remote sensing

MDPI

Research Article - for submission to Special Issue "Earth Observations for Ecosystem Resilience"

Title: FIRED (Fire Events Delineation): An open, flexible algorithm & database of US fire events derived from the MODIS burned area product (2001-19)

Jennifer K. Balch[1,2,*], Lise A. St. Denis[1], Adam L. Mahood[1,2], Nathan P. Mietkiewicz[1,3], Travis Williams[1], Joe McGlinchy[1], and Maxwell C. Cook[1,2]

[1] Earth Lab, CIRES, University of Colorado Boulder, Boulder, CO 80309 USA;

[2] Department of Geography, University of Colorado Boulder, Boulder, CO 80309 USA;

[3] National Ecological Observatory Network (NEON), Boulder, Colorado 80301 USA

* Correspondence: jennifer.balch@colorado.edu

Received: date; Accepted: date; Published: date

Abstract:
Harnessing the fire data revolution, i.e., the abundance of information from satellites, government records, social media, and human health sources, now requires complex and challenging data integration approaches. Defining fire events is key to that effort. In order to understand the spatial and temporal characteristics of fire, or the classic fire regime concept, we need to critically define fire events from remote sensing data. Events, fundamentally a geographic concept with delineated boundaries around a specific phenomena that is homogenous in some property, are key to understanding fire regimes and more importantly how they are changing. Here, we describe FIRED, an event-delineation algorithm, that has been used to derive fire events (N = 51,871) from the MODIS MCD64 burned area product for the coterminous US from January 2001 to May 2019. The optimized spatial and temporal thresholds to cluster burned area pixels into events were an 11-day window and a 5-pixel distance, when optimized against 13,741 wildfire perimeters in the coterminous US from the Monitoring Trends in Burn Severity record. The linear relationship between FIRED and MTBS events size for the US was strong (R^2 = 0.92 for all events). Importantly, this algorithm is open source and flexible, allowing the end user to modify the spatio-temporal threshold or even the underlying algorithm as they see fit. We expect the optimized criteria to vary across regions, based on regional distributions of fire event size and rate of spread. We describe the derived metrics provided in a new national database and how they can be used to better understand US fire regimes. The open, flexible FIRED algorithm could be utilized to derive events in any satellite product. We hope that this open science effort will help catalyze a community-driven, data-integration effort (termed OneFire) to build a more complete picture of fire.

Keywords: data harmonization; event-builder algorithm; fire regimes; open fire science; satellite fire detections

1. Introduction

What is a fire? Defining the spatial and temporal boundaries of fire events is critical for understanding the drivers and trends in fires [1], ecological consequences [2], and adaptation implications [3]. Answering this question is, in fact, necessary for defining fire regimes, or the spatial and temporal characteristics of fires in a strict sense [4–6], which have been explored in recent decades via ground-based measurements [7–9], government records [10,11], or satellite sensors [12–14]. Here, we define an event [15] as a geographic concept with delineated boundaries around a specific phenomena that is homogenous in some property and distinct from adjacent areas. Delineation of events for environmental applications (e.g., wildfires, insect infestations, precipitation, or hurricanes) is a rich area of active research in Geographic Information Science [16], and fires have properties that are both discrete and continuous, requiring object and field-based approaches to defining them [16].

Remote sensing has fundamentally changed the way we quantify fire, and has consequently challenged how we define fire events. There are three classes of information that satellite sensors generally capture about fire behavior: active fires based on thermal threshold exceedance [17–19], fire radiative power as a metric of heat flux [20–22], and burned area derived from a change detection algorithm [23–25], sometimes also informed by active fire detections [26]. These properties of fires are estimated at the pixel level, which ranges in size for these products from 10s to 1000s of meters. In order to explore fire behavior patterns these pixel-level detections are aggregated in some way, necessitating the assumption of homogenous fire characteristics across that pixel. Global burned area products tend to underestimate total burned area due to missing small fires [33] and within-fire burned area due to underestimation of burned areas within an event [32]. Further, global scale studies explore total burned area summed across larger units or the density of hot pixels as a metric of fire frequency [13,27–29], which leaves understanding of actual events missing. Given the abundance of satellite fire data (e.g., Table 1), and that they do not "see" the same aspects of fire [12,31,32], we fundamentally need landscape-scale event delineation to integrate across products and build greater understanding of how fire regimes vary at regional and global scales [30].

There are several different approaches for delineating fire events based on proximity of burned area or hot pixels in space and time. Some studies have clustered the MODIS active fire hotspots (MODIS MOD12) to derive events in Europe and northern Africa to understand what drives large fires [31] and Indonesian tropical rainforests [32]. Others have used clustering of MODIS burned area pixels [12,13,33,34]. Most studies require pixel adjacency (Table 1), but a more relaxed spatial criteria facilitates exploring fires that have unburned patches within their perimeters—critical refugia that are necessary for regeneration [35], and is also less likely to over-segment events that are imperfectly detected, due to low fire severity or cloudiness, for example.

Given the number of studies that use the MODIS burned area product (e.g., studies in Table 1) and emerging new fire data products [19,24,30,36–38] that conduct some sort of event delineation as part of the processing, there is a great need to develop an open and well-documented algorithm for defining fire events from remotely sensed detections of fire. Moreover, event delineation enables joining of different data products to build a more complete picture of regional and global fire. Better delineation of the boundaries, and thus events, could lead to better estimates of total burned area, as well as exploration of derived spatiotemporal metrics around events that constitute the fire regime (e.g., event size, event shape, ignition point, unburned refugia within a fire, and fire spread rate).

There is a need however, to validate the approach, as the temporal and spatial thresholds chosen can substantially alter the number of detected wildfire events. And further, fire metrics can be sensitive to how boundaries are delineated [39]. Moreover, we expect the optimum temporal and spatial thresholds to vary based on size distribution and spread differences that will vary across ecoregions (e.g., fast, large grassland fires vs. small, slow temperate forest fires) and land use types (e.g., agricultural fires vs. deforestation fires). But even so, ground-based delineations of fire perimeters also have their challenges, incident command delineations may overestimate wildfire perimeters, as delineating unburned patches is difficult on the ground. Also, multiple fire patches may start independently and in proximity (e.g., when a lightning storm starts multiple events), which then merge into one fire complex.

Table 1: Studies using a spatio-temporal flooding algorithm for delineating fire events, first utilized by Archibald et al. [40].

