensity of fire weather, a trend that is projected to worsen with global warming (Hessilt et al., 2022; Jones et al., 2022). This trend in fire weather frequency has already led to approximately 50 % increases in burned area between 2001–2019 in certain extratropical forest ecoregions, such as the Pacific US and high-latitude forests, with an expected increase in forest fire activity and severity in some higher-latitude regions (De Groot et al., 2013; Descals et al., 2022; Flannigan et al., 2000; Stocks et al., 1998; Zheng et al., 2023). Yet, for a forest to burn there needs to be a source of ignition, with lightning being a key source of wildfire ignition in boreal forests (Gao et al., 2024; Moris et al., 2020; Pérez-Invernón et al., 2023; Sofronov et al., 1998; Veraverbeke et al., 2017). Lightning-ignited wildfires are a significant cause of burn area in boreal forests (Hanes et al., 2019; Kasischke et al., 2002; Nash and Johnson, 1996; Veraverbeke et al., 2017), and they are more difficult to detect, suppress and extinguish than human-caused fires (Flannigan and Wotton, 1991; Kourtz and Todd, 1991; Podur et al., 2003; Wotton and Martell, 2005). Under the influence of climate change, the scale and occurrence of lightning-ignited wildfires in boreal forests are projected to increase significantly. Janssen et al. (2023), for instance, estimated that 77 % of the burned areas in extratropical intact forests are attributable to lightning, with lightning occurrences projected to increase by 11 % to 31 % for each degree of warming. Additionally, Krause et al. (2014) suggest a potential 21.3 % rise in cloud-to-ground lightning activity by the end of the century under the RCP8.5 scenario, potentially doubling the burned area in high-latitude regions. That said, the lightning-ignition efficiency, i.e. the number of fires ignited per lightning, can differ significantly by region and land cover (Podur et al., 2003). Lightning storms can also result in concentrated clusters of large numbers of fires (Flannigan and Wotton, 1991; Kourtz and Todd, 1991; Podur et al., 2003; Woolford et al., 2021) and lightning-ignited wildfires can smoulder for long periods of time before igniting fully, in what is referred to as the holdover time (time between fire ignition and detection; Flannigan and Wotton, 1991; Wotton and Martell, 2005). Holdover time presents a true challenge for real-time lightning-ignited wildfire detection in boreal forests. They are still not fully understood and likely depend heavily on initial detection and ignition characteristics. Their duration can vary heavily, with some literature suggesting up to around a week (Anderson, 2002; Flannigan and Wotton, 1991; Moris et al., 2023), while others suggest they can last up to a few weeks in boreal forests (Moris et al., 2023; Nash and Johnson, 1996; Scholten et al., 2021; Wotton and Martell, 2005). Scholten et al. (2021) reported that fire managers in Alaska and Canada have started reporting extreme holdover times, where fires “hibernate” over winter (up to seven or eight months) only to re-emerge the following fire season as “overwintering fires.” The spatial distribution,

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

characteristics and potentially complex interactions of lightning-ignited wildfires in boreal forests are thus not fully understood, partly due to a lack of global data on lightning-ignited fires.

Recent advancements in remote sensing and lightning location systems have significantly enhanced our understanding of lightning-ignited wildfires. More specifically, the precise geolocation of lightning paired with spatially explicit information on wildfires allows us to distinguish lightning from human-caused ignitions by matching wildfire and lightning location data (Larjavaara et al., 2005; Nash and Johnson, 1996; Wotton and Martell, 2005). Current methods search for the most likely individual lightning event based on its temporal and spatial distance to the fire's ignition point. These methods use a buffer area around the ignition point and a backward temporal window to account for location and holdover time and typically stem from Larjavaara et al.'s (2005) proximity index: a spatio-temporal index to match candidate lightning to fires using holdover time (t) and spatial distance (S). Variations of lightning and wildfire matching methodologies are used often in scientific literature and have been implemented on a local, regional, and global scale; a few case studies specifically from a boreal perspective include: Finland (Larjavaara et al., 2005), northeastern China (Gao et al., 2024), North America (Hessilt et al., 2022; Nash and Johnson, 1996), and Siberia (Xu et al., 2022). Testing the effectiveness of some of these variations, Moris et al. (2020) found that the Maximum Index A (MaxA) and Daily Minimum Distance (DMin) produced a high match between candidate lightning (lightning that was most likely to start the ignition) and the ignition, as well as the lowest values of distance and holdover time (ex: MaxA within a 10 km radius and a 14-day holdover time to be highly effective, with 80 % of fires detected within 3 days). However, these existing methodologies rely on the availability of ignition point locations, which are often absent in datasets produced without input from local fire and forest authorities. Most research on lightning-ignited wildfires has thus been conducted in North America, where fire point and polygon datasets are available from the Canadian National Fire Database (Canadian Forest Service, 2024) and the Bureau of Land Management's Alaska Fire Service (U.S. Department of the Interior, Bureau of Land Management, Alaska Fire Service, 2024a&b). Both datasets include key information such as fire start and end date and ignition cause and location. Unfortunately at this time, there are no other publicly available agency datasets within the boreal biome that provide this crucial information. Both existing products are built upon local context and knowledge provided by forest authorities and personnel; which is not reproducible on a global scale. Although global remote sensing-based wildfire datasets do exist like Collection 6 MCD64A1 burned area (Giglio et al., 2018), FRY (Laurent et al., 2018), and FIRECCI5 (Lizundia-Loiola et al., 2020), they lack ignition location and cause. Consequently, a comprehensive pan-boreal forest lighting-ignited wildfire dataset does not currently exist.

Here, we introduce the pan-boreal forest lightning-ignited wildfire dataset (BoLtFire), an Earth-observation-based database on lightning-ignited wildfire across global boreal forests. Our database overcomes challenges with existing methods by introducing a new methodology specifically designed for detecting lightning-ignited wildfires in boreal forests and without exact ignition location. More specifically, the objectives of our study were to:

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

1. Develop and implement a new approach to identify lightning-ignited wildfires that do not rely on ignition locations and benchmark this new approach with a regional test case using the Canadian National Fire Database and Alaska Fire Service datasets
2. Apply the new approach globally across the boreal forest biome to create a pan-boreal dataset of lightning-ignited wildfires across the global boreal forest, leveraging lightning location data from the Earth Networks Total Lightning Network (ENTLN) and fire location data from the GlobFire Fire Perimeters dataset (Artés et al., 2019)
3. Enhance understanding of holdover time, distance to fire perimeter, lightning-ignition efficiency, frequency of fires, total burned area, and spatial distribution of lightning-ignited wildfires in boreal forests

2. Data and methods

2.1 Study area

The boreal zone, a vast circumpolar vegetation region situated between latitudes 45˚N and 70˚N, represents one of the world's largest biogeoclimatic zones. It plays a critical role in providing renewable resources, supporting diverse habitats, regulating global climate, sequestering carbon, and is recognized as one of the fastest-warming biomes (Brandt, 2009; Brandt et al., 2013; Zheng et al., 2023). Spanning 1.89 billion hectares, the boreal zone encompasses 1,133 million hectares in Russia, 552 million hectares in Canada, and 73.7 million hectares in the United States, with additional areas across Finland, Sweden, China, Norway, Mongolia, Iceland, Kazakhstan, Greenland, and Saint Pierre and Miquelon (Brandt, 2009; Brandt et al., 2013).

2.2 Boreal forest fire regimes

A lightning-ignited fire progresses through three primary phases: (1) ignition, (2) survival (or smouldering), and (3) arrival (Anderson, 2002; Kourtz and Todd, 1991; Martell and Sun, 2008; Pérez-Invernón et al., 2023). The survival and arrival phases of ignition depend heavily on the fuel availability and composition, weather conditions and topography (Anderson, 2002; Flannigan and Wotton, 1991; Kourtz and Todd, 1991; Martell and Sun, 2008; Pérez-Invernón et al., 2023). During the ignition phase, lightning triggers an ignition which can then smoulder within the duff layers. Certain duff layers, such as the needle-covered ground beneath conifers, are particularly susceptible to ignition and prolonged smouldering (Flannigan and Wotton, 1991). The ignition will either “self-extinguish” or if the conditions are conducive to fire, the fire will enter the survival and arrival phase (Anderson, 2002; Flannigan and Wotton, 1991; Martell and Sun, 2008; Pérez-Invernón et al., 2023).

Wildfires are a natural process and important drivers of forest dynamics (Latham and Williams, 2001; Podur et al., 2003; Seidl et al., 2020; Vajda et al., 2013). Over long time scales, fires will create a mosaic of burned, recovering and unburned patches, with patch characteristics (i.e. size and shape), fire frequency and recovery dynamics largely determining the resulting landscape mosaic (Turner, 2010). Fire-adapted forests can typically be classified into two distinct categories; (1) forests with species that can regenerate independently and (2) forests that require species members to regenerate. The first category includes

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

species such as conifers which store their seeds in insulated serotinous cones that only open to heat, and secondly, hardwoods which regenerate from the root layer after a fire (Stocks et al., 2001). North American boreal forests, for example, contain Pinus banksiana (jack pine) and Picea mariana (black spruce), which both require fire to regenerate (Rogers et al., 2015; Stocks et al., 2001). In contrast, Eurasian boreal forests are dominated by non-serotinous conifers and other species that release seeds annually, resulting in heterogeneous, uneven-aged forests (Stocks et al., 2001). Due to the predominant tree species, Rogers et al. (2015) suggest that North American boreal forests are more prone to high-intensity crown fires, which consume large amounts of vegetation and detritus, whereas Eurasian forests typically experience lower-intensity surface fires that burn less vegetation and fewer trees (De Groot et al., 2013; Rogers et al., 2015).

2.3 Summary of datasets

2.3.1 Lightning data

The Earth Networks Total Lightning Network is a global network of 1800 sensors across more than 100 countries (Liu and Heckman, 2011; Zhu et al., 2022). Their network is further enhanced by an integration with the World Wide Lightning Location Network (Rodger et al., 2004). The ENTLN detects wideband electric field signals (1 Hz to 12 MHz) that are emitted by cloud-to-ground and intracloud lightning. Only cloud-to-ground flash data are used in this study. Each cloud-to-ground flash has at least one return stroke (Liu and Heckman, 2011). Intracloud pulses occur within the cloud and do not strike the ground as a return stroke, making them irrelevant for wildfire ignition. Using the time-of-arrival technique, they report the time, location (latitude and longitude), type of discharge, peak current, and polarity (positive or negative), of each lightning event. Individual strokes are clustered into a flash if they are located within 10 kilometres and within 0.7s of each other (Liu and Heckman, 2011). For each flash, multiplicity, start time, end time, and duration are reported.

Detection efficiency refers to the percentage of flashes or strokes detected by the network and can vary depending on the location and the distance between the sensors and the lightning event. Relative detection efficiency assumes a uniform detection efficiency across the network. The ENTLN's cloud-to-ground stroke detection efficiency across the CONUS is reported to be greater than 90 % (Lapierre et al., 2020), with relative detection efficiency values ranging from 85 % to 100 % across the Americas and similar levels in Europe and Australia (Bui et al., 2015). Significant improvements have been observed over the years due to advancements in processor technology (Mallick et al., 2013; Mallick et al., 2015; Zhu et al., 2017; Zhu et al., 2022). The most recent processor has achieved a stroke classification accuracy of 94 % and has reduced the median location error from 215 meters to 92 meters when compared to ground truth data in Florida (Zhu et al., 2022). Globally, this upgrade has resulted in a 149 % increase in the overall global detection of pulses, with North America showing a 145 % gain, Asia a 142 % gain, and Europe a 152 % gain. However, in more remote areas, detection efficiency tends to be lower, and location errors can extend up to several kilometres. Additionally, there are no sensors located in Russia, all lightning is detected by sensors in neighbouring countries.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

2.3.2 Wildfire data

The MODIS Collection 6 MCD64A1 burned area product (Giglio et al., 2018) utilizes daily 500-meter resolution MODIS surface reflectance data in conjunction with 1-kilometer resolution MODIS active fire observations (Giglio et al., 2020). The GlobFire algorithm was applied to the MCD64A1 product to create the GlobFire Fire Perimeters [2002-2023] dataset (Artés et al., 2019). Each yearly fire shapefile has a unique fire id, an initial date, a final date, the geometry and the final area in hectares for each fire within. Within their algorithm, a fire event is a set of burned areas that are intersecting or touching. Each fire event comprises interconnected burned areas that are only treated as separate fires if there is a temporal distance of more than five days between each event. If there is no new burned area after 16 days, the fire is no longer considered active.

2.3.3 Agency reference fire occurrence records

We used the Canadian National Fire Database Fire Point and Polygon datasets along with the Alaska Fire History Location Points and Perimeter Polygons datasets to assess our BoLtFire dataset. The Canadian National Fire Database fire points and polygons are compiled from various jurisdictions, including provinces, territories, and Parks Canada; and represent a collaborative effort by all Canadian fire agencies. This database is the most comprehensive resource for fire points and perimeters in Canada. The point data includes critical fire information such as fire ID, fire report date, out date, cause, and ignition point. Hanes et al. (2019) states that these points are just the presumed points of ignition. The polygon data includes similar information, with fire ID, fire report date, out date, cause, and map source, available until December 2020. Fire perimeters are typically derived through the interpretation of Landsat (30m) or other satellite imagery, and are sometimes derived from aerial or field surveys by fire management agencies.

The Bureau of Land Management's Alaska Fire Service maintains a detailed record of all detected natural or human-caused wildfire events in Alaska from 1940 to 2023. The Alaska Fire History Location Points dataset provides fire-specific information, including fire ID, discovery date, out date, estimated fire size, and cause. The perimeters in the Alaska Fire History Perimeter dataset were mapped using ground and airborne surveys, as well as aerial photography and Sentinel-2 satellite imagery (10m). This dataset includes details such as fire ID, estimated burned area size, map source, geographical coordinates, and fire out date, but does not include the fire start date or cause.