Study	**Purpose**	**Satellite fire product**	**Spatial criteria**	**Temporal criteria**
Archibald et al. 2009	Examined environmental and anthropogenic drivers of fire in South Africa	MODIS MCD45	Adjacency	8 days
Balch et al. 2013	Tested whether cheatgrass occurrence increases fire activity	MODIS MCD45, RMGSC	2 pixels (1000 m)	2 days
Hantsen et al. 2015	Explored global fire size distribution	MODIS MCD45	Adjacency	14 days
Frantz et al. 2016	Aggregated raster grids from burn date to event objects	MODIS MCD64	10 pixels (5000 m)	5 days

Andela et al. 2017	Examined global fire activity	GFED4s, MODIS MCD64	Local spread rate x distance	Spatially varying fire persistence threshold
Laurent et al. 2018	Derived patch functional traits and other morphological features of fire events	MODIS MCD64, MERIS	1 pixel (500 m)	3, 5, 9 and 14 days
Andela et al. 2018	Created global fire atlas product	MODIS MCD64	1 pixel (500 m)	Spatially varying

Here, we: i) develop an open, refined, and adaptable algorithm for defining events; ii) derive events and companion metrics for fires in the conterminous US from the MODIS MCD64 burned area product, based on the optimum spatial and temporal thresholds; iii) validate the MODIS-derived events against the Landsat-derived MTBS product [41]; and iii) demonstrate how defining events enables us to explore additional metrics of the fire regime across the US. The algorithm is designed in a way that makes it adaptable to data source, regional context, and even event type: it could be used with newer burned area products (e.g., Fire_cci based on MODIS images at 250 m resolution [36] or VIIRS [19]) and the spatiotemporal criteria can be altered.

2. Materials and Methods

a. Study area and data acquisition and processing

The study area was the coterminous United States (CONUS). We chose this study area because of the availability of other fire datasets like MTBS [41] which we were able to use to gauge the accuracy of our aggregation of burned pixels to events from the MCD64 dataset. We used the MODIS Collection 6 MCD64 burned area product [37] [available at ftp://fuoco.geog.umd.edu/MCD64A1/C6/]. These data contain five layers: burn date, first day, last day, a quality assessment, and error. The data are available worldwide, via a sinusoidal projection that is divided into 648 tiles (268 of which are terrestrial), each with 2400 rows and columns at 463-m resolution. We downloaded the entire monthly time series available for each tile that overlaps with the coterminous US, and extracted the burn date layer.

b. Accounting for pixels that burn more than once per year

Some other studies that have aggregated pixels into fire events from the MODIS burned area product have aggregated the input data to a yearly time-step [42,43], taking either the earliest or latest burn date in the case of pixels that burn twice in one year. This assumes a minimal occurrence of pixels that actually burn twice in one year (e.g. the land burns first in spring and then again in fall). This makes the processing of the data much less complex, however it presents two problems. First, any significant occurrence of reburned pixels within a year would result in an underestimate of burned area. Second, fires that burn from one year to the next become split into two events.

To investigate whether reburned pixels would have a significant confounding effect on our data, we calculated intra-year reburns for each of the tiles we used in our analysis for each year. We converted each monthly tile in the coterminous US to binary (1 for burned, 0 for unburned), summed each monthly pixel per year and calculated the percentage of pixels that burned more than once per tile, per year. For 2001 - 2016 for the coterminous US, there were a total of 12,676 pixels that burned more than once in a given year (excluding the h10v06 tile), or about 0.48% of pixels. The tile that includes Florida (h10v06) had a rate of 5% (sd 2.3%) reburns per year (Table 2). We suspect that this high reburn occurrence is due to the year-round growing season combined with year-round occurrence of lightning strikes. This would present a problem if this algorithm were expanded globally, because there are many ecosystems with year-round growing seasons combined with year-round ignition sources, be they anthropogenic sources or lightning.

Because of the relatively high reburn occurrence, and also due to concern over segmenting winter fires into multiple events, we decided not to aggregate the input rasters by year or fire season. Instead, we created a space-time cube for each monthly tile for the entire time series, where the julian day of the year for each pixel in each month layer was converted to a number along a continuous series starting on January 1, 1970.

Table 2. Number of reburned pixels per year, per tile calculated for 2001-2018 from the monthly MODIS MCD64 Burned area product.

Tile	Mean Reburn %	Std Reburn %
h08v04	0.17	0.18
h08v05	0.35	0.27
h08v06	1.35	1.05
h09v04	0.36	0.30
h09v05	0.23	0.19
h09v06	0.73	0.47
h10v04	0.12	0.09
h10v05	0.67	0.29
H10v06 (Florida)	5.12	2.31
h11v04	0.35	0.33
h11v05	0.32	0.35
h12v04	0.35	0.61
h13v04	0.32	0.29
Total (excluding h10v06)	**0.48**	**0.55**

c. Defining events with a flexible, fast algorithm
To define events, we used a 3-dimensional moving window to aggregate burned pixels into distinct events. The algorithm takes as input a spatial variable, representing the number of pixels, and a temporal variable, representing the number of days, within which to group burn detections. It then aggregates by assigning each burned pixel an event identification number.

The event perimeter script reads in a pre-processed MODIS 3-dimensional netCDF file for each tile, where each band represents one month, and for each burned pixel the date of fire detection is represented as the number of days since January 1, 1970. The netCDF file is converted into a three-dimensional array and the moving window traverses the array. To avoid unnecessary computation, we did not check cells in which there was no burned area assignment throughout the study period.

For each cell where at least one fire detection occurred, the program creates a mask identifying all burned pixels that fall within the spatial and temporal range of the current cell. If the current cell is already part of an existing event, any new burned pixels are assigned the event ID for that event. If it is a new event, the current cell and all overlapping cells are given the next sequential event ID. If there are multiple event IDs within the mask, two perimeters have grown together and they are merged into the first event ID. After the event perimeters are delineated within each tile, all event perimeters that potentially overlap with an adjacent tile are flagged. After all tiles are processed, the flagged events are partitioned and those that overlap spatially and temporally are merged. Finally, events across all tiles are merged into a final dataset and given a new sequential event ID.

d. Sensitivity analysis: identifying the optimal spatiotemporal thresholds for delineating fire events
In order to find which combination of spatial and temporal variables outputs best defined fire events, we used MTBS fire perimeters [41] as a validation reference. MTBS is a dataset of fire perimeters from 1984-2016, derived from Landsat satellite data, which has a minimum size threshold of 404 ha in the western US and 202 ha in the eastern US (separated by the 97th parallel). It documents 21,673 fire events throughout the entire US, and 13,741 in the overlapping study area and timeframe, beginning in 2001. One problematic feature of the MTBS data for this comparison is that fire complexes are not dealt with uniformly. Fire complexes are "two or more individual incidents located in the same general area which are assigned to a single incident commander or unified command [44]." In some cases each fire patch is assigned its own ID number and is represented as a single perimeter, and in other cases these complexes are lumped into a multipolygon with a single ID number. To address this issue, we split all multipolygons into single polygons, assigned unique ID numbers to each polygon, and then calculated the area for each individual polygon. This way, our sensitivity analysis would objectively assess how individual polygons matched, without the confounding factor of aggregated multipolygons.

We ran the fire event classifier for all spatiotemporal combinations between 1-15 days and 1-15 pixels, resulting in 225 spatiotemporal combinations for the conterminous US For each combination we associated the FIRED events that were >404 ha in the west and >202 ha in the eastern US to the associated MTBS wildfire perimeter.