2.4 Processing of the lightning and wildfire datasets

Both the lightning and wildfire location datasets utilized in this study are global products and were initially filtered in order to only include those points which fell within the Boreal Forests/Taiga biome as defined by Ecoregions 2017 (Olson et al., 2001). Due to detection efficiency and location errors of lightning, the ENTLN was only filtered to the Biome location, while the GlobFire dataset went through additional filtering. Given MODIS burned area product's inability to adequately detect smaller fires, those less than 200 ha were excluded from the analysis (Giglio et al., 2009). This should not significantly influence the overall outcome of this dataset, as typically within boreal forests, large fires account for 85 % of the burned area (Macias

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Fauria and Johnson, 2007). Fires were then filtered based on the majority land cover class using the MODIS MCD12Q1v061 Land Cover Type 1: Annual IGBP classification datasets from 2012-2022, with a 500 m spatial resolution (Friedl and Sulla-Menashe, 2022). Non-forest and non-native forest land covers were excluded from the analysis (Table A1). Fire class size was then added based on Talucci et al. (2022) (Table A2). Each fire was then assigned a country using the same methodology along with the World Bank Official Boundaries dataset (World Bank, 2020). We removed three fires as they did not have a country location. A complete list of input datasets used can be found in Table 1.

Both the lightning and wildfire location datasets were then converted from their original coordinate system to their respective Global LANd Cover mapping and Estimation (GLANCE) Grids - Version 01 CRS (North American, Europe, and Asia) for the matching process (Arevalo et al., 2022). Though the ENTLN dataset was only filtered by its biome, it did receive land cover and country labels for additional analysis. The filtering process identified a total of 27,525,747 lightning flashes within the boreal forest biome, with 12,526,781 occurring in North America and 14,998,966 in Eurasia. Data for 4 November 2014, 9 May 2016, and 19 May 2019 were unavailable. Additionally, the process identified 31,363 total GlobFire events, comprising 5,290 in North America and 26,073 in Eurasia.

Table 1: Description of variables provided for each lightning-ignited wildfire within the BoLtFire dataset.

Column Name	Description
FireID	Unique fire identification number
StartDate	Start date of the fire
EndDate	End date of the fire
FireYear	Year fire was discovered
AreaHa	Total burned area of the fire in hectares
ClassSize	Fire class size; Small ≤ 1,000 ha, Moderate 1,000 ≤ 10,000 ha, Large 10,000 ≤ 50,000 ha, Extremely Large 50,000 ≤ 100,000 ha, Mega Fires > 100,000 ha
BiomeName	Biome name based on Olson et al. (2001)
EcoBiome	Ecoregion Biome number based on Olson et al. (2001)
EcoName	Ecoregion name based on Olson et al. (2001)

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

EcoID	Ecoregion ID based on Olson et al. (2001)
Realm	Realm based on Olson et al. (2001)
LCDN	Land cover type number based on Friedl and Sulla-Menashe (2022)
LCName	Land cover type name based on Friedl and Sulla-Menashe (2022)
Country	Country where fire is located based on World Bank (2020)
Continent	Continent on which the fire was located
HoldoverD	Holdover in days
HoldoverRD	Holdover in days, rounded to the day
IgnLat	Latitude location of the candidate lightning
IgnLong	Longitude location of the candidate lightning
DisPol	Distance to polygon if ignition point is outside polygon
PerCheck	True(1)/False(0) if the candidate lightning is within the fire perimeter

2.5 Processing of the agency reference datasets

To ensure a consistent assessment between the agency reference dataset and the BoLtFire dataset, the agency reference datasets were filtered by size, start date, and ignition cause. Prescribed burns were removed. The fire perimeters and points from the agency references were initially joined based on their respective fire ID keys (Canadian National Fire Database: NFDBFIREID and CFS_REF_ID; Alaska Fire Service: ID and FIREID). Polygons without points were removed from the dataset. Points without corresponding perimeter IDs were categorized into two groups: those located within an existing fire perimeter and occurring within seven days of the fire start date and those that do not. For the latter (as not all fire parameters have been mapped yet; Canadian Forest Service, 2024), the fire area was used to determine the fire radius, which was then used to generate a fire buffer perimeter. For perimeters without corresponding point IDs, if there was an unmatched point within the perimeter within seven days of the fire start date, that became the fire's ignition point. If there were no points located within the perimeter, the fire was left without an ignition point and was removed from the dataset. Both datasets were then merged

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

and filtered based on the majority land cover class. The filtering process yielded an agency reference dataset consisting of 2,424 fires across Alaska and Canada.

2.6 BoLtFire dataset creation method

Existing methodologies for fire ignition detection require an ignition point for each fire, which is not provided by the original MCD64A1 nor the GlobFire dataset. To address this, a Temporal Minimum Distance (TMin) methodology is proposed. This approach searches for a candidate lightning within the fire's perimeter starting from its ignition date and searches until a candidate lightning is found or a 14-day window has been reached. The first candidate lightning found within the perimeter is designated as the candidate lightning and the ignition point. If multiple potential candidate lightnings are found, the one closest in time to the ignition date is chosen. If no candidate lightning is found within the perimeter, the search extends to a 10 km radius outside the perimeter, following the same temporal process. If multiple potential candidate lightning are found, the stroke closest both spatially to the perimeter and temporally to the start date is selected as the ignition point (Larjavaara et al., 2005). A visualization of the process can be seen in Fig. 1. The proposed TMin methodology was then applied to the filtered ENTLN and GlobFire datasets to create the BoLtFire dataset.

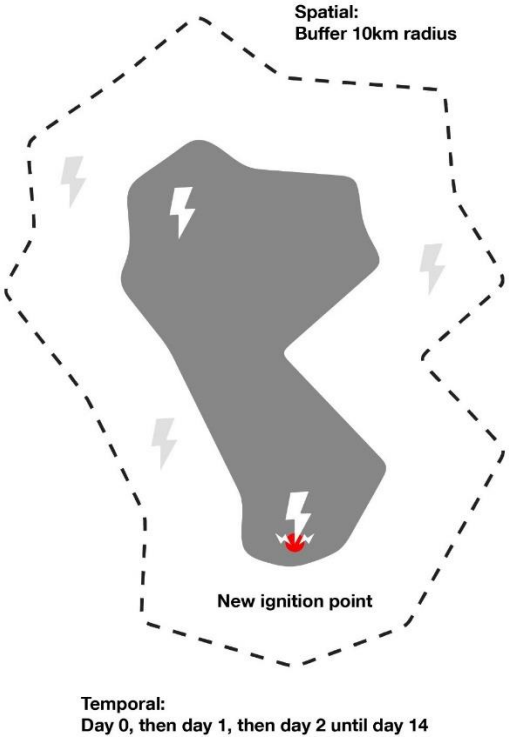

Figure 1: Visualization of the temporal minimum distance process, where lightning candidates are first searched for within the fire perimeter before then searching the buffered area if one is not found.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024


Open Access Earth System Science Data Discussions

2.7 Agency reference comparison and analysis

We conducted three different comparison assessments using the proposed TMin approach: (1) a candidate lightning methodology comparison, (2) a spatial and temporal accuracy comparison, and (3) fire count and size comparison. The matching lightning methodology comparison was implemented to gauge the performance of the proposed (TMin) methodology to currently established methodologies (DMin and MaxA). To align our methodologies to the current literature, all three methodologies were implemented using input parameters of tmax of 14 days and a Smax of 10 km. The methodologies were applied to the agency reference dataset as both datasets provide a start date and an ignition point, allowing for a more precise comparison of the candidate lightning's distance to ignition location and holdover time across the methodologies. The following variables were analysed: candidate lightning agreement, fire count, fire size, holdover time, and distance to the ignition point. The mean, median, and histograms of distance and holdover time were also produced. Additionally, a detailed analysis of the location of the candidate lightning in or outside (and distance there-from) the fire perimeter was conducted, to gauge the approach's ability to accurately locate the ignition point. We conducted a secondary benchmark comparison, where we compared the temporal and spatial locations of the fire perimeters to determine the overlap and difference between both datasets (agency vs North America (NA) BoLtFire). The NA BoLtFire dataset successfully located a fire if it was within seven days (before or after the start date) and 10 km of the agency's fire perimeter. A large spatial and temporal window was used to encompass reporting errors by both the agencies and errors from MODIS. The BoLtFire dataset successfully located a fire if it was within seven days (before or after the start date) and 10 km (Fusco et al., 2019) of the agency's fire perimeter. Multiple NA BoLtFire fires could be matched to one agency fire. If a NA BoLtFire fire was not within this spatial and temporal window, then it was unsuccessful. A confusion matrix was used to evaluate the performance of the dataset. A third and final comparison was conducted in line with that by Artés et al. (2019), where the total fire count and fire size between the NA BoLtFire and the agency reference datasets were compared.

3 Results

3.1 Matching lightning methodology comparison

3.1.1 Candidate lightning agreement

The TMin method slightly outperformed the MaxA and DMin methods, matching 74.71 % of the total agency reference dataset fires while DMin matched 71.49 %, and MaxA matched 66.67 % (Table 2). Each methodology selected different candidate lightnings, with general agreement between all three methodologies at 35.97 %. DMin and MaxA selected 55.86 % of the same overlapping lightning. There was a general decreasing trend in identical candidate lightnings as fire size increased. For 21.82 % of the fires, all three methodologies failed to find a matching lightning candidate; and 42.20 % of the matched fires had

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

different or no candidate lightning. Breaking down candidate lightning agreement by fire size, the DMin method shows the highest candidate lightning matching for small fires (74.25 %), but its performance declines significantly for larger fires, particularly for XLarge and Mega fires, where it drops to 57.14 % and 64.00 %, respectively. Conversely, the TMin method shows improvement with larger fires, achieving its highest matching for Mega fires at 92.00 %. The matching accuracy for Moderate and Large fires using the TMin method is also noteworthy at 74.54 % and 78.59 %, respectively. MaxA, while less accurate overall, shows a similar trend to DMin with a decreasing accuracy to increasing fire size, particularly for XLarge fires where it drops to 52.38 %.

Table 2: Total count and total burned area in hectares (from the agency reference dataset) of the candidate lightnings found by the three methodologies: TMin, MaxA and DMin. The total count and burned area as percentages of the total count and total burned ha of the agency reference dataset arealso included for all methodologies. Significant results are in bold.

	Agency Reference Dataset		North America BoLtFire dataset											
Fire Size	LIW Count	Total Ha	TMin				MaxA				DMin			
			LIW Count	% Total	Total Burned Area (Ha)	% Ha Total	LIW Count	% Total	Total Burned Area (Ha)	% Ha Total	LIW Count	% Total	Total Burned Area (Ha)	% Ha Total
Small	932	460,119.82	680	72.96 %	332,957.17	72.36 %	646	69.31 %	314,454.29	68.34 %	692	**74.25 %**	336,010.78	**73.03 %**
Moderate	1,084	3,768,131.15	808	**74.54 %**	2,823,847.99	**74.94 %**	724	66.79 %	2,503,849.55	66.45 %	778	71.77 %	2,658,545.57	70.55 %
Large	341	7,096,947.69	268	**78.59 %**	5,433,741.73	**76.56 %**	208	61.00 %	4,120,002.18	58.05 %	223	65.40 %	4,487,389.33	63.23 %
Xlarge	42	2,972,699.20	32	**76.19 %**	2,235,446.92	**75.20 %**	22	52.38 %	1,528,884.62	51.43 %	24	57.14 %	1,681,481.77	56.56 %
Mega	25	5,359,454.62	23	**92.00 %**	5,002,774.83	**93.34 %**	16	64.00 %	3,309,407.70	61.75 %	16	64.00 %	3,309,407.70	61.75 %
Total	2,424	19,657,352.47	1,811	**74.71 %**	15,828,768.65	**80.52 %**	1,616	**66.67 %**	11,776,598.34	59.91 %	1,733	**71.49 %**	12,472,835.15	63.45 %

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

3.1.2 Fire count and size

For Small fires, the TMin method matches more lightning candidates (72.36 % by total burned area) compared to MaxA (68.34 % by ha), with DMin slightly outperforming TMin (Table 2). For all other fire sizes, TMin has a similar advantage (74.94 %) over MaxA (66.45 %) and DMin (70.55 %). For Large fires, TMin continues to perform better (76.56 %) than MaxA (58.05 %) and DMin (63.23 %), showing a significant drop in accuracy for MaxA. XLarge fires exhibit a consistent pattern with TMin at 75.20 %, ahead of MaxA (51.43 %) and DMin (56.56 %), demonstrating the lowest accuracy for MaxA. Mega fires also see the highest accuracy for TMin (93.34 %), outperforming both MaxA (61.75 %) and DMin (61.75 %).

3.1.3 Holdover times

In all three methods, the holdover time showed an exponential decrease over time (Figure 2). Over 50 % of candidate lightnings among the three methodologies (TMin: 55.27 %, MaxA: 57.49 %, and DMin: 64.80 %) were found within the first 3 days (Day +1/0-2; Table B1); over 88 % within the first 10 days (TMin: 88.74 %, MaxA: 91.34 %, and DMin: 92.96 %). The largest frequencies for the TMin, MaxA, and DMin were found on Day 0 (503, 449, and 546 respectively). These results align with the overall trend that most wildfires occur within the first few days following a lightning event, with the majority occurring within two weeks. The TMin method shows the highest median and mean holdover times (2.20 and 3.97 days, respectively; Table B2), with overlapping confidence intervals for both medians and means when compared to MaxA, which has slightly lower values (2.12 and 3.64 days). The DMin method, while showing the shortest median and mean holdover times (1.84 and 3.05 days), does not have overlapping confidence intervals with the other methods.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

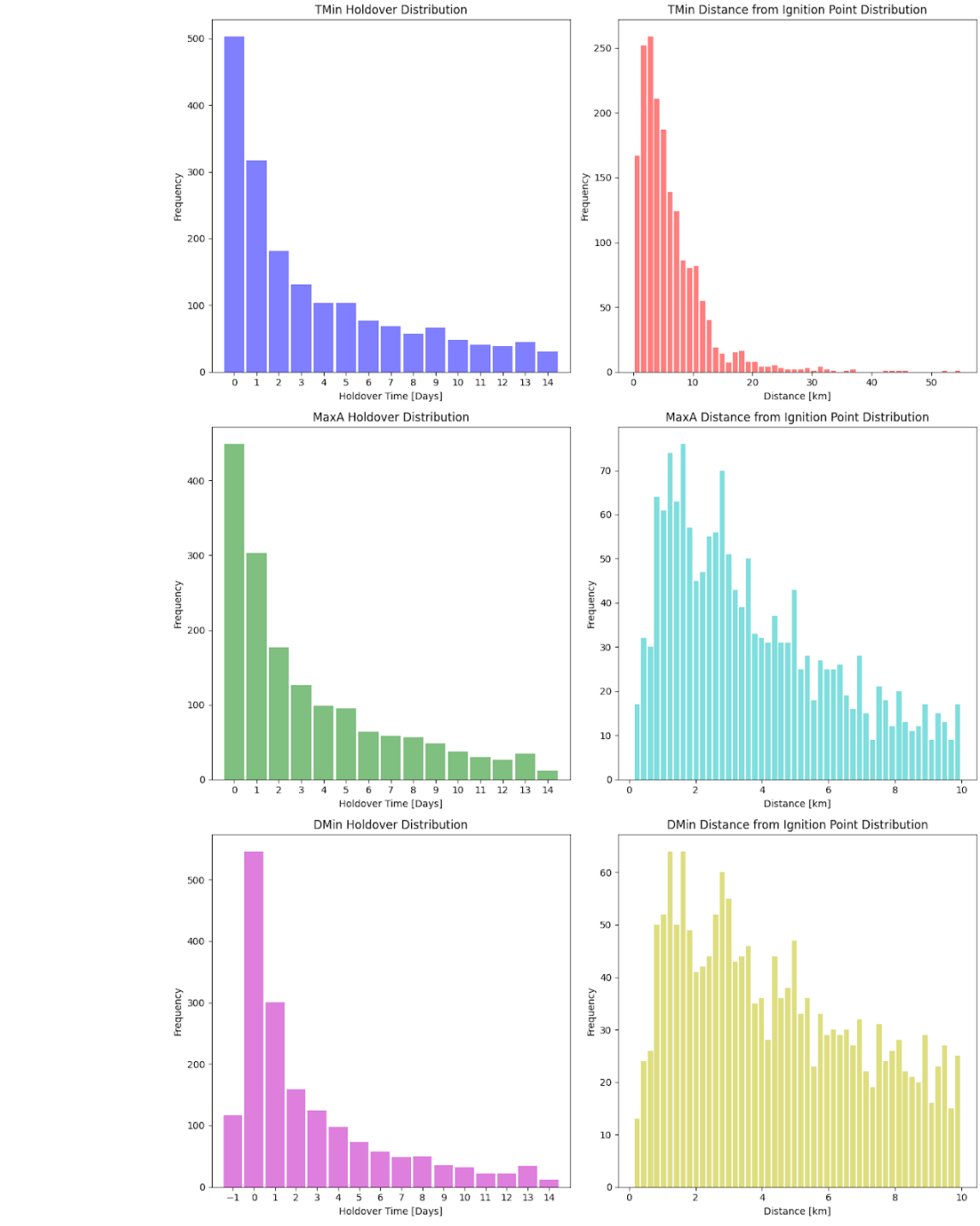

Figure 2: Distribution of lightning candidates by holdover time and distance to the ground truth ignition point. Holdover time is calculated as the difference between the start date of each lightning-ignited wildfire and the time of the occurrence of the candidate lightning. Distance is calculated as the distance of the lightning candidate from the ignition point of each lightning-ignited wildfire.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

3.1.4 Distance to the ignition point

The median distance of the selected candidate lightnings from the ground truth ignition point varied greatly for all three methodologies. Both the MaxA and DMin methods identify candidate lightnings with shorter distances from the ignition point compared to the TMin method. The TMin has a median of 4,592.00 meters and a mean of 6,110.32. MaxA identified the lightning candidates closest to the ignition point with a median distance of 3,155.20 meters and a mean distance of 3,771.49 meters (Table B2). DMin also identifies candidate lightnings with a shorter distance than TMin, with a median of 3,968.77 meters and a mean of 4,391.07 meters. These results were expected as both the MaxA and the DMin approaches rely on the ignition point in order to select the candidate lightning while the TMin looks to select the lightning candidate first within the fire perimeter, prioritizing candidates that are only closer temporally to the start date, before then checking outside the perimeter.

Within the agency reference dataset, 246 of the fires had ignition points located outside of the fire perimeter, and 2,178 within the fire perimeter (where the ignition of the fire should most likely occur). All three methodologies show varying levels of effectiveness in selecting a lightning candidate point inside or outside the ignition perimeter. The TMin methodology offers a balanced performance with a moderate overall accuracy of 32.29 %, making it the most reliable across both inside and outside the perimeter (Table B3). Both the MaxA and DMin methodologies did well correctly locating points outside the fire perimeter, but both struggled to locate them inside the fire perimeter leading to an overall accuracy of 26.11 % and 24.50 % respectively. Overall, while all three struggle to accurately determine if the ignition point is inside or outside the perimeter, TMin appears to be the most consistent with a higher overall accuracy.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

3.2 Evaluation of TMin and the BoLtFire dataset

3.2.1 Confusion matrix results

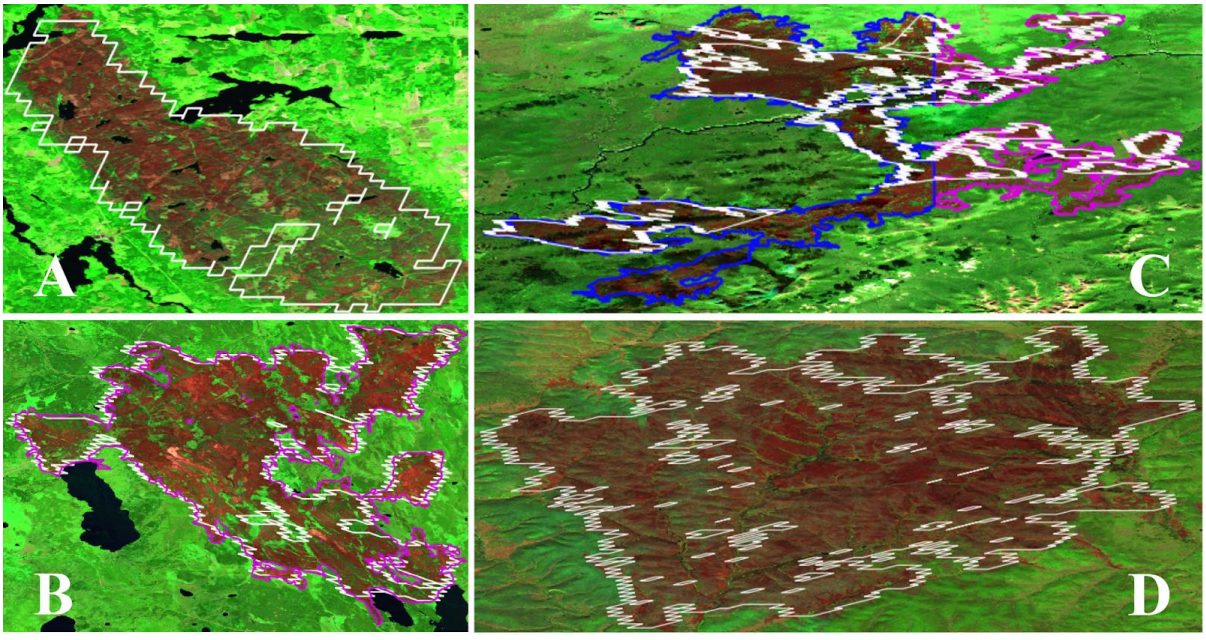

Figure 3: Within this figure there are four different fires: A, B, C, D. Fire A: BoLtFire 16389528 ignited on 26 July 2014 and was extinguished on 12 August 2014. It burned 11,699.48 ha and was located in Sweden. Fire B: Visual comparison of matched agency reference dataset fire 15BN-MACK (magenta) and BoLtFire 17455090 (blue). The lightning-ignited wildfire, 15BN-MACK, occurred in Saskatchewan, Canada on 26 June 2015 and lasted until 29 July 2015 burning roughly 38,249.32 ha. Its perimeter was mapped and created using Aerial GPS. Fire C is unique as occurred on the border between both Alaska and Canada. On the Alaska side, Fire C shows two fires (blue), fire 39800 started on 26 June 2017 and burned 37,846.33 ha and fire 39832 started on 03 July 2017 and burned 31,912.67 ha. On the Canadian side (magenta), fire 2017OC010 started on 22 June 2017 and was 97102.29 ha. BoLtFire 19856819 shows these fires as one fire that started on 29 July 2017 but was only 67,932.6 ha in size. Fire D: BoLtFire 24442738 occurred on 22 August 2021 in Russia and burned 77,645.45 ha. Base imagery is Landsat 8 ©USGS.

The overall accuracy was 63.04 %, commission error was 30.06 % and the omission error was 53.63 %, indicating that approx. 70% of the fires detected in BoLtFire were correctly classified, but that approx. 54 % of all occurring fires were missed in BoLtFire (Table C1). The BoLtFire database is thus conservative in comparison to the agency datasets. Within the agency reference dataset, of the 2,424 lightning-ignited wildfires, 1,124 wildfires had a BoLtFire dataset match while 1,300 remained unmatched. Of the 2,042 lightning-ignited wildfires in the BoLtFire dataset for North America, 1,559 lightning-ignited wildfires were matched while 483 remained unmatched. Overall, the matched fires accounted for 46.37 % of the total fires in the agency reference dataset but covered 70.15 % of the total burned area from the agency reference dataset. An example of agency fires and its matched BoLtFire can be seen in Fig. 3B and Fig. 3C. Within the agency reference dataset, fire count and size fluctuated drastically across the years, with a recorded total of 2,424 fires, burning approximately 19.66 million ha (Table C2). The year 2015 recorded the highest number of fires, with 514 incidents burning approximately 4.53 million ha, while

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

2021 witnessed the lowest number of fires, with only seven incidents, burning only around 30,259.57 ha. In terms of matched fires, 2017 had the highest match rate at 186 fires (55.19 % of the 2017 agency reference dataset total lightning-ignited wildfire count), and 2019 had the lowest, with just 30 of the 158 agency reference dataset lightning-ignited wildfires, representing 18.99 % of the total lightning-ignited wildfire count for that year.

One noticeable trend was that as fire size increased, both the amount of total fires as well as total burned area generally improved. Small fires have the lowest matching success, as only 24.36 % of the total Small fires successfully matched the agency reference dataset (Table 3), with a delta of -71.53 % of the total burned area matched. Moderate and Large fires show moderate success, with about 55-70 % of fires and burned areas matched, while XLarge and Mega fires exhibit a strong matching performance, with 76.19 % of XLarge fires matched with 73.76 % of the total burned area and 80.00 % of the Mega fires matched with 82.30 % of the total burned area. The 10 unmatched XLarge and 5 unmatched Mega fires accounted for approximately 1.73 million unmatched hectares, representing 29.45 % of the total unmatched burned area. When Small and Moderate fires were removed, 287 of the remaining 408 fires were matched, increasing the overall matching from 46.37 % to 70.34 % of total fires, and increasing the matched total burn area from 70.15 % to 74.04 %. Overall, we found that the matching process is more effective for larger fires, both in terms of the number of fires and the total burned area matched.

Table 3: Compares the total number of fires and the total burned area by fire size category between the agency reference dataset (ARD) and matched fires in the North American BoLtFire dataset, using the total burned area for both from the agenda reference dataset.

	Agency Reference Dataset		Matched Fires in the NA BoLtFire Dataset				Delta (Agency Reference Dataset to North America BoLtFire)			
Fire Size	**Total LIW**	**Total Burned Area (Ha)**	**Total LIW**	**Percent of Total LIW**	**Total Burned Area from ARD (Ha)**	**Percent of Total Burned Area from ARD (Ha)**	**Total LIW**	**Percent Total LIW**	**Total Burned Area from ARD (Ha)**	**Percent Total Burned Area (Ha)**
Small	932	460,119.82	227	24.36 %	130,977.77	28.47 %	-705	-75.64 %	-329,142.05	-71.53 %
Moderate	1,084	3,768,131.15	610	56.27 %	2,233,730.82	59.28 %	-474	-43.73 %	-1,534,400.33	-40.72 %
Large	341	7,096,947.69	235	68.91 %	4,820,762.69	67.93 %	-106	-31.09 %	-2,276,185.00	-32.07 %
Xlarge	42	2,972,699.20	32	76.19 %	2,192,740.13	73.76 %	-10	-23.81 %	-779,959.07	-26.24 %
Mega	25	5,359,454.62	20	80.00 %	4,410,940.26	82.30 %	-5	-20.00 %	-948,514.36	-17.70 %
Total	2,424	19,657,352.47	1,124	46.37 %	13,789,151.67	70.15%	-1,300	-53.63 %	- 5,868,200.80	-29.85 %

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

3.2.2 Fire count and fire size comparison

The agency reference dataset reports 2,424 fires, burning approximately 19.66 million ha, while the NA BoLtFire dataset records 2,042 lightning-ignited wildfires with a total burned area of approximately 12.01 million ha (Table 4). Small fires were more frequently recorded in the NA BoLtFire dataset (1,003 fires) compared to the agency reference dataset (932 fires), though the difference in the burned area for Small fires was relatively minor. However, the agency reference dataset records a higher number of Moderate (delta of 287), Large fires (delta of 137), XLarge (delta of 21) and Mega (delta of 8) fires, burning significantly more hectares than in the NA BoLtFire dataset. Moderate fires also show a significant difference, with the agency reference dataset reporting 287 more incidents and a 30.14 % larger total burned area. The most notable differences between the two datasets can be seen in the larger fire categories, with the agency reference dataset showing significantly more Large, XLarge, and Mega fires, resulting in a 38 % to 53 % larger burned area (40.79 %, 53.40 %, 38.17 % respectively). Overall, the agency reference dataset reflects a higher total count and larger burned area, particularly in the larger fire size categories, while the NA BoLtFire dataset is slightly higher in small fires but substantially lower across all other fire sizes.