An accuracy assessment was conducted for each spatiotemporal combination of the MODIS-based events, based on how well the MODIS events matched the MTBS events. For each unique fire polygon in the MTBS database, we extracted the ID numbers for each MODIS event overlapping the MTBS polygon. Then, for each unique MODIS event, we extracted each MTBS ID that overlapped. We then calculated the ratio of the number of unique MTBS events that contained a MODIS event divided by the number of unique MODIS events that contained at least one MTBS event, with the optimum value being one. We used this ratio to approximate the spatio-temporal combination that minimized both over- and under-segmentation of the MODIS fire events based on known MTBS fire perimeters.

After evaluating the outputs of all space/time pairwise combinations, we estimated an optimal combination for the US of 5 pixels and 11 days. We calculated commission and omission errors for both the MODIS-based events and the MTBS events.

e. Calculating statistics for each event, and daily statistics within events

Once the optimal spatial-temporal aggregation level was identified, we created two vector products for the conterminous US: one where individual pixels were aggregated to polygons representing each fire event, and one where individual pixels were aggregated to each date within each event. For the event-level vector product, we calculated ignition location and timing, duration, spread rate, burned area, date of maximum growth, area burned on the dates of maximum and minimum growth, and the mean daily growth for each event. We also extracted the mode of the International Geosphere-Biosphere Programme land cover classification from the MODIS MCD12Q1 landcover product for the year before the fire, and the level 1-3 ecoregions, for each event (Table 3). For the daily-level vector product, we calculated the daily burned area, cumulative burned area per day, days since ignition, mode landcover per day, and mode ecoregion per day, in addition to the metrics calculated for the event-level product (Table 4).

Table 3. Attributes included in the event-level FIRED product.

Attribute	**Units**
Ignition	date, day of year, month, year, location
Duration	days
Burned Area	km^2, ha, acres, pixels
Fire Spread Rate	pixels/day, km^2/day, ha/day, acres/day

Maximum, minimum, and mean growth rate	km^2, ha, acres, pixels, date (max only)
Land Cover (for the year before the fire)	mode land cover classification / event
Ecoregion	mode ecoregion, Levels 1-3

Table 4. Attributes included in the daily-level FIRED product

Attribute	**Units**
Daily Burned Area	km^2, ha, acres, pixels
Daily Landcover	mode land cover classification / day
Daily Ecoregion	mode ecoregion, Levels 1-3
Cumulative Burned Area	km^2, ha, acres, pixels
Ignition Date	date
Last Burn Date	date
Duration	days
Event Day	days from ignition date
Percent Total Area	percent (%)
Percent Cumulative Area	percent (%)
Fire Spread Rate	pixels/day, km^2/day, ha/day, acres/day

f. Data and code availability

Code for the python command line interface used to download data, classify events, calculate event- or daily-level statistics, and write tables and shapefiles is available as the "firedpy" python package at www.github.com/earthlab/firedpy. R code for the analysis presented here is available at https://github.com/earthlab/modis-fire-events-delineation. R code for the sensitivity analysis is available at www.github.com/admahood/fired_optimization. Data is available at CU Scholar [at time of publication will have a DOI].

3. Results

a. Classification accuracy assessment

The MODIS-derived events had a 55% omission and 62% commission error, compared to the MTBS reference dataset (Table 5). Commission error was calculated as: 11,412/(11,412+7,054) .

Omission error was calculated as 8,721/(8,721+7,054). An additional 24,163 events were detected below the MTBS size thresholds and were not included in these calculations.

$$CE = \frac{FIRED_{true}MTBS_{false}}{(FIRED_{true}MTBS_{false}+FIRED_{true}MTBS_{true})} = \frac{11{,}412}{(11{,}412+7{,}054)} = 0.62$$

$$OE = \frac{FIRED_{false}MTBS_{True}}{(FIRED_{true}MTBS_{false}+FIRED_{true}MTBS_{true})} = \frac{8{,}721}{(8{,}721+7{,}054)} = 0.55$$

Table 5. Confusion matrix for the MODIS MCD64-derived events. The MTBS event-size threshold is 404 ha in the western US, 202 ha in the eastern US.

	MTBS True	MTBS False (Commission)	MTBS False (Commission)
FIRED True	7,054	11,412 (over threshold only)	24,163 (under threshold only)
FIRED False (Omission)	8,721	-	-

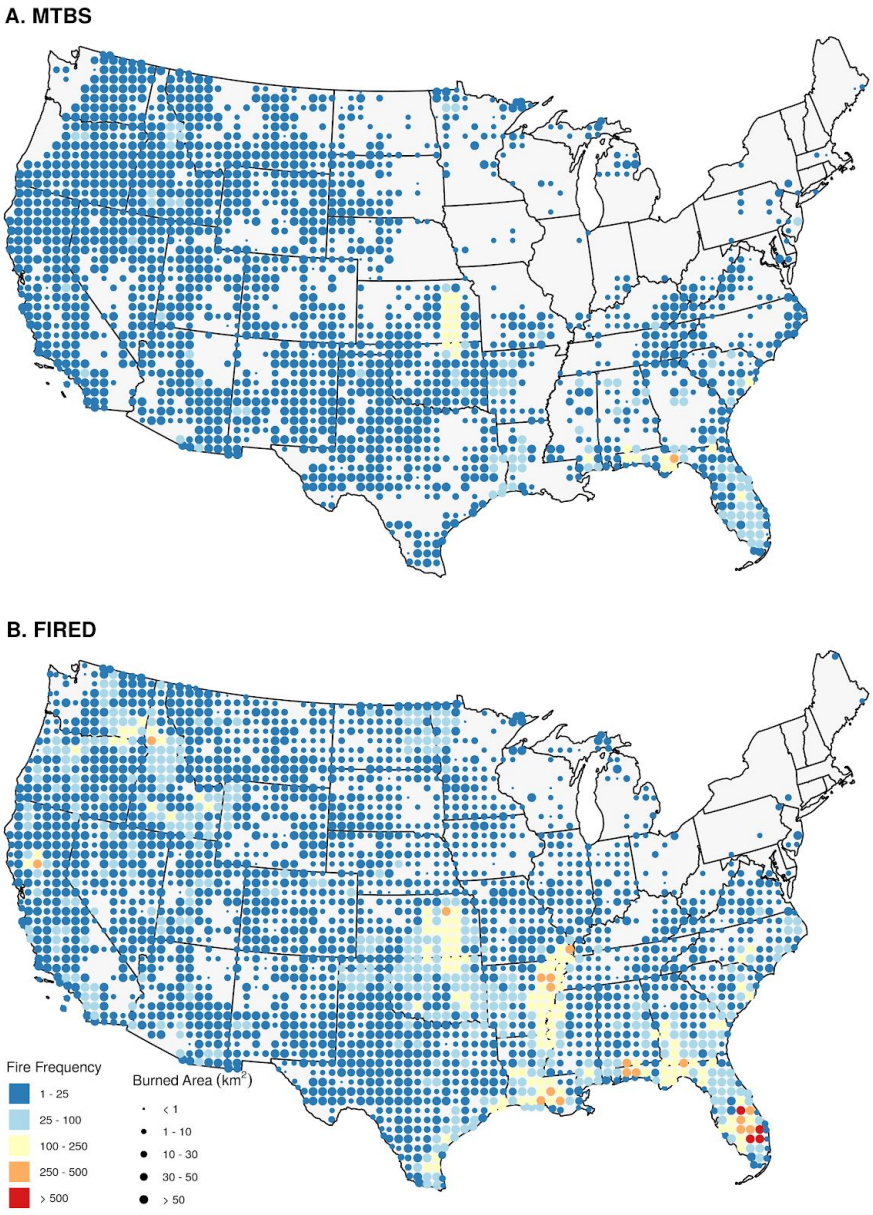

Figure 1. A comparison of the spatial distribution of fire events from the FIRED and MODIS products shows a similar distribution of fire events and burned area in general, but the FIRED algorithm picks up many more events and burned area in the midwest, southeastern US and eastern Washington.

b. Comparison to MTBS:

There were approximately 3.3 times more wildfire events and 65,000 km^2 (18%) more burned area captured in the FIRED product compared to MTBS. The FIRED burned area represents 97% of the NIFC reported totals from 2001-2016 (Table 6).

Table 6: Fire events and burned area by level one ecoregion, 2001-2016.

Level 1 Ecoregions	MTBS		FIRED		NIFC	
	Events	Burned Area (km^2)	Events	Burned Area (km^2)	Events	Burned Area (km^2)
Eastern Temperate Forests	5,644	47,116	20,556	103,615	-	-
Great Plains	3,350	94,068	11,818	112,907	-	-
Marine West Coast Forest	22	379	249	978	-	-
Mediterranean California	368	17,971	1,432	21,251	-	-
North American Deserts	1,739	80,430	5,689	72,012	-	-
Northern Forests	134	2,130	141	2,086	-	-
Northwestern Forested Mountains	1,614	81,189	3,815	68,006	-	-
Southern Semi-Arid Highlands	159	5,494	260	4,459	-	-
Temperate Sierras	431	19,374	447	13,674	-	-
Tropical Wet Forests	266	4,818	1,394	19,424	-	-
Conterminous US	13,727	352,967	45,801	418,414	1,153,896	432,733

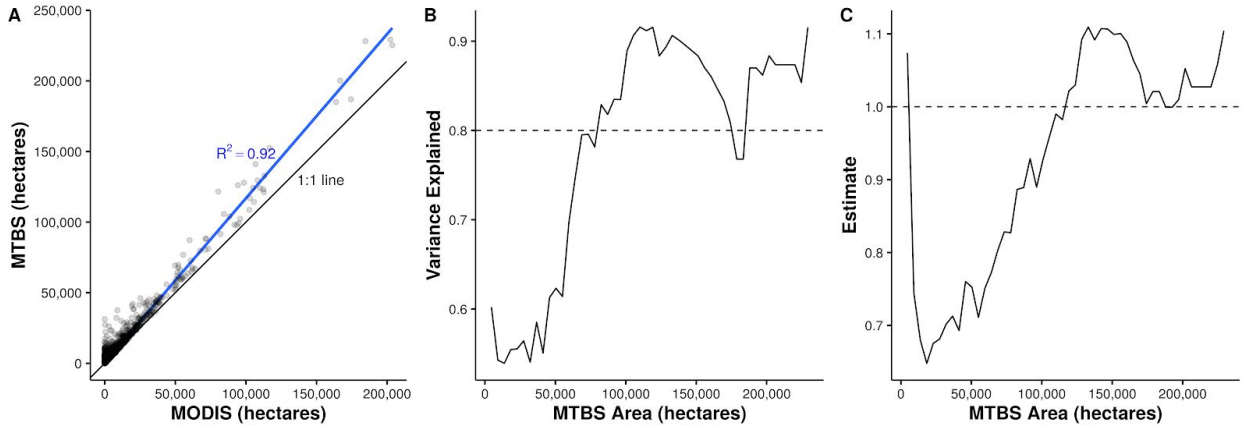

Figure 2: Panel A shows the relationship between area burned for MTBS and FIRED events, for fires captured by both products. While the relationship is generally strong (R^2 = 0.92 for all events), it is weaker for smaller fires. For panels B and C we binned the data into 50 equal size classes (each bin spans ~ 5000 hectares), and ran a linear regression (MTBS burned area predicted by MODIS burned area) on each bin. Panel B shows the R^2 values, which do not

consistently stay above 0.8 until about 70,000 hectares. Panel C shows the relationship between the slope of the regression line and MTBS burned area, illustrating that the MODIS MCD64A1 burned area product consistently underestimates burned area for fires below 100,000 hectares.

d. Ecoregion comparisons between FIRED and MTBS

One of the primary differences between the two products is the detection of small fires, which is a function of the ~200-ha and ~400-ha cut-off for the eastern and western US in the MTBS product [41]. In the east and central US, where fires are generally smaller, FIRED captured 37,724 fires while MTBS captured 11,008 fires (Figure 1, Table 5). There were several ecoregions where FIRED captured more events, but less burned area (e.g., in North American Deserts; Table 5). This is either due to the lack of smaller events in the MTBS dataset, or that MTBS does not delineate unburned patches within its fire perimeters, and so can overestimates burned area for many fires (e.g., see Figure 3).

Ecoregions with the highest maximum fire spread rates were those with large areas of grasslands - the Great Plains and desert ecoregions (Table 7). However, the three ecoregions with the highest mean fire spread rates were all forested ecosystems - the temperate Sierras, southern semi-arid highlands, and northern forests, and these ecoregions also had the highest variability in fire spread rates.

Table 7. Fire spread rate by ecoregion.

Level 1 Ecoregions	Fire Events	Fire Spread Rate (ha/day)					
	n	Max	Lower 95%tile	Mean	Upper 95%tile	SD	SE
Eastern Temperate Forests	20,556	2,756	9	43	119	60	0.4
Great Plains	11,818	13,584	12	95	279	293	2.7
Marine West Coast Forest	249	301	7	42	143	45	2.8
Mediterranean California	1,432	5,883	11	126	497	329	8.7
North American Deserts	5,689	14,620	11	137	481	487	6.5
Northern Forests	141	2,442	10	144	614	312	26.3
Northwestern Forested Mountains	3,815	3,878	10	105	415	233	3.8
Southern Semi-Arid Highlands	260	1,755	17	162	550	244	15.2
Temperate Sierras	447	6,365	16	194	627	541	25.6
Tropical Wet Forests	1,394	1,220	8	45	117	85	2.3