Table 4: Total count and total burned area of both the agency reference dataset and the North America BoLtFire Dataset from 2012-2022 organized by fire size. The delta was calculated between the Agency Reference Dataset and the North America BoLtFire Dataset.

	Agency Reference Dataset		North America BoLtFire Dataset		Delta (Agency Reference Dataset to North America BoLtFire)			
	LIW Count	Total Burned Area (Ha)	LIW Count	Total Burned Area (Ha)	LIW Count	Percent LIW Count	Total Burned Area (Ha)	Percent Total Burned Area (Ha)
Small	932	460,119.82	1003	481,707.10	-71	-7.62 %	-21,587.28	-4.69 %
Moderate	1,084	3,768,131.15	797	2,632,230.98	287	26.48 %	1,135,900.17	30.14 %
Large	341	7,096,947.69	204	4,202,232.89	137	40.18 %	2,894,714.80	40.79 %
XLarge	42	2,972,699.20	21	1,385,300.87	21	50.00 %	1,587,398.33	53.40 %
Mega	25	5,359,454.62	17	3,313,507.60	8	32.00 %	2,045,947.02	38.17 %
Total	2,424	19,657,352.47	2042	12,014,979.44	382	15.76 %	7,642,373.03	38.88 %

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

380

3.3 BoLtFire dataset overview

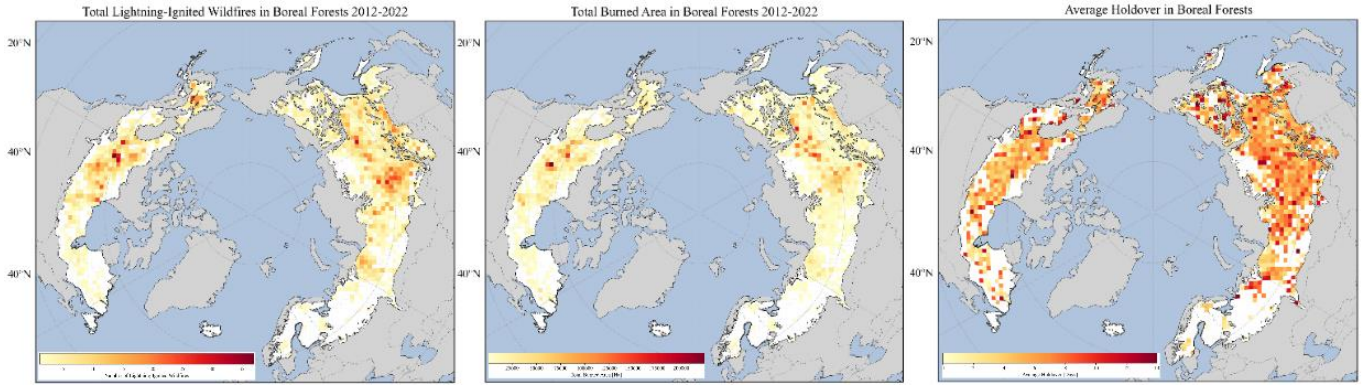

Figure 5: Visual depiction of all located in forest and forest-like land covers in boreal forests from 2012-2022 that are equal to or greater than 200 ha.

385 Of the original 31,363 total filtered GlobFire events, 6,228 were identified through the TMin methodology (within 14 days and 10 km of the GlobFire event) to be lightning-ignited (19.86 %). The BoLtFire dataset encompasses lightning-ignited wildfires located in forest and forest-like land covers in boreal forests from 2012-2022 that are at least 200 ha in size. Of the 6,228 fires, 4,186 were located in Eurasia and 2,042 were located in North America (Figure 5). Within those 4,184 fires identified as lightning-ignited wildfires within Eurasia, 4,137 are located in Russia, 40 are in Mongolia, 8 are in Sweden and
390 1 is located in Finland. Of the 2,042 in North America BoLtFire, 1,759 are located in Canada and 283 are located in Alaska, United States. The BoLtFire dataset encompasses a total of approximately 33.09 million ha of burned area (21.07 million ha in Eurasia and 12.01 million ha in North America) with an overall average burned area per fire of 5,312.84 ha. The information included for each fire within the dataset can be found in Table 1.

3.4 Characterization of lightning-ignited wildfires

395 We summarized the following metrics for each continent: holdover time, distance to fire perimeter, lightning-ignition efficiency, frequency of fires per year and total burned area, area burned by fire size class, and total burned area per land cover.

3.4.1 Holdover

Similar to the agency reference dataset, the holdover time within the BoLtFire dataset exhibited an exponential decline, with the highest frequency occurring within the first 24 hours (Figure 4). Specifically, 711 lightning-ignited wildfires (11.42 %)
400 were identified on the first day, and 1,963 lightning-ignited wildfires (31.52 %) ignited within the first three days (Day 0 to Day 2; Table D1). The cumulative number of lightning-ignited wildfires increased quickly, with nearly half of all fires recorded within 5 days (Day 0 to Day 4) and 78.23 % occurring within the first 10 days.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

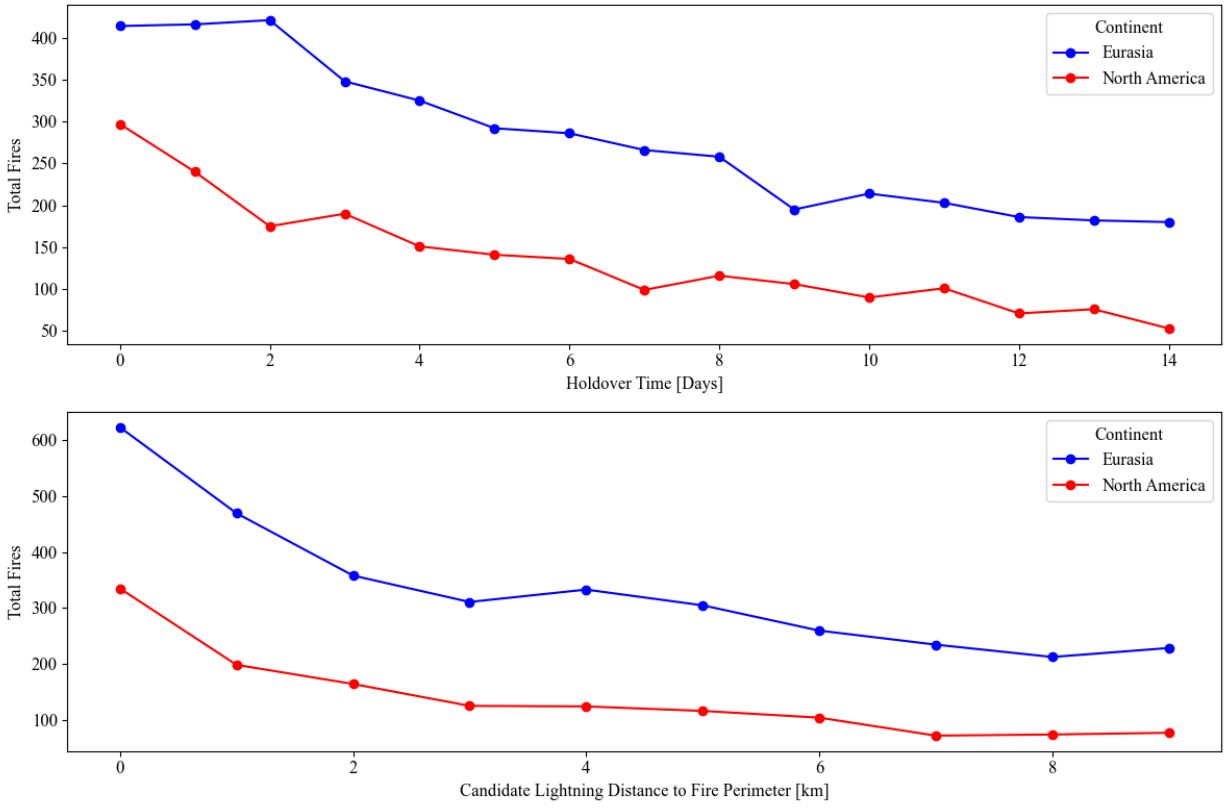

Figure 4: Top graph compares the holdover time for the total number of fires over a 15-day period between the Eurasia and North America BoLtFire dataset continents. Bottom graph compares the distance of the candidate lightning from the fire perimeter for all of the fires whose candidate lighting was located outside the fire perimeter within the Eurasia and North America BoLtFire dataset continents.

3.4.2 Distance to fire perimeter

The analysis of the distance of the ignition point from the perimeter of the fire revealed a pattern similar to that of the holdover, a decrease in total fires as the distance from the perimeter increases (Figure 4). A total of 1,495 of the ignition points were located within the perimeter of the fire. The remaining 4,733 ignition points were located outside of the perimeter, 174 were within 100 m of the fire perimeter, 957 (20.22 %) were within 1 km and over half (2585, 54.62 %) were within 4 km (Table D2). This trend was mirrored in both Eurasia and North America.

3.4.3 Lightning-ignition efficiency

The BoLtFire dataset reveals variations in the incidence of lightning ignition efficiency (LIE) across both continents (Table D3). Eurasia recorded a higher number of total strokes (approximately 15 million) and a greater number of lightning-ignited wildfires (4,186), compared to North America's 12.53 million strokes and 2,042 lightning-ignited wildfires. Approximately 0.0279 of lightning strokes resulted in ignition in Eurasia, while just 0.0163 of strokes led to an ignition in North America.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Overall, the combined data across both continents shows a total of nearly 27.53 million strokes, resulting in 6,228 lightning-ignited wildfires, with an overall lightning-ignition efficiency of 0.0226. To better understand the lightning-ignition efficiency at a more regional level, it was calculated by country: The United States (Alaska) has the highest lightning-ignition efficiency at 0.5079. Mongolia has the next highest efficiency at 0.075 with Russia following with 0.029. Canada has a much lower lightning-ignition efficiency of 0.0141, while Sweden and Finland have the lowest lightning-ignition efficiencies at 0.0037 and 0.0003, respectively. This data suggests that the likelihood of lightning igniting wildfires varies considerably across these countries for larger fires.

3.4.4 Frequency of fires per year and total burned area

Understanding the number of fires per year is critical for assessing the frequency and trends of wildfire occurrences. The total frequency of lightning-ignited wildfires exhibited substantial year-to-year variability across both continents, with the most amount of lightning-ignited wildfires found in 2013 (923) and the least found in 2019 (104) (Table D4). From a total burned area perspective, even though 2013 had the most lightning-ignited wildfires, it burned the third most amount of area at close to 4.31 million ha; while 2021 had the most burned area at 6.07 million ha and 2019 had the least at 832,307 total ha burned. In 2012, the dataset shows that Eurasia had no lightning-ignited wildfires, while North America had 219. This discrepancy is most likely due to the ENTLN having fewer sensors available at the time, as only 244 total lightning strokes were found in that year in Eurasia (the average amount is approximately 1,363,542 per year). In 2013, despite Eurasia having more lightning-ignited wildfires, North America experienced a significantly larger burned area (2.62 million ha compared to 1.69 million ha in Eurasia).

3.4.5 Area burned by fire size class

Within the BoLtFire dataset, Small fires are the most common, with 3,299 LIW, burning a total of approximately 1.56 million ha, resulting in an average of 141,428.26 ha per year and 471.57 ha per fire (Table D5 and D6). Moderate fires follow with 2,344 lightning-ignited wildfires, contributing to a total burned area of about 7.45 million ha, and an average of 677,225.82 ha per year and 3,178.11 ha per fire. Although Large fires are less frequent, with 484 incidents, they account for a significant burned area of approximately 10.15 million ha, averaging 923,008.77 ha per year and 20,977.47 ha per fire. XLarge and Mega fires are the least common (52 and 49 incidents respectively), but they contribute disproportionately to the total burned area. XLarge fires burned 3.56 million ha, with an average of 323,826.30 ha per year and 68,501.72 ha per fire, while Mega fires had the largest average burned area at 942,543.65 ha per year and 211,591.43 ha per fire, totaling approximately 10.37 million ha. When reviewing all fire sizes, North America has an average burned area of 5,883.93 ha per fire, which is relatively close to Eurasia's 5,034.25 ha per fire. When averaging the smaller fire sizes (Small and Moderate fires), North America has a higher average burned area of 1,729.97 ha per fire, compared to Eurasia's 1,532.98 ha per fire. When looking at just the larger fire sizes (Large, XLarge, and Mega), Eurasia has a significantly larger average burned area of 44,262.75 ha compared to North

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

America's 36,781.16 ha burned area. Overall, North American Small and Moderate fires tend to be larger than Eurasia's, but Eurasia tends to have larger Large, XLarge, and Mega fires.