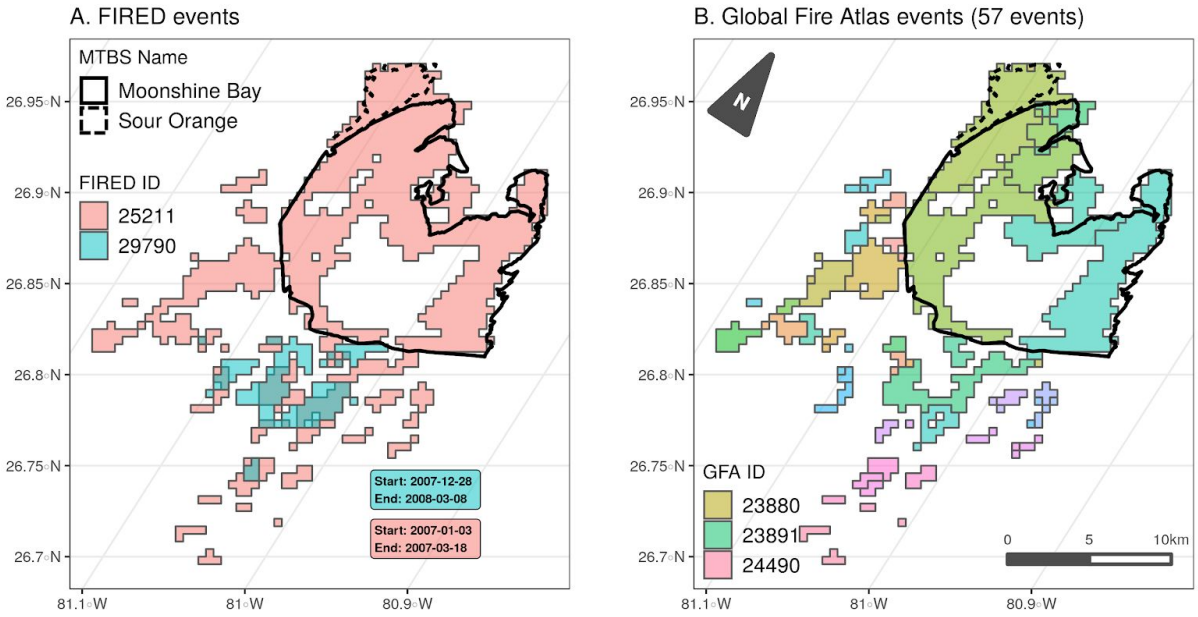

Figure 3. Comparison of A) FIRED and B) Global Fire Atlas delineated events for the Sour Orange fire (started February 9, 2007), the Moonshine Bay fire (started February 24, 2007), and a third unnamed event, FIRED event #29790 (started December 28, 2007, and continued into March of 2008). The FIRED product joins the two intra-year burns (#25211) and delineates a third event (#29790) that reburns some of the same pixels. The dark outlines, bold and dashed, show the MTBS fire perimeters for the Sour Orange and Moonshine Bay fires. Note that MTBS does not include unburned patches within perimeters. Panel B) shows the Global Fire Atlas (with an abridged legend), which segments the same MODIS burned area pixels into 57 events and no delineation of overlapping reburns.

4. Discussion

Remote sensing has fundamentally changed our ability to quantify fire, and has consequently challenged how we define fire events. The active fire, burned area, and fire radiative power and severity products [19,21,22,24,25,37,41] have fundamentally changed how we can conceptualize fire regimes. Key to translating this wealth of information is defining fire events in space and time so that we can understand how modern fire regimes are changing. Parallel efforts such as the Global Fire Atlas (based on the MODIS MCD64 product [34]) have converged on identifying the same need, with a key motivation to improve global fire modeling [35]. We argue that the need is more profound, that in order to understand how fire regimes are changing at regional to global scales we need an open, and flexible methodology to identify events and integrate fire data across sources based on these events. This event-based approach could be utilized to derive events in any satellite product to build a more complete picture of fire.

There are several unique aspects of this dataset, algorithm, and approach that are worth highlighting. The primary difference between FIRED and other algorithms is that FIRED uses the entire monthly time series as a space-time cube input, upon which a 3-dimensional moving window is applied, compared to aggregating fire seasons or years into one layer on which a 2-dimensional moving window is applied. Our approach enables proper identification of intra-year reburns and ensures that fires at the edges of seasons/years are not artificially split into multiple events (Figure 3). Second, the FIRED database delineates small fire events, expanding our ability to understand how fire size and burned area is changing, beyond just the large events [45]. Smaller events are difficult to capture systematically but we know these events can be incredibly important in the US, contributing large additional burned areas and emissions [46,47]. Third, this product provides several attributes that are new pieces of information, refined across the conterminous US. For example, fire spread rate is a unique attribute, derived from events, which is a critical piece of information not easily accessed in other datasets (e.g., MTBS or ICS-209s). FIRED also provides the landcover the year before the fire for each event, a coarse metric of fuels information.

A key advantage of this approach is that the algorithm is open and flexible; we hope for community input and we expect it to be improved over time. The spatio-temporal criteria can be altered based on other information, regionally-specific fire perimeters such as Canada's National Burned Area Composite (https://cwfis.cfs.nrcan.gc.ca/datamart), or known delineations of intentional land use fires or prescribed burns. Further, we anticipate that this algorithm has wide applicability to other fire products and other efforts to build events based on any geospatial data that has both spatial and temporal information. Previous studies, including this team's previous efforts (Balch et al. 2013), have not made their workflow and code publicly available, limiting the potential to facilitate community development of an integrated, global fire database.

With the plethora of remote sensing data about fire and fire effects, there is a great need to delineate events at large regional and global scales. There are at least three other recent studies that have created fire events from the MODIS burned area product (Table 1), two of which [42,48] have created global fire event databases. In addition to the global efforts, Frantz *et al.* [43] created an algorithm based on a study area in sub-Saharan Africa which uses a top–down multilevel segmentation strategy that starts by defining potential ignition points and gradually refines the individual object membership. All three efforts use an approach that starts by identifying potential ignition points and grows objects from the ignition point using only adjacent pixels. The code for the algorithm created by Andela et al. [42] is not publicly available and the code created by Frantz et al. [43] is available upon request. Laurent et al. [48] created a publicly available database and the code is also available upon request. Their output data contains what they term fire patch functional traits, including patch area and other morphological features, but does not preserve daily fire spread information or polygons containing the perimeter shapes of the derived events. Our approach differs in that we use a spatiotemporal window that can capture isolated burned pixels that may be part of the same event, but may be isolated because of the inability of the MODIS sensor to detect burned area in

the area between patches due to cloudiness, low vegetation density, low severity, or unburned patches (i.e., fire refugia) that are important elements of an event. It is worth noting that the spatial-temporal thresholds we derived (i.e., 11-day window and a 5-pixel distance) are much greater than those used in most previous studies (e.g. [12,34] but see Frantz et al. [43]), leading to less artificial truncation, or oversplitting, of events. For example, the Rim fire which occurred in California in 2013 was delineated into more than 10 separate events by the Global Fire Atlas algorithm, whereas our algorithm delineated a single event that more closely matches the MTBS delineation (Figure 4). Future improvements could include: i) validation with smaller events, such as those contained in the US-based GeoMac dataset [49] or others; ii) estimates of uncertainty around start and end dates of the fires; and iii) regionally-varying thresholds based on fire regime characteristics.

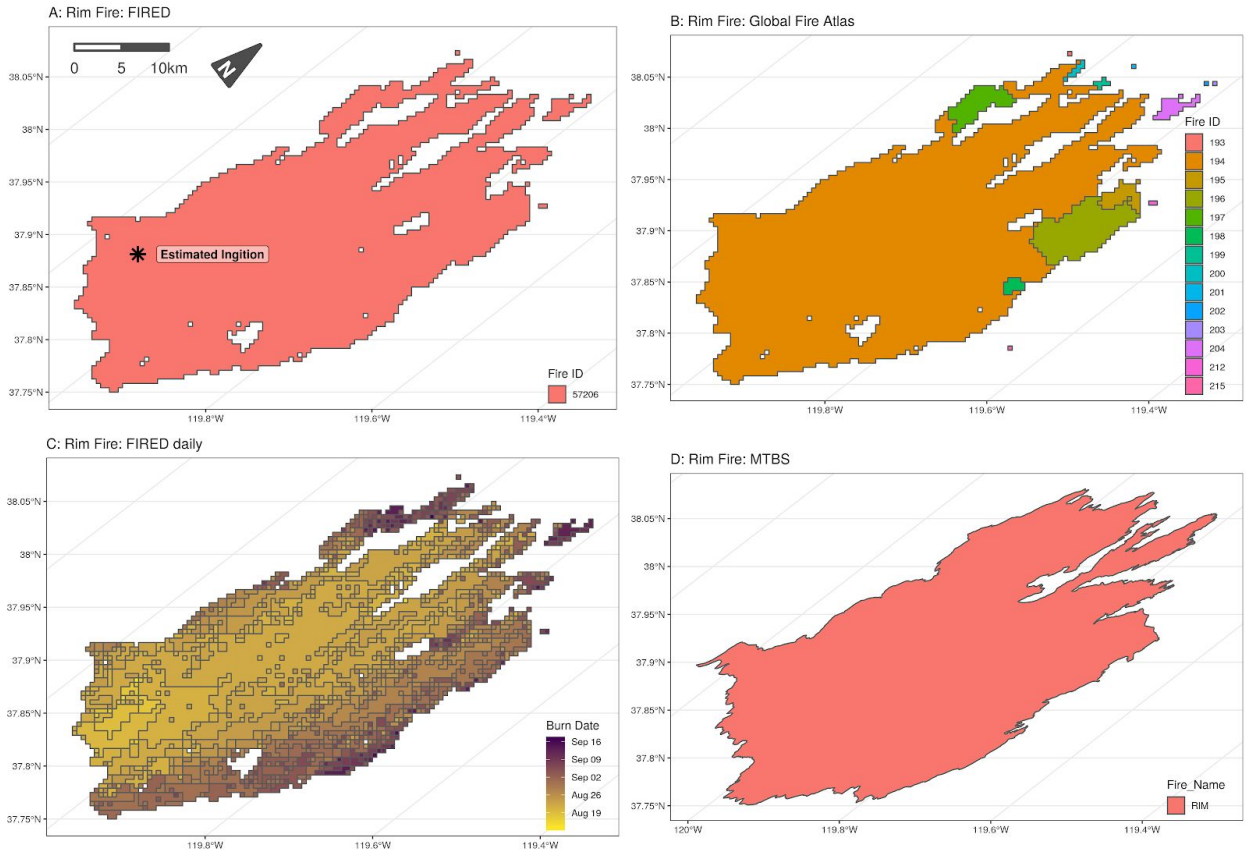

Figure 4: The 2013 Rim Fire, which lasted over a month and was more than 100,000 ha in total size according to incident reports, as delineated by the A) FIRED event product; B) global fire atlas C) FIRED daily event product; and D) MTBS. The optimized spatial-temporal criteria we used allowed us to correctly classify it as a single event, while the global fire atlas has segmented the Rim Fire into 14 separate events. The FIRED ignition point is estimated as the average location of all pixels occurring on the first day of the event.

This is a unique moment in the history of fire science, given the abundance of fire data across spatial scales, that requires the fire science community to better coordinate efforts on fire data

harmonization challenges and opportunities. We see great potential to build a community-driven, fire data infrastructure that we term OneFire. OneFire is a coordinated architecture that would enable a community of researchers and stakeholders to use, repurpose, and contribute to fire data, code, and workflows. The vision for OneFire is that it will be a coordinated, community-inspired data architecture that connects and integrates the many global, national, and regional fire databases. This is no small task, but integrating these datasets is key to unlocking a transformation in fire science and rapidly accelerating new discoveries about why fire regimes are changing and how societies and ecosystems are vulnerable. There is an enormous amount of data and work relevant for fire science that could be leveraged, if only it was open, reproducible, and scalable. For example, we anticipate that a newly published ICS-209-PLUS dataset that is an integrated database of over 120,000 incident command reports could be connected to MODIS FIRED events to join physical attributes with social impact and response on a daily scale [50]. Social media information around wildfires could also be leveraged, and provide a view of social response that before would not have been possible [51,52]. Or additional satellite sensors, e.g., active fire products, could be merged to fill holes in the burned area record. Key elements of a vision for OneFire include: i) identified fire events across many datasets utilizing the FIRED event-builder algorithm or other approach the delineates events in space and time; ii) integration workflows that then connect those same events across data sources to build a fuller suite of attributes around commonly identified events; iii) data and computational infrastructure that allows for community contributions of data, code, and compute environments; iv) formal linkages to other important climate, environment, and social data sources that provide insights into driving forces or responses; and v) support for community building, engagement, and training that facilitates large, diverse team science. Ultimately, no single sensor is going to provide all the information we need about fires, and we will never anticipate all the ways that such an integrated source of fire information would get used. OneFire would help us build a fuller, global picture of fire.

5. Conclusions

There is a clear need to derive events from remotely sensed detections of fires, as fundamentally without events we cannot explore how the spatio-temporal properties of fire regimes are changing. Further, there are dozens of fire products available, for the US and globally (Table 1), that could, if combined and harmonized, shed new insights on the drivers and consequences of changing fire. Delineating fire events is key to this process, and we argue that this US database and algorithm offer the opportunity to begin to build OneFire, a community data-integration effort for fire science. No one research group can predict the variables that will be needed for all studies, and there is no one satellite that captures all the needed information about fire [53]. We envision that our algorithm will be optimized at different scales to best capture regional fire size distributions. We also envision that this algorithm can be used across any satellite-based fire product, from active fire detections to burned area products, and particularly new efforts, such as the BAECV product or VIIRS. Moreover, this algorithm can be used with any spatiotemporal data and is not constrained to fire data. As other efforts are built to understand natural hazards, these efforts may help to better delineate the spatial and temporal dimensions of floods,

hurricanes, disease outbreaks, and other events. The fire science community can better harmonize fire observations for a larger network of researchers and practitioners who need this information to better help society more sustainably live with fire.