3.4.6 Total burned area per land cover

Within the BoLtFire dataset, Woody Savannas and Savannas are the most affected forest or forest-like land cover types, with 2,861 and 2,051 fires respectively, contributing to the largest total burned areas of approximately 15.81 million ha and 13.86 million ha; 89.69 % of the total burned area (Table D7). Evergreen Needleleaf Forests also had an impact, with 418 fires burning approximately 1.47 million ha, accounting for 4.45 % of the total burn area. Open Shrublands and Mixed Forests experienced moderate impacts, accounting for 3.58 % and 1.44 % of the total burn area, respectively. Unsurprisingly, Savannas had the largest average burn area per fire at 6,759.60 ha, followed closely by Woody Savannas at 5,526.64 ha, Evergreen Needleleaf Forests at 3,520.17 ha, and Open Shrublands at 3,076.51. Deciduous Broadleaf Forests, Permanent Wetlands, and Deciduous Needleleaf Forests had the fewest total burned ha, and some of the smallest average burned area per ha. The total burned area was driven strongly by the Woody Savannas and Savannas land covers and this trend was also mirrored in both Eurasia and North America.

When all other land covers besides forest-specific ones (Deciduous Broadleaf Forests, Deciduous Needleleaf Forests, Evergreen Needleleaf Forests, and Mixed Forests) were removed, Eurasia, contained a total of 579 fires, contributing to a combined fire area of 1,108,490.08 ha, with an overall average fire size of 1,914.49 ha. In North America, there were 176 fires, with a total area burned of 940,773.94 hectares and an overall average fire size of 5,345.31. While Eurasia had a larger overall combined burn area and more fires, North America experienced larger fires on average. North America had on average larger Small (460.74 ha vs 484.77 ha, respectively), Moderate (2,921.58 ha vs 3,550.06 ha), and Large (17,323.29 ha vs 21,621.91 ha) fires. Eurasia did not contain any forest land cover-based XLarge fires, while North America had five, totaling 321,079.76 ha with an average fire size of 64,216.95 ha. While Eurasia experienced a significantly higher number of fires, North America had fewer but larger fires on average, particularly in the extreme fire size categories. This suggests potential differences in fire management or environmental conditions influencing fire behavior between the two continents.

4 Discussion

4.1 Agency reference dataset and NA BoLtFire dataset comparison

The candidate lightning agreement highlighted the important methodological choices when matching lightning candidates (Moris et al., 2020). Although all three methods performed relatively well, finding candidate lightnings for at least 66 % of the agency reference dataset, they had difficulty agreeing on the "correct" lightning candidate, with only 35.97 % agreement across all three methodologies. This discrepancy is likely due to the differences in methodological approaches, as MaxA and DMin, who have similar methodologies, agreed on 55.86 % of candidates, while TMin agreed with MaxA on only 41.50 % and DMin

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

on 37.13 %. Overall, the TMin matched the most lightning candidates (74.71 %); while DMin and MaxA matched slightly fewer at 71.49 % and 66.67 %, respectively. However, significant differences became clear in the total burned area matched: out of approximately 19.66 million ha in the agency reference dataset, TMin identified candidate lightning for approximately 15.83 million ha, whereas DMin identified only 12.47 million ha (3.36 million ha less than TMin) and MaxA identified only 11.78 million ha (4.05 million ha less than TMin).

While TMin matched more large fire sizes, MaxA identified the lightning candidates closest to the ignition point but correctly located only 22.68 % within the fire perimeter. While the agency reference dataset did have 246 ignition locations outside the fire perimeter, 89.85 % of ignitions occurred within the fire perimeter. Of the 2,178 ignitions inside the fire perimeter, TMin had the highest accuracy, correctly identifying 719 of those ignitions within the perimeter (33.01 % accuracy). This low accuracy could be attributed to several factors, one of which could be due to the TMin methodology, though it originally looks for only candidate lightnings within the perimeter; if one is not found within a certain window (14 days), then it starts to look outside the perimeter. The holdover for these fires could be longer than 14 days, as holdovers in boreal forests can last for a significant period of time before flaming (Scholten et al., 2021). Additionally, it could be due to location error of the lightning locations, as quite a large portion of the boreal forest is remotely located. Furthermore, as noted by Hanes et al. (2019), the Canadian National Fire Database points are just the presumed points of ignition, which may also explain why 246 of the ignition locations were not within their corresponding fire polygons. Furthermore, as Crowley et al. (2023) highlighted, both points and perimeter polygons in the Canadian National Fire Database are compiled and submitted by different provinces, potentially adding additional discrepancies and leading to variations in definitions across the datasets.

The results of the confusion matrix comparing the temporal and spatial overlap of the agency reference dataset and the NA BoLtFire dataset indicated that our proposed TMin methodology was promising; with low commission and moderate omission errors of 30.06 % and 53.63 %, respectively. Interestingly, Small and Moderate fires matched the fewest fires, and when they were removed, the overall matching increased from 46.37 % to 70.34 % of total fires; indicating that fires that are less than 10,000 ha are more difficult to match than larger fires. This may be due to the difficulty MODIS has in detecting smaller wildfires (discussed further in Sect. 4.3, Limitations). Another possible explanation could be the size of the matching temporal window. In their study, Fusco et al. (2019) used a seven-day window when spatially matching their MODIS fires, which is what we also implemented. After applying this window, we had five unmatched Mega fires which we then reviewed visually to better understand why they were unmatched. We noticed that three of the five had fire perimeters within the agency reference dataset that matched well visually, but all had start dates outside of seven days, but within 30 days of the agency reference fire. Due to the remoteness of some of these fires and possible fire recognition issues with the underlying MODIS data, an increase in the window size could include more fires. This increase in window size could also cause fires to be incorrectly associated with neighboring fires due to the increase in window size. Overall errors of omission are likely due to the detailed limitations mentioned below, as well as differences in pixel resolution, satellite overpass time, and cloud and smoke cover (Giglio et al., 2009; Hantson et al., 2013; Hawbaker et al., 2008; Johnston et al., 2018; Roy et al., 2005). Errors of commission are likely due to fire-free surfaces that are highly reflective (Cardoso et al., 2005; Giglio et al., 2009; Hantson et al., 2013); though less likely

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

in this study as both urban and agricultural land uses were removed. As we removed fires that were smaller than 200 ha, our errors of omission and commission are likely to be lower than expected.

4.2 BoLtFire dataset

From a global perspective, the BoLtFire dataset provided interesting insights that could help enhance our understanding of holdover time, lightning-ignition efficiency, primary characteristics, frequency, and spatial distribution of lightning-ignited wildfires in boreal forests. The holdover phenomenon poses a significant challenge to the real-time detection of lightning-ignited wildfires. While our results generally align with current literature, as the frequency of the holdover decreased over time, it differs in its initial detection, as the majority of lightning-ignited wildfires are typically detected within the first 24 hours (Gao et al., 2024; Moris et al., 2023). Our lightning-ignited wildfire holdovers resulted in just 11.42 % occurring within the first 24 hours, with 47.80 % within a five-day window (Day 0 to Day 4) days. This could be due to a variety of factors, including the global scale of our results (most studies are done at a local or more regional level). Additionally, lightning-ignited wildfire ignition is strongly influenced by fuel type, weather, and topography; holdover times could reflect this dependency.

The analysis of the distance of the ignition point from the perimeter of the fire also revealed a decrease in total fires as the distance from the perimeter increased. While the agency reference dataset did have ignition points outside the fire perimeter, these accounted for 10.15 % of the total dataset. Within the BoLtFire dataset, 76.00 % of fires have an ignition point outside the fire perimeter. This could be due to a couple of different reasons, either the ignition point was associated with the wrong fire perimeter (i.e., the correct fire perimeter was not identified by MODIS, or was removed in the filtering process), the correct lightning candidate within the fire perimeter was not within the ENTLN dataset, or possibly, it could be due to spatial location errors from either, or both, the underlying MODIS dataset or the ENTLN dataset. Benali et al. (2016) found that there could be up to a 12 hr temporal and 2 km spatial lag between data reported by agencies and data that was derived by a satellite. We think it is most likely both, as within the 4,733 ignition points located outside the perimeter, 20.22 % were located within 1 km of the perimeter, 33.33 % were within 2 km, and over half were within 4 km.

While thunderstorms generate thousands of strikes a year, the likelihood of ignition occurring from a lightning strike is low, as only a few manage to persist long enough to enter a flaming stage (Wotton and Martell, 2005; Pineda and Rodríguez, 2023; Podur et al., 2003). Latham and Williams (2001) found that within much of North America, 0.01 to 0.04 actually ignites a fire; on average about 0.00167 caused an ignition in Alaska (Wendler et al., 2010), 0.001329 in Alberta and Saskatchewan (Nash and Johnson, 1996), 0.02 in British Columbia and 0.00071 in Alberta (Wierzchowski et al., 2002). Our lightning-ignition efficiency across North America and Eurasia varies from that found in the literature, though likely due to the regionality of the other studies and to our filtering process. When reviewed more closely, the lightning-ignition efficiency of our lightning-ignited wildfires in Canada were roughly 0.000141 and Alaska's lightning-ignition efficiency was roughly 0.005079, or 3/595. Though Canada's lightning-ignition efficiency rate is not in line with the literature, Alaska's is. This could be due to the 200

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

ha filter, which would have removed a significant amount of smaller fires. Canada's lightning-ignition efficiency is still not in line with more local studies. This could still be due to the filtering, but could also indicate that lightning-ignition efficiency needs to be conducted at a more local and regional level. Furthermore, upon visual inspection, some lightning data was located within a country but was not labeled as such due to imperfections within the World Bank (2020) Official Boundaries dataset. This could lead to a higher lightning-ignited efficiency than what was reported.

There was no distinct increase or decrease in fire count and size over the years, with both variables experiencing high year-to-year variation. As mentioned, weather and other climatic factors play a significant role in lightning-ignited wildfire ignition; which could help explain the general variation between, and within, the total count and burned area in fire seasons. However, one of the most prevalent trends in the BoLtFire dataset shows that larger fires, despite the smaller overall fire count, are disproportionately responsible for the largest portion of the total burned area; whereas Small and Moderate fires represent 90.61 % of the total fire count but account for just 27.22 % of the total burned area. Conversely, Large, XLarge, and Mega fires make up 9.39 % of the total fires but are responsible for 72.28 % of the total burned area. This trend is also observed in the agency reference dataset, Small and Moderate fires constitute 83.17 % of the total fire count but only contribute to 21.51 % of the total burned area. In contrast, Large, XLarge, and Mega fires, although comprising only 16.83 % of the total fires, are responsible for 78.49 % of the total burned area. These results are in line with current literature (Grünig et al, 2022; Hanes et al., 2019; Stocks et al., 2002). Interestingly, when reviewing larger fire sizes (Large, XLarge, and Mega) between Eurasia and North America, Eurasian fires are on average 16.90 % larger than fires in North America. This seems to be due to the dominant land cover type underlying the fires. The overwhelming majority of the burned area occurred within either the Savannas or Woody Savanna (accounting for a combined 89.69 % of the total burned land cover). When only reviewing forest-specific land covers, while Eurasia still had a larger total burned area, North America had on average fire sizes that were 64.18 % bigger.

4.3 Limitations

While our overall results were promising, there are several uncertainties and limitations that need to be considered when working with our data. First, it is important to emphasize that the comparison between the MODIS-based dataset and the agency reference dataset should be viewed as an assessment rather than a validation. The production of similar values between the datasets enhanced our confidence in the reliability of our resulting BoLtFire datasets. However, interpreting discrepancies between the results is complex due to a multitude of influencing factors. Second, our current understanding of global fire regimes relies heavily on satellite-based products, MODIS active fire and burned area products are one of the most widely used (e.g., Crowley et al., 2023). The efficacy of MODIS products, along with other satellite-based fire products, is constrained by limitations in resolution and sensitivity, as demonstrated by comparative validation with higher-resolution sensors like Landsat. Within the U.S., Hawbaker et al. (2008) found that MODIS active fire product detected 82 % of all Landsat reference fires, but their detection efficacy decreased with fire size and just 50 % of fires were detected at 105 ha. The mean detected

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

fire sizes were 915 and 1,044 ha, which could heavily impact the frequency and size of fires found within BoLtFire datasets. When reviewing small (> 0.2 ha) sub-canopy fires in Canada, Johnston et al. (2018) found that coarser spatial resolution sensors would be incapable of early fire detection for 44 - 70 % of historic wildfires in boreal forests in Canada. Additionally, Loboda et al. (2011) found that MODIS burned area products can underestimate the extent of fires by 15 - 70 %. This could be an indication that the overall total count and fire size of the BoLtFire dataset are lower than expected. Furthermore, there are limitations within the agency reference datasets that affect the overall commission and omission errors. Some fires that were identified by MODIS might not be identified by an agency due to the remoteness of their location (Fusco et al., 2019), incomplete lightning data or incorrect occurrence date (Flannigan and Wotton, 1991), or misidentified as lightning-ignited due to a nearby thunderstorm activities (Müller et al., 2013). However, the former reason is unlikely due to the methods and experience of those investigating the fire (Schultz et al., 2019). Additionally, fires that are mapped later in the season could have fire scars that are insensitive to multiple fire events, making it difficult to distinguish between singular fire events and fire complexes. Yet, this is also an issue for agency databases as they can also fail to distinguish between the two (Benali et al., 2016). Third, since we used a lightning detection network, there is a possibility that not all lightning strikes were detected. As mentioned, in more remote areas, detection efficiency can be relatively low, and location errors are expected to be larger (up to 10 km for the ENTLN). Additionally, as the ENTLN currently has no sensors located in Russia, this could create larger spatial errors or possibly not detect some lightning strikes. Since our lightning matching methodology heavily depends on the detection of lightning strikes, these factors could influence not only the likelihood of a correct match but also the selection of the candidate lightning strike in the absence of the "correct" match. Finally, as the currently available reference datasets are limited to boreal forests, we were only able to assess the BoLtFire dataset with those resulting fires within Canada and Alaska. Fire records from other agencies, especially those in Siberia and Russia are difficult to obtain or might be inaccurate (Stocks et al., 2001). As tree canopies can obscure fires (Johnston et al., 2018; Kolden et al., 2012), and as surface fires are more prevalent in Eurasia than the crown fires dominant in North America, omission rates could be higher in Eurasia. Additionally, Talucci et al. (2022) found when comparing MODIS to Landsat fire detection in Siberia from 2001-2020, Landsat captured 47.9 % more burn area, which could indicate that our datasets do not fully reflect the true burn area in Siberia, and thus, within Eurasia.

4.4 Outlook

The effectiveness of lightning in igniting wildfires is expected to be amplified by climate change; under the RCP8.5 scenario, Hessilt et al. (2022) predict that changes in fire weather and vegetation will increase the effectiveness of lightning ignition by 31 ± 28 % in Canada's Northwest Territories and 14 ± 9 % in Alaska per 1°C of warming. Understanding the characteristics of these fires is crucial for improving our knowledge of where, why, and how they ignite, which could enhance our ability to model and mitigate their occurrence in the future. Furthermore, research by Zhu et al. (2017) suggests that the effects of Eurasian boreal wildfire emissions on Arctic warming may be underestimated. Given that lightning is a predominant cause of

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

wildfires in boreal forests, improving the identification of lightning-ignited fires will enhance our understanding of their emissions and their contribution to global warming.

5 Data availability

All input data used in this dataset (Engle et al., 2024) was publicly available (with the exception of the lightning data, as this data was provided by Earth Networks, Inc., an AEM company) and can be found in Table A3. Data described in this manuscript can be accessed at Zenodo under data doi 10.5281/zenodo.13897163.

6 Code availability

The code used in this dataset is available at https://github.com/BrittanyEngle/BoLtFire_Code.

7 Conclusion

Challenges in identifying lightning-ignited wildfires, coupled with limitations in data availability, have hindered our understanding of the characteristics of these wildfires in boreal forests. In this paper, we introduced the Temporal Minimum Distance (TMin) methodology, a novel approach specifically developed to match lightning strikes without an ignition point. It outperformed current methods, allowing us to create the largest pan-boreal forest dataset. This groundbreaking dataset consists of 6,228 lightning-ignited wildfires located in forest and forest-like land covers from 2012-2022 that are at least 200 ha. When benchmarked to the agency reference dataset, this new dataset performed reasonably well, with an overall commission error of 30.06 % and omission error of 53.63 %; though further fine-tuning of the TMin methodology and input parameters could lead to an even better performance. This dataset can be further used in conjunction with climate and additional environmental data to help better model lightning-ignited wildfire ignition characteristics and provide additional insights into lightning-ignited wildfires in boreal forests.

Acknowledgements

CS received funding from the European Space Agency through the Climate-Space RECCAP2 project (4000144908/24/I-LR).

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

Appendix A

Table A1: Descriptive list of which land cover types were either forest or forest-like and thus labeled "Forest". Those labeled "Forest" were included within the dataset while those that were not, were removed.

Forest	Non-Forest
• Evergreen Needleleaf Forests • Evergreen Broadleaf Forests • Deciduous Needleleaf Forests • Deciduous Broadleaf Forests • Mixed Forests • Closed Shrublands • Open Shrublands • Woody Savannas • Savannas • Grasslands • Permanent Wetlands • Permanent Snow and Ice	• Urban and Built-up Lands • Cropland/Natural Vegetation Mosaics • Croplands • Water Bodies • Barren

Table A2: Description of the total burned area per fire size class.

Fire Size Class	Size in HA
Small	200 ≤ 1,000 ha
Moderate	1,000 ≤ 10,000 ha
Large	10,000 ≤ 50,000 ha
Extremely Large	50,000 ≤ 100,000 ha
Mega Fires	> 100,000 ha

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

Table A3: List of every dataset used to create the BoLtFire dataset, their spatial and temporal extent, citation, and link to their location.

Dataset	Spatial extent	Temporal extent	Citation	Notes	Link
Alaska Fire Service Fire Perimeters	Alaska, USA	1942-2023	(U.S. Department of the Interior, Bureau of Land Management, Alaska Fire Service, 2024b)	Almost all of the fire perimeters had a matching fire point	https://fire.ak.blm.gov/
Alaska Fire Service All Fire Points	Alaska, USA	1939-2022	(U.S. Department of the Interior, Bureau of Land Management, Alaska Fire Service, 2024a)	Meaning of point locations unclear, no fire perimeters	https://fire.ak.blm.gov/
Canadian National Fire Database - National Fire Database fire polygon data	Canada	1917-2020	(Canadian Forest Service, 2024)	Fire perimeters do not always have a matching fire point	https://cwfis.cfs.nrcan.gc.ca/ha/nfdb
Canadian National Fire Database - National Fire Database fire point data	Canada	1930-2022	(Canadian Forest Service, 2024)	Meaning of point locations differs across provinces, no fire perimeters	https://cwfis.cfs.nrcan.gc.ca/ha/nfdb
Ecoregions 2017	Global	2017	(Olson et al., 2001)	global classification of	https://ecoregions.appspot.com/

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

				terrestrial ecoregions	
GlobFire Fire Perimeters	Global, 500m	2002-2023	(Artés et al., 2019)	Global, vectorized MODIS burned area product	https://gwis.jrc.ec.europa.eu/apps/country.profile/downloads
World Bank Official Boundaries: Admin 0, 10 meter boundary Dataset	Global, 10m	2020	(World Bank, 2020)	Boundaries are 10 m, some imperfections along the borders found	https://datacatalog.worldbank.org/search/dataset/0038272/World-Bank-Official-Boundaries
Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 MCD64A1 burned area product	Global, 500m	2000-11-01 - Present	(Giglio et al., 2018)	Global fire dataset, raster	https://lpdaac.usgs.gov/products/mcd64a1v061/
MODIS MCD12Q1v061 Land Cover Type 1	Global, 500m	2001-01-01 - 2022-12-31	(Friedl and Sulla-Menashe, 2022)	Global land cover dataset, downloadable in small regions	https://lpdaac.usgs.gov/products/mcd12q1v061/
Global LANd Cover mapping and Estimation (GLANCE) Grids	Global	-	(Arevalo et al., 2022)	Global land cover that splits North America, Europe and Asia	https://measures-glance.github.io/glance-grids/

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

- Version 01 CRS					
Earth Networks Total Lightning Network (ENTLN)	Global	2012-Current	(Zhu et al., 2022)	Global dataset of lightning locations	https://www.earthnetworks.com/
Worldwide Lightning Location Network (WWLLN)	Global	15 August 2004 to the present	(Rodger et al., 2004)	Global dataset of lightning locations	https://wwlln.net/

Appendix B

Table B1: Cumulative percentage values of the holdover time distribution for all three methodologies within the BoLtFire Dataset, TMin, MaxA and DMin. Holdover time is calculated as the difference between the start date of each lightning-ignited wildfire and the time of the occurrence of the candidate lightning.

	Holdover time for BoLtFire dataset								
	TMin			**MaxA**			**DMin**		
Day	Per Day	Cumulative	Cumulative Percent	Per Day	Cumulative	Cumulative Percent	Per Day	Cumulative	Cumulative Percent
+1							117	117	6.75 %
0	503	503	27.77 %	449	449	27.78 %	546	663	38.26 %
1	317	820	45.28 %	303	752	46.53 %	301	964	55.63 %
2	181	1,001	**55.27 %**	177	929	**57.49 %**	159	1,123	**64.80 %**
3	131	1,132	62.51 %	126	1,055	65.28 %	125	1,248	72.01 %
4	103	1,235	68.19 %	99	1,154	71.41 %	98	1,346	77.67 %

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

5	103	1,338	73.88 %	95	1,249	77.29 %	73	1,419	81.88 %
6	77	1,415	**78.13 %**	64	1,313	**81.25 %**	58	1,477	**85.23 %**
7	69	1,484	81.94 %	58	1,371	84.84 %	49	1,526	88.06 %
8	57	1,541	85.09 %	57	1,428	88.37 %	50	1,576	90.94 %
9	66	1,607	**88.74 %**	48	1,476	**91.34 %**	35	1,611	**92.96 %**
10	48	1,655	91.39 %	37	1,513	93.63 %	32	1,643	94.81 %
11	41	1,696	93.65 %	30	1,543	95.48 %	22	1,665	96.08 %
12	39	1,735	95.80 %	26	1,569	97.09 %	22	1,687	97.35 %
13	45	1,780	98.29 %	35	1,604	99.26 %	34	1,721	99.31 %
14	31	1,811	**100.00 %**	12	1,616	**100.00 %**	12	1,733	**100.00 %**

Table B2: Median and mean holdover time, rounded to the second decimal for the TMin, MaxA and DMin lighting matching methodologies. Holdover time is calculated as the difference between the start date of each lightning-ignited wildfire and the time of the occurrence of the candidate lightning. Median and mean distance from the agency reference dataset ignition point locations to each of the TMin, MaxA, and DMin's candidate lightning.

Method	Holdover time for North America BoLtFire dataset		Distance from ignition point for the North America BoLtFire dataset	
	Median	Mean	Median (meters)	Mean (meters)
TMin	2.20	3.97	4,592.00	6,110.32
MaxA	2.12	3.64	3,155.20	3,771.49
DMin	1.84	3.05	3,968.77	4,391.07

Table B3: Comparison of ignition point location (inside or outside the perimeter) of the agency reference dataset with that predicted by the TMin, MaxA, and DMin methodologies.

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

660

Location of LIW	Agency Reference Dataset	North America BoLtFire dataset					
		TMin		MaxA		DMin	
0 (Outside fire Perimeter)	246						
Incorrectly Matched		126	51.22 %	31	12.60 %	24	9.76 %
Correctly Matched		63	25.61 %	139	56.50 %	154	62.60 %
Not Matched		57	23.17 %	76	30.89 %	68	27.64 %
1 (Inside fire Perimeter)	2,178						
Incorrectly Matched		903	41.46 %	952	43.71 %	1115	51.19 %
Correctly Matched		719	33.01 %	494	22.68 %	440	20.20 %
Not Matched		556	25.53 %	732	33.61 %	623	28.60 %
Overall Accuracy	**2,424**	**782**	**32.26 %**	**633**	**26.11 %**	**594**	**24.50 %**

Appendix C

Table C1: The table shows the Confusion Matrix for the Agency Reference Dataset and the North America BoLtFire dataset.

		Predicted		
		Positive	**Negative**	
Actual	**Positive**	1124	1300	0.4637
	Negative	483	NA	NA
		0.6994	NA	0.6304

665

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

Table C2: The table compares the total number of fires and total burned area in hectares from the agency reference dataset to the matched fires from 2012 to 2022, focusing on the percentage of fires and burned areas successfully matched to the BoLtFire dataset each year.

	Agency Reference Dataset		Matched LIW in the NA BoLtFire Dataset			
Year	**Total LIW**	**Total Burned Area from ARD (Ha)**	**Total LIW**	**% of Total ARD LIW**	**Total Burned Area from ARD (Ha)**	**% of Total Burned Area of ARD LIW (Ha)**
2012	277	1,194,093.07	129	46.57 %	932,268.24	78.07 %
2013	358	4,509,476.07	181	50.56 %	3,767,148.46	83.54 %
2014	260	3,681,170.78	133	51.15 %	2,850,437.03	77.43 %
2015	514	4,533,388.97	242	47.08 %	2,750,248.49	60.67 %
2016	185	790,941.98	74	40.00 %	570,483.72	72.13 %
2017	337	1,911,371.68	186	55.19 %	1,344,255.95	70.33 %
2018	194	770,780.33	96	49.48 %	545,100.90	70.72 %
2019	158	1,285,196.45	30	18.99 %	387,674.26	30.16 %
2020	62	76,127.56	12	19.35 %	25,499.31	33.50 %
2021	7	30,259.57	2	28.57 %	7,102.05	23.47 %
2022	72	874,546.02	39	54.17 %	608,933.25	69.63 %
Total	2,424	19,657,352.47	1,124	46.37 %	13,789,151.67	70.15 %

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Open Access Earth System Science Data Discussions

Appendix D

Table D1: Distribution of holdover time across the BoLtFire dataset. Holdover time refers to the time difference between when the fire was ignited by the lightning-ignited candidate and when it was detected. The detection time used was the startdate.

North America BoLtFire dataset			
Holdover Days	**Count per Day**	**Cumulative Count**	**Cumulative % Total**
0	711	711	11.42 %
1	656	1,367	21.95 %
2	596	1,963	31.52 %
3	538	2,501	40.16 %
4	476	2,977	47.80 %
5	433	3,410	54.75 %
6	422	3,832	61.53 %
7	365	4,197	67.39 %
8	374	4,571	73.39 %
9	301	4,872	78.23 %
10	304	5,176	83.11 %
11	304	5,480	87.99 %
12	257	5,737	92.12 %
13	258	5,995	96.26 %
14	233	6,228	100.00 %

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

680 **Table D2: Distribution of candidate lightning ignition points relative to their corresponding fire perimeters, categorized by distance in kilometres. It shows the number of ignition points at each distance, along with their cumulative counts and cumulative percentages.**

North America BoLtFire dataset			
Distance	**Count per Day**	**Cumulative Count**	**Cumulative % Total**
1 km	957	957	20.22 %
2 km	668	1,625	34.33 %
3 km	523	2,148	45.38 %
4 km	437	2,585	54.62 %
5 km	458	3,043	64.29 %
6 km	422	3,465	73.21 %
7 km	365	3,830	80.92 %
8 km	308	4,138	87.43 %
9 km	288	4,426	93.51 %
10 km	307	4,733	100.00 %

Table D3: Total count of lightning-ignited wildfires and total count of ENTLN strokes per region. Lightning-ignition efficiency was calculated as the percent of total strokes that caused an ignition.