Author Contributions: Conceptualization, JB, AM and NM; methodology, JB, LS, AM, NM and TW.; software, LS, AM, NM, TW, JM, and MC; validation, LS, AM and NM.; formal analysis, LS, AM and NM.; investigation, JB, LS, AM and NM; data curation, AM, NM and TW; writing—original draft preparation, JB, LS and AM; writing—review and editing, JB, LS, AM, NM and JM; visualization, AM and NM.; supervision, JB and LS.; project administration, JB; funding acquisition, JB. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Aeronautics and Space Administration Terrestrial Ecology Program (grant number NNX14AJ14G) and by CIRES and the Grand Challenge Initiative at the University of Colorado, Boulder.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bowman, D.M.J.S.; Balch, J.K.; Artaxo, P.; Bond, W.J.; Carlson, J.M.; Cochrane, M.A.; D'Antonio, C.M.; Defries, R.S.; Doyle, J.C.; Harrison, S.P.; et al. Fire in the Earth System. *Science* **2009**, *324*, 481–484, doi:10.1126/science.1163886.
2. Fortin, M.-J.; Drapeau, P. Delineation of ecological boundaries: comparison of approaches and significance tests. *Oikos* **1995**, 323–332.
3. Schoennagel, T.; Balch, J.K.; Brenkert-Smith, H.; Dennison, P.E.; Harvey, B.J.; Krawchuk, M.A.; Mietkiewicz, N.; Morgan, P.; Moritz, M.A.; Rasker, R.; et al. Adapt to more wildfire in western North American forests as climate changes. *Proc. Natl. Acad. Sci.* **2017**, *114*, 4582–4590, doi:10.1073/pnas.1617464114.
4. Krebs, P.; Pezzatti, G.B.; Mazzoleni, S.; Talbot, L.M.; Conedera, M. Fire regime: history and definition of a key concept in disturbance ecology. *Theory Biosci.* **2010**, *129*, 53–69.
5. Gill, D.E. Spatial patterning of pines and oaks in the New Jersey pine barrens. *J. Ecol.* **1975**, 291–298.
6. Pyne, S.; Andrews, P.; Laven, R.D. Introduction to Wildland Fire, John Wiley and Sons. *N. Y.* **1996**.
7. Balch, J.K.; Nepstad, D.C.; Brando, P.M.; Curran, L.M.; Portela, O.; de Carvalho Jr, O.; Lefebvre, P. Negative fire feedback in a transitional forest of southeastern Amazonia. *Glob. Change Biol.* **2008**, *14*, 2276–2287.
8. Veblen, T.T.; Kitzberger, T.; Donnegan, J. Climatic and human influences on fire regimes in ponderosa pine forests in the colorado front range. *Ecol. Appl.* **2000**, *10*, 1178–1195, doi:10.1890/1051-0761(2000)010[1178:CAHIOF]2.0.CO;2.
9. Veblen, T.T.; Hadley, K.S.; Nel, E.M.; Kitzberger, T.; Reid, M.; Villalba, R. Disturbance Regime and Disturbance Interactions in a Rocky Mountain Subalpine Forest Reid and Ricardo Villalba Published by : British Ecological Society Stable URL : http://www.jstor.org/stable/2261392 REFERENCES Linked references are available on JSTOR f. *J. Ecol.* **1994**, *82*, 125–135.
10. Balch, J.K.; Bradley, B.A.; Abatzoglou, J.T.; Nagy, R.C.; Fusco, E.J.; Mahood, A.L. Human-started wildfires expand the fire niche across the United States. *Proc. Natl. Acad. Sci.* **2017**, doi:10.1073/pnas.1617394114.
11. Williams, M.A.; Baker, W.L. Spatially extensive reconstructions show variable-severity fire and heterogeneous structure in historical western United States dry forests. *Glob. Ecol. Biogeogr.* **2012**, *21*, 1042–1052, doi:10.1111/j.1466-8238.2011.00750.x.
12. Balch, J.K.; Bradley, B.A.; D'Antonio, C.M.; Gómez-Dans, J. Introduced annual grass increases regional fire activity across the arid western USA (1980-2009). *Glob. Change Biol.* **2013**, *19*, 173–183, doi:10.1111/gcb.12046.
13. Archibald, S.; Lehmann, C.E.R.; Gómez-dans, J.L.; Bradstock, R.A. Defining pyromes and global syndromes of fire regimes. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 6442–6447, doi:10.1073/pnas.1211466110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1211466110.
14. Morton, D.C.; Collatz, G.J.; Wang, D.; Randerson, J.T.; Giglio, L.; Chen, Y. Satellite-based assessment of climate controls on US burned area. *Biogeosciences* **2013**, *10*, 247–260, doi:10.5194/bg-10-247-2013.
15. Worboys, M. Event-oriented approaches to geographic phenomena. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 1–28, doi:10.1080/13658810412331280167.
16. Yuan, M. Representing Complex Geographic Phenomena in GIS. *Cartogr. Geogr. Inf. Sci.* **2001**, *28*, 83–96, doi:10.1559/152304001782173718.