		BoLtFire dataset	
Type	**Total ENTLN Strokes**	**Total LIW**	**LIE**
North America	12,526,781	2,042	0.0163
Eurasia	14,998,966	4,186	0.0279
Total	27,525,747	6,228	0.0226

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

685

Table D4: Total count of lightning-ignited wildfires and total burned area per continent across the study period

	Total BoLtFire Dataset		Eurasia BoLtFire Dataset		North America BoLtFire Dataset	
Fire Year	**LIW Count**	**Total Burned Area (Ha)**	**LIW Count**	**Total Burned Area (Ha)**	**LIW Count**	**Total Burned Area (Ha)**
2012	219	948,303.52	0	0	219	948,303.52
2013	923	4,305,664.56	570	1,686,514.47	353	2,619,150.10
2014	822	4,848,885.08	524	2,569,868.70	298	2,279,016.38
2015	883	4,047,524.75	425	1,489,829.85	458	2,557,694.91
2016	544	2,147,570.12	406	1,655,154.47	138	492,415.64
2017	871	3,447,717.05	600	1,980,088.18	271	1,467,628.86
2018	513	2,779,725.83	384	2,180,896.38	129	598,829.45
2019	104	832,307.40	27	54,756.72	77	777,550.68
2020	585	2,312,353.08	562	2,218,195.59	23	94,157.49
2021	404	6,074,648.61	401	6,068,468.83	3	6,179.78
2022	360	1,343,660.87	287	1,169,608.24	73	17,4052.63
Total	6,228	33,088,360.88	4,186	21,073,381.44	2042	12,014,979.44

690

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

695 **Table D5: Overall view of both the Eurasia and North America BoLtFire datasets by total lightning-ignited wildfire count, toral burned area, and average burned area, within the BoLtFire dataset by fire size.**

	Total BoLtFire Dataset			Eurasia BoLtFire Dataset			North America BoLtFire Dataset		
Fire Size	**LIW Count**	**Total Burned Area (Ha)**	**Average Burned Area per Fire (Ha)**	**LIW Count**	**Total Burned Area (Ha)**	**Average Burned Area per Fire (Ha)**	**LIW Count**	**Total Burned Area (Ha)**	**Average Burned Area per Fire (Ha)**
Small	3,299	1,555,710.84	471.57	2,296	1,074,003.74	467.77	1,003	481,707.10	480.27
Moderate	2,344	7,449,484.07	3,178.11	1,547	4,817,253.09	3,113.93	797	2,632,230.98	3,302.67
Large	484	10,153,096.47	20,977.47	280	5,950,863.57	21,253.08	204	4,202,232.89	20,599.18
Xlarge	52	3,562,089.32	68,501.72	31	2,176,788.45	70,218.98	21	1,385,300.87	65,966.71
Mega	49	10,367,980.18	211,591.43	32	7,054,472.58	220,452.27	17	3,313,507.60	19,4912.21
Total	6,228	33,088,360.88	5,312.84	4,186	21,073,381.44	5,034.25	2,042	12,014,979.44	5,883.93

Table D6: Total count and average fire size of fires in evergreen needleleaf forests

Fire Size	LIW Count	Total Burned Area (Ha)	Average Burned Area Per Year
Small	3,299	1,555,710.84	141,428.26
Moderate	2,344	7,449,484.07	677,225.82
Large	484	10,153,096.47	923,008.77
Mega	49	10,367,980.18	942,543.65
Xlarge	52	3,562,089.32	323,826.30
Grand Total	6,228	33,088,360.88	3,008,032.81

700

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Table D7: Total count and average fire size of fires in evergreen needleleaf forests

BoLtFire Dataset			
Land Cover Type	**LIW Count**	**Total Burned Area (Ha)**	**Average Burned Area per Fire (Ha)**
Eurasia	4,186	21,073,381.44	5,034.25
Deciduous Broadleaf Forests	4	3,578.54	894.63
Deciduous Needleleaf Forests	71	90,233.37	1,270.89
Evergreen Needleleaf Forests	259	551,094.26	2,127.78
Grasslands	138	145,285.19	1,052.79
Mixed Forests	245	463,583.91	1,892.18
Open Shrublands	330	1,091,449.61	3,307.42
Permanent Wetlands	8	2,763.93	345.49
Savannas	1,306	9,583,546.13	7,338.09
Woody Savannas	1,825	9,141,846.49	5,009.23
North America	2,042	12,014,979.44	5,883.93
Deciduous Broadleaf Forests	2	884.62	442.31
Deciduous Needleleaf Forests	7	5288.50	755.50
Evergreen Needleleaf Forests	159	920,338.84	5,788.29
Grasslands	13	19,643.94	1,511.07
Mixed Forests	8	14,261.98	1,782.75
Open Shrublands	55	93,007.79	1,691.05

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Permanent Wetlands	17	11,285.43	663.85
Savannas	745	4,280,388.01	5,745.49
Woody Savannas	1,036	6,669,880.33	6,438.11
Total	**6,228**	**33,088,360.88**	**5,312.84**

Author Contributions

BE developed the methodology and workflow with IB and CS. Supervision and oversight was provided by CS. BE and IB developed the code, curated and validated the data and created the visualizations. BE wrote the original draft, review and editing of the final draft was conducted by CS, IB, MC, and YZ.

Competing interests

The contact author has declared that none of the authors has any competing interests.

References

Anderson, K. (2002). A model to predict lightning-caused fire occurrences. International Journal of Wildland Fire, 11(4), 163. doi:10.1071/wf02001

Arevalo, P., Stanimirova, R., Bullock, E., Zhang, Y., Tarrio, K., Turlej, K., Hu, K., McAvoy, K., Pasquarella, V., Woodcock, C., Olofsson, P., Zhu, Z., Gorelick, N., Loveland, T., Barber, C., Friedl, M. (2022). *Global Land Cover Mapping and Estimation Yearly 30 m V001* [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. Accessed 2024-09-21 from https://doi.org/10.5067/MEaSUREs/GLanCE/GLanCE30.001

Benali, A., Russo, A., Sá, A., Pinto, R., Price, O., Koutsias, N., & Pereira, J. (2016). Determining Fire Dates and Locating Ignition Points With Satellite Data. Remote Sensing, 8(4), 326. doi:10.3390/rs8040326

Bond, T. C., Doherty, S. J., Fahey, D. W., Forster, P. M., Berntsen, T., DeAngelo, B. J., … Zender, C. S. (2013). Bounding the role of black carbon in the climate system: A scientific assessment. Journal of Geophysical Research: Atmospheres, 118(11), 5380–5552. doi:10.1002/jgrd.50171

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Brandt, J. P. (2009). The extent of the North American boreal zone. Environmental Reviews, 17(NA), 101–161. doi:10.1139/a09-004

Brandt, J. P., Flannigan, M. D., Maynard, D. G., Thompson, I. D., & Volney, W. J. A. (2013). An introduction to Canada's boreal zone: ecosystem processes, health, sustainability, and environmental issues. Environmental Reviews, 21(4), 207–226. doi:10.1139/er-2013-0040

Bui, V., Chang, L.-C., & Heckman, S. (2015, December). A Performance Study of Earth Networks Total Lighting Network (ENTLN) and Worldwide Lightning Location Network (WWLLN). 2015 International Conference on Computational Science and Computational Intelligence (CSCI). doi:10.1109/csci.2015.120

Canadian Forest Service. (2024). Canadian National Fire Database – Agency Fire Data [Data set]. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta. http://cwfis.cfs.nrcan.gc.ca/en_CA/nfdb

Cardoso, M. F., Hurtt, G. C., Moore, B., Nobre, C. A., & Bain, H. (2005). Field work and statistical analyses for enhanced interpretation of satellite fire data. Remote Sensing of Environment, 96(2), 212–227. doi:10.1016/j.rse.2005.02.008

Crowley, M. A., Stockdale, C. A., Johnston, J. M., Wulder, M. A., Liu, T., McCarty, J. L., Rieb, J. T., Cardille, J. A., & White, J. C. (2023). Towards a whole-system framework for wildfire monitoring using Earth observations. Global Change Biology, 29, 1423–1436. https://doi.org/10.1111/gcb.16567

de Groot, W. J., Flannigan, M. D., & Cantin, A. S. (2013). Climate change impacts on future boreal fire regimes. Forest Ecology and Management, 294, 35–44. doi:10.1016/j.foreco.2012.09.027

Descals, A., Gaveau, D. L. A., Verger, A., Sheil, D., Naito, D., & Peñuelas, J. (2022). Unprecedented fire activity above the Arctic Circle linked to rising temperatures. Science, 378(6619), 532–537. doi:10.1126/science.abn9768

Engle, B., Bratoev, I., Crowley, M. A., Zhu, Y., & Senf, C. (2024). Distribution and Characteristics of Lightning-Ignited Wildfires in Boreal Forests - the BoLtFire database (1.0) [Dataset]. https://doi.org/10.5281/zenodo.13897163

Friedl, M., Sulla-Menashe, D. (2022). MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. Accessed 2024-09-20 from https://doi.org/10.5067/MODIS/MCD12Q1.061

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Flannigan, M. D., Stocks, B. J., & Wotton, B. M. (2000). Climate change and forest fires. Science of The Total Environment, 262(3), 221–229. doi:10.1016/s0048-9697(00)00524-6

Flannigan, M. D., Stocks, B. J., Turetsky, M. R., & Wotton, M. (2009). Impacts of climate change on fire activity and fire management in the circumboreal forest. Global Change Biology, 15(3), 549–560. doi:10.1111/j.1365-2486.2008.01660.x

Flannigan, M. D., & Wotton, B. M. (1991). Lightning-ignited forest fires in northwestern Ontario. Canadian Journal of Forest Research, 21(3), 277–287. doi:10.1139/x91-035

Fusco, E. J., Finn, J. T., Abatzoglou, J. T., Balch, J. K., Dadashi, S., & Bradley, B. A. (2019). Detection rates and biases of fire observations from MODIS and agency reports in the conterminous United States. Remote Sensing of Environment, 220, 30–40. doi:10.1016/j.rse.2018.10.028

Gao, C., Shi, C., Li, J., Yuan, S., Huang, X., Zhang, Q., … Wu, G. (2024). Igniting lightning, wildfire occurrence, and precipitation in the boreal forest of northeast China. Agricultural and Forest Meteorology, 354, 110081. doi:10.1016/j.agrformet.2024.110081

Giglio, L., Boschetti, L., Roy, D., Hoffmann, A., Humber, M., & Hall, J. (2020). Collection 6 MODIS Burned Area Product User's Guide Version 1.3. National Aeronautics and Space Administration (NASA). https://lpdaac.usgs.gov/documents/875/MCD64_User_Guide_V6.pdf

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., & Justice, C. O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. Remote Sensing of Environment, 217, 72–85. doi:10.1016/j.rse.2018.08.005

Giglio, L., Loboda, T., Roy, D. P., Quayle, B., & Justice, C. O. (2009). An active-fire based burned area mapping algorithm for the MODIS sensor. Remote Sensing of Environment, 113(2), 408–420. doi:10.1016/j.rse.2008.10.006

Gillett, N. P., Weaver, A. J., Zwiers, F. W., & Flannigan, M. D. (2004). Detecting the effect of climate change on Canadian forest fires. Geophysical Research Letters, 31(18). doi:10.1029/2004gl020876

Hanes, C. C., Wang, X., Jain, P., Parisien, M.-A., Little, J. M., & Flannigan, M. D. (2019). Fire-regime changes in Canada over the last half century. Canadian Journal of Forest Research, 49(3), 256–269. doi:10.1139/cjfr-2018-0293

Hansen, J., & Nazarenko, L. (2003). Soot climate forcing via snow and ice albedos. Proceedings of the National Academy of Sciences, 101(2), 423–428. doi:10.1073/pnas.2237157100

Hantson, S., Padilla, M., Corti, D., & Chuvieco, E. (2013). Strengths and weaknesses of MODIS hotspots to characterize global fire occurrence. Remote Sensing of Environment, 131, 152–159. doi:10.1016/j.rse.2012.12.004

Hawbaker, T. J., Radeloff, V. C., Syphard, A. D., Zhu, Z., & Stewart, S. I. (2008). Detection rates of the MODIS active fire product in the United States. Remote Sensing of Environment, 112(5), 2656–2664. doi:10.1016/j.rse.2007.12.008

Hessilt, T. D., Abatzoglou, J. T., Chen, Y., Randerson, J. T., Scholten, R. C., van der Werf, G., & Veraverbeke, S. (2022). Future increases in lightning ignition efficiency and wildfire occurrence expected from drier fuels in boreal forest ecosystems of western North America. Environmental Research Letters, 17(5), 054008. doi:10.1088/1748-9326/ac6311

Janssen, T. A. J., Jones, M. W., Finney, D., van der Werf, G. R., van Wees, D., Xu, W., & Veraverbeke, S. (2023). Extratropical forests increasingly at risk due to lightning fires. Nature Geoscience, 16(12), 1136–1144. doi:10.1038/s41561-023-01322-z

Jones, M. W., Abatzoglou, J. T., Veraverbeke, S., Andela, N., Lasslop, G., Forkel, M., … Le Quéré, C. (2022). Global and Regional Trends and Drivers of Fire Under Climate Change. Reviews of Geophysics, 60(3). doi:10.1029/2020rg000726

Johnston, J. M., Johnston, L. M., Wooster, M. J., Brookes, A., McFayden, C., & Cantin, A. S. (2018). Satellite Detection Limitations of Sub-Canopy Smouldering Wildfires in the North American Boreal Forest. Fire, 1(2). doi:10.3390/fire1020028

Kasischke, E. S., Williams, D., & Barry, D. (2002). Analysis of the patterns of large fires in the boreal forest region of Alaska. International Journal of Wildland Fire, 11(2), 131. doi:10.1071/wf02023

Kim, Y., Hatsushika, H., Muskett, R. R., & Yamazaki, K. (2005). Possible effect of boreal wildfire soot on Arctic sea ice and Alaska glaciers. Atmospheric Environment, 39(19), 3513–3520. doi:10.1016/j.atmosenv.2005.02.050

Kolden, C. A., Lutz, J. A., Key, C. H., Kane, J. T., & van Wagtendonk, J. W. (2012). Mapped versus actual burned area within wildfire perimeters: Characterizing the unburned. Forest Ecology and Management, 286, 38–47. doi:10.1016/j.foreco.2012.08.020

Kourtz, P. H., Todd, J. B., & Others. (1991). Predicting the daily occurrence of lightning-caused forest fires (Vol. 112). Petawawa National Forestry Institute.

Krause, A., Kloster, S., Wilkenskjeld, S., & Paeth, H. (2014). The sensitivity of global wildfires to simulated past, present, and future lightning frequency. Journal of Geophysical Research: Biogeosciences, 119(3), 312–322. doi:10.1002/2013jg002502

Lapierre, J. L., Laughner, J. L., Geddes, J. A., Koshak, W. J., Cohen, R. C., & Pusede, S. E. (2020). Observing U.S. Regional Variability in Lightning NO2 Production Rates. Journal of Geophysical Research: Atmospheres, 125(5). doi:10.1029/2019jd031362

Larjavaara, M., Pennanen, J., & Tuomi, T. J. (2005). Lightning that ignites forest fires in Finland. Agricultural and Forest Meteorology, 132(3–4), 171–180. doi:10.1016/j.agrformet.2005.07.005

Latham, D., & Williams, E. (2001). Lightning and Forest Fires. In Forest Fires (pp. 375–418). doi:10.1016/b978-012386660-8/50013-1

Laurent, P., Mouillot, F., Yue, C. et al. FRY, a global database of fire patch functional traits derived from space-borne burned area products. Sci Data 5, 180132 (2018). https://doi.org/10.1038/sdata.2018.132

Lavoué, D., Liousse, C., Cachier, H., Stocks, B. J., & Goldammer, J. G. (2000). Modeling of carbonaceous particles emitted by boreal and temperate wildfires at northern latitudes. Journal of Geophysical Research: Atmospheres, 105(D22), 26871–26890. doi:10.1029/2000jd900180

Liu, C., & Heckman, S. (2011). The application of total lightning detection and cell tracking for severe weather prediction. 91st American Meteorological Society Annual Meeting, 1–10.

Lizundia-Loiola, J., Otón, G., Ramo, R. & Chuvieco, E. A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data. Remote Sens. Environ. 236, 111493 (2020).

Loboda, T. V., Hoy, E. E., Giglio, L., & Kasischke, E. S. (2011). Mapping burned area in Alaska using MODIS data: a data limitations-driven modification to the regional burned area algorithm. International Journal of Wildland Fire, 20(4), 487. doi:10.1071/wf10017

Macias Fauria, M., & Johnson, E. A. (2007). Climate and wildfires in the North American boreal forest. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1501), 2315–2327. doi:10.1098/rstb.2007.2202

Mallick, S., Rakov, V. A., Hill, J. D., Gamerota, W. R., Uman, M. A., Heckman, S., … Liu, C. (2013, October). Calibration of the ENTLN against rocket-triggered lightning data. 2013 International Symposium on Lightning Protection (XII SIPDA), 39–46. doi:10.1109/sipda.2013.6729186

Mallick, S., Rakov, V. A., Hill, J. D., Ngin, T., Gamerota, W. R., Pilkey, J. T., … Liu, C. (2015). Performance characteristics of the ENTLN evaluated using rocket-triggered lightning data. Electric Power Systems Research, 118, 15–28. doi:10.1016/j.epsr.2014.06.007

Martell, D. L., & Sun, H. (2008). The impact of fire suppression, vegetation, and weather on the area burned by lightning-caused forest fires in Ontario. Canadian Journal of Forest Research, 38(6), 1547–1563. doi:10.1139/x07-210

Melillo, J. M., Richmond, T. T., Yohe, G., & Others. (2014). Climate change impacts in the United States. Third National Climate Assessment, 52, 150–174.

Moris, J. V., Conedera, M., Nisi, L., Bernardi, M., Cesti, G., & Pezzatti, G. B. (2020). Lightning-caused fires in the Alps: Identifying the igniting strokes. Agricultural and Forest Meteorology, 290, 107990. doi:10.1016/j.agrformet.2020.107990

Müller, M. M., Vacik, H., Diendorfer, G., Arpaci, A., Formayer, H., & Gossow, H. (2012). Analysis of lightning-induced forest fires in Austria. Theoretical and Applied Climatology, 111(1–2), 183–193. doi:10.1007/s00704-012-0653-7

Nash, C. H., & Johnson, E. A. (1996). Synoptic climatology of lightning-caused forest fires in subalpine and boreal forests. Canadian Journal of Forest Research, 26(10), 1859–1874. doi:10.1139/x26-211

Natali, S. M., Watts, J. D., Rogers, B. M., Potter, S., Ludwig, S. M., Selbmann, A.-K., … Zona, D. (2019). Large loss of CO2 in winter observed across the northern permafrost region. Nature Climate Change, 9(11), 852–857. doi:10.1038/s41558-019-0592-8

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., … Kassem, K. R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth. BioScience, 51(11), 933. doi:10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., … Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. Science, 333(6045), 988–993. doi:10.1126/science.1201609

Pérez-Invernón, F. J., Moris, J. V., Gordillo-Vázquez, F. J., Füllekrug, M., Pezzatti, G. B., Conedera, M., … Huntrieser, H. (2023). On the Role of Continuing Currents in Lightning-Induced Fire Ignition. Journal of Geophysical Research: Atmospheres, 128(21). doi:10.1029/2023jd038891

Pineda, N., & Rodríguez, O. (2023). ERA5 Reanalysis of Environments Conducive to Lightning-Ignited Wildfires in Catalonia. Atmosphere, 14(6), 936. doi:10.3390/atmos14060936

Podur, J., Martell, D. L., & Csillag, F. (2003). Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998. Ecological Modelling, 164(1), 1–20. doi:10.1016/s0304-3800(02)00386-1

Post, E., Alley, R. B., Christensen, T. R., Macias-Fauria, M., Forbes, B. C., Gooseff, M. N., … Wang, M. (2019). The polar regions in a 2°C warmer world. Science Advances, 5(12). doi:10.1126/sciadv.aaw9883

Rodger, C. J., Brundell, J. B., Dowden, R. L., & Thomson, N. R. (2004). Location accuracy of long distance VLF lightning location network. Annales Geophysicae, 22(3), 747–758. doi:10.5194/angeo-22-747-2004

Rogers, B. M., Soja, A. J., Goulden, M. L., & Randerson, J. T. (2015). Influence of tree species on continental differences in boreal fires and climate feedbacks. Nature Geoscience, 8(3), 228–234. doi:10.1038/ngeo2352

Roy, D. P., Jin, Y., Lewis, P. E., & Justice, C. O. (2005). Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. Remote Sensing of Environment, 97(2), 137–162. doi:10.1016/j.rse.2005.04.007

Scholten, R. C., Jandt, R., Miller, E. A., Rogers, B. M., & Veraverbeke, S. (2021). Overwintering fires in boreal forests. Nature, 593(7859), 399–404. doi:10.1038/s41586-021-03437-y

Schultz, C. J., Nauslar, N. J., Wachter, J. B., Hain, C. R., & Bell, J. R. (2019). Spatial, Temporal and Electrical Characteristics of Lightning in Reported Lightning-Initiated Wildfire Events. Fire, 2(2), 18. doi:10.3390/fire2020018

Seidl, R., Honkaniemi, J., Aakala, T., Aleinikov, A., Angelstam, P., Bouchard, M., … Senf, C. (2020). Globally consistent climate sensitivity of natural disturbances across boreal and temperate forest ecosystems. Ecography, 43(7), 967–978. doi:10.1111/ecog.04995

Sofronov, M. A., Volokitina, A. V., & Shvidenko, A. Z. (1998). Wildland fires in the north of Central Siberia. The Commonwealth Forestry Review, 124–127.

Soja, A. J., Tchebakova, N. M., French, N. H. F., Flannigan, M. D., Shugart, H. H., Stocks, B. J., … Stackhouse, P. W. (2007). Climate-induced boreal forest change: Predictions versus current observations. Global and Planetary Change, 56(3–4), 274–296. doi:10.1016/j.gloplacha.2006.07.028

Stocks, B. J., Fosberg, M. A., Lynham, T. J., Mearns, L., Wotton, B. M., Yang, Q., … McKenney D. W. (1998). Climate Change and Forest Fire Potential in Russian and Canadian Boreal Forests. Climatic Change, 38(1), 1–13. doi:10.1023/a:1005306001055

Stocks, B. J., Mason, J. A., Todd, J. B., Bosch, E. M., Wotton, B. M., Amiro, B. D., … Skinner, W. R. (2002). Large forest fires in Canada, 1959–1997. Journal of Geophysical Research: Atmospheres, 107(D1). doi:10.1029/2001jd000484

Stocks, B. J., Wotton, B. M., Flannigan, M. D., Fosberg, M. A., Cahoon, D. R., & Goldammer, J. G. (2001). Boreal Forest Fire Regimes And Climate Change. In Remote Sensing and Climate Modeling: Synergies and Limitations (pp. 233–246). doi:10.1007/0-306-48149-9_10

Stohl, A. (2006). Characteristics of atmospheric transport into the Arctic troposphere. Journal of Geophysical Research: Atmospheres, 111(D11). doi:10.1029/2005jd006888

Talucci, A. C., Loranty, M. M., & Alexander, H. D. (2022). Siberian taiga and tundra fire regimes from 2001–2020. Environmental Research Letters, 17(2), 025001. doi:10.1088/1748-9326/ac3f07

Turner, M. G. (2010). Disturbance and landscape dynamics in a changing world. Ecology, 91(10), 2833–2849. doi:10.1890/10-0097.1

U.S. Department of the Interior, Bureau of Land Management (BLM), Alaska Fire Service. (2024a). BLM AFS Alaska Statewide fire History Locations 1939-2023 points, NAD83 [Dataset]. https://fire.ak.blm.gov/

U.S. Department of the Interior, Bureau of Land Management (BLM), Alaska Fire Service. (2024b). BLM AFS Alaska Statewide Fire History Perimeters 1940-2023 Polygons, AKAlbersNAD83 [Dataset]. https://fire.ak.blm.gov/

Vajda, A., Venäläinen, A., Suomi, I., Junila, P., & Mäkelä, H. M. (2013). Assessment of forest fire danger in a boreal forest environment: description and evaluation of the operational system applied in Finland. Meteorological Applications, 21(4), 879–887. doi:10.1002/met.1425

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Veraverbeke, S., Rogers, B. M., Goulden, M., Jandt, R. R., Miller, C. E., Wiggins, E. B., & Randerson, J. T. (2017). Lightning as a major driver of recent large fire years in North American boreal forests. Nature Climate Change, 7(7), 529–534. doi:10.1038/nclimate3329

Watts, J. D., Farina, M., Kimball, J. S., Schiferl, L. D., Liu, Z., Arndt, K. A., … Oechel, W. C. (2023). Carbon uptake in Eurasian boreal forests dominates the high-latitude net ecosystem carbon budget. Global Change Biology, 29(7), 1870–1889. doi:10.1111/gcb.16553

Wendler, G., Conner, J., Moore, B., Shulski, M., & Stuefer, M. (2011). Climatology of Alaskan wildfires with special emphasis on the extreme year of 2004. Theoretical and Applied Climatology, 104(3), 459–472. doi:10.1007/s00704-010-0357-9

Wierzchowski, J., Heathcott, M., & Flannigan, M. D. (2002). Lightning and lightning fire, central cordillera, Canada. International Journal of Wildland Fire, 11(1), 41. doi:10.1071/wf01048

Woolford, D. G., Martell, D. L., McFayden, C. B., Evens, J., Stacey, A., Wotton, B. M., & Boychuk, D. (2021). The development and implementation of a human-caused wildland fire occurrence prediction system for the province of Ontario, Canada. Canadian Journal of Forest Research, 51(2), 303–325. doi:10.1139/cjfr-2020-0313

Wotton, B. M., & Martell, D. L. (2005). A lightning fire occurrence model for Ontario. Canadian Journal of Forest Research, 35(6), 1389–1401. doi:10.1139/x05-071

World Bank. (2020). World Bank Official Boundaries: Admin 0, 10 meter boundary [Dataset]. World Bank Group.

Xu, W., Scholten, R. C., Hessilt, T. D., Liu, Y., & Veraverbeke, S. (2022). Overwintering fires rising in eastern Siberia. Environmental Research Letters, 17(4), 045005. doi:10.1088/1748-9326/ac59aa

Zheng, B., Ciais, P., Chevallier, F., Yang, H., Canadell, J. G., Chen, Y., … Zhang, Q. (2023). Record-high CO 2 emissions from boreal fires in 2021. Science, 379(6635), 912–917. doi:10.1126/science.ade0805

Zhu, C., Kobayashi, H., Kanaya, Y., & Saito, M. (2017). Size-dependent validation of MODIS MCD64A1 burned area over six vegetation types in boreal Eurasia: Large underestimation in croplands. Scientific Reports, 7(1). doi:10.1038/s41598-017-03739-0

https://doi.org/10.5194/essd-2024-465
Preprint. Discussion started: 29 October 2024

Zhu, Y., Rakov, V. A., Tran, M. D., Stock, M. G., Heckman, S., Liu, C., … Hare, B. M. (2017). Evaluation of ENTLN Performance Characteristics Based on the Ground Truth Natural and Rocket-Triggered Lightning Data Acquired in Florida. Journal of Geophysical Research: Atmospheres, 122(18), 9858–9866. doi:10.1002/2017jd027270

Zhu, Yanan, Stock, M., Lapierre, J., & DiGangi, E. (2022). Upgrades of the Earth Networks Total Lightning Network in 2021. Remote Sensing, 14(9), 2209. doi:10.3390/rs14092209