17. Dwyer, E.; Pinnock, S.; Grégoire, J.-M.; Pereira, J. Global spatial and temporal distribution of vegetation fire as determined from satellite observations. *Int. J. Remote Sens.* **2000**, *21*, 1289–1302.
18. Justice, C.; Giglio, L.; Korontzi, S.; Owens, J.; Morisette, J.; Roy, D.; Descloitres, J.; Alleaume, S.; Petitcolin, F.; Kaufman, Y. The MODIS fire products. *Remote Sens. Environ.* **2002**, *83*, 244–262.
19. Schroeder, W.; Oliva, P.; Giglio, L.; Csiszar, I.A. The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment. *Remote Sens. Environ.* **2014**, *143*, 85–96, doi:https://doi.org/10.1016/j.rse.2013.12.008.
20. Li, F.; Zhang, X.; Kondragunta, S.; Csiszar, I. Comparison of Fire Radiative Power Estimates From VIIRS and MODIS Observations. *J. Geophys. Res. Atmospheres* **2018**, *123*, 4545–4563, doi:10.1029/2017JD027823.
21. Wooster, M.J.; Xu, W.; Nightingale, T. Sentinel-3 SLSTR active fire detection and FRP product: Pre-launch algorithm development and performance evaluation using MODIS and ASTER datasets. *Remote Sens. Environ.* **2012**, *120*, 236–254, doi:https://doi.org/10.1016/j.rse.2011.09.033.
22. Freeborn, P.H.; Wooster, M.J.; Roy, D.P.; Cochrane, M.A. Quantification of MODIS fire radiative power (FRP) measurement uncertainty for use in satellite-based active fire characterization and biomass burning estimation. *Geophys. Res. Lett.* **2014**, *41*, 1988–1994, doi:10.1002/2013GL059086.
23. Roy, D.P.; Boschetti, L.; Justice, C.O.; Ju, J. The collection 5 MODIS burned area product — Global evaluation by comparison with the MODIS active fire product. *Remote Sens. Environ.* **2008**, *112*, 3690–3707, doi:https://doi.org/10.1016/j.rse.2008.05.013.
24. Hawbaker, T.J.; Vanderhoof, M.K.; Beal, Y.-J.; Takacs, J.D.; Schmidt, G.L.; Falgout, J.T.; Williams, B.; Fairaux, N.M.; Caldwell, M.K.; Picotte, J.J.; et al. Mapping burned areas using dense time-series of Landsat data. *Remote Sens. Environ.* **2017**, *198*, 504–522, doi:https://doi.org/10.1016/j.rse.2017.06.027.
25. Mouillot, F.; Schultz, M.G.; Yue, C.; Cadule, P.; Tansey, K.; Ciais, P.; Chuvieco, E. Ten years of global burned area products from spaceborne remote sensing—A review: Analysis of user needs and recommendations for future developments. *Int. J. Appl. Earth Obs. Geoinformation* **2014**, *26*, 64–79.
26. Giglio, L.; Loboda, T.; Roy, D.P.; Quayle, B.; Justice, C.O. An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sens. Environ.* **2009**, *113*, 408–420, doi:https://doi.org/10.1016/j.rse.2008.10.006.
27. CHUVIECO, E.; GIGLIO, L.; JUSTICE, C. Global characterization of fire activity: toward defining fire regimes from Earth observation data. *Glob. Change Biol.* **2008**, *14*, 1488–1502, doi:10.1111/j.1365-2486.2008.01585.x.
28. Krawchuk, M.A.; Moritz, M.A.; Parisien, M.-A.; Van Dorn, J.; Hayhoe, K. Global pyrogeography: the current and future distribution of wildfire. *PloS One* **2009**, *4*, e5102.
29. Van der Werf, G.R.; Randerson, J.T.; Giglio, L.; Collatz, G.; Mu, M.; Kasibhatla, P.S.; Morton, D.C.; DeFries, R.; Jin, Y. van; van Leeuwen, T.T. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chem. Phys.* **2010**, *10*, 11707–11735.
30. Chuvieco, E.; Yue, C.; Heil, A.; Mouillot, F.; Alonso-Canas, I.; Padilla, M.; Pereira, J.M.; Oom, D.; Tansey, K. A new global burned area product for climate assessment of fire impacts. *Glob. Ecol. Biogeogr.* **2016**, *25*, 619–629, doi:10.1111/geb.12440.
31. Loepfe, L.; Rodrigo, A.; Lloret, F. Two thresholds determine climatic control of forest fire size in Europe and northern Africa. *Reg. Environ. Change* **2014**, *14*, 1395–1404,

doi:10.1007/s10113-013-0583-7.
32. Cattau, M.E.; Harrison, M.E.; Shinyo, I.; Tungau, S.; Uriarte, M.; DeFries, R. Sources of anthropogenic fire ignitions on the peat-swamp landscape in Kalimantan, Indonesia. *Glob. Environ. Change* **2016**, *39*, 205–219, doi:10.1016/j.gloenvcha.2016.05.005.
33. Hantson, S.; Arneth, A.; Harrison, S.P.; Kelley, D.I.; Prentice, I.C.; Rabin, S.S.; Archibald, S.; Mouillot, F.; Arnold, S.R.; Artaxo, P.; et al. The status and challenge of global fire modelling. *Biogeosciences* **2016**, *13*, 3359–3375.
34. Dadashi, Sepideh What is a fire? Identifying individual fire events using the MODIS burned area product. Masters thesis, University of Colorado Boulder: Boulder, CO, 2018.
35. Meddens, A.J.H.; Kolden, C.A.; Lutz, J.A.; Abatzoglou, J.T.; Hudak, A.T. Spatiotemporal patterns of unburned areas within fire perimeters in the northwestern United States from 1984 to 2014: *Ecosphere* **2018**, *9*, doi:10.1002/ecs2.2029.
36. Chuvieco, E.; Lizundia-Loiola, J.; Lucrecia Pettinari, M.; Ramo, R.; Padilla, M.; Tansey, K.; Mouillot, F.; Laurent, P.; Storm, T.; Heil, A.; et al. Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. *Earth Syst. Sci. Data* **2018**, *10*, 2015–2031, doi:10.5194/essd-10-2015-2018.
37. Giglio, L.; Boschetti, L.; Roy, D.P.; Humber, M.L.; Justice, C.O. The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sens. Environ.* **2018**, *217*, 72–85, doi:10.1016/j.rse.2018.08.005.
38. Giglio, L.; Schroeder, W.; Justice, C.O. The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sens. Environ.* **2016**, *178*, 31–41, doi:10.1016/j.rse.2016.02.054.
39. Andison, D.W. The influence of wildfire boundary delineation on our understanding of burning patterns in the Alberta foothills. *Can. J. For. Res.* **2012**, *42*, 1253–1263, doi:10.1139/X2012-074.
40. Archibald, S.; Roy, D.P.; van Wilgen, B.W.; Scholes, R.J. What limits fire? An examination of drivers of burnt area in Southern Africa. *Glob. Change Biol.* **2009**, *15*, 613–630, doi:10.1111/j.1365-2486.2008.01754.x.
41. Eidenshink, J.; Schwind, B.; Brewer, K.; Zhu, Z.-L.; Quayle, B.; Howard, S. A project for monitoring trends in burn severity. *Fire Ecol.* **2007**, *3*, 3–21.
42. Andela, N.; Morton, D.C.; Giglio, L.; Paugam, R.; Chen, Y.; Hantson, S.; van der Werf, G.R.; Randerson, J.T. The Global Fire Atlas of individual fire size, duration, speed and direction. *Earth Syst. Sci. Data* **2019**, *11*, 529–552, doi:10.5194/essd-11-529-2019.
43. Frantz, D.; Stellmes, M.; Röder, A.; Hill, J. Fire spread from MODIS burned area data: obtaining fire dynamics information for every single fire. *Int. J. Wildland Fire* **2016**, *25*, 1228, doi:10.1071/wf16003.
44. USDA Forest Service Fire Terminology Available online: https://www.fs.fed.us/nwacfire/home/terminology.html.
45. Picotte, J.J.; Peterson, B.; Meier, G.; Howard, S.M. 1984-2010 trends in fire burn severity and area for the conterminous US. *Int. J. Wildland Fire* **2016**, *25*, 413–420, doi:10.1071/WF15039.
46. Randerson, J.; Chen, Y.; Van Der Werf, G.; Rogers, B.; Morton, D. Global burned area and biomass burning emissions from small fires. *J. Geophys. Res. Biogeosciences* **2012**, *117*.
47. Short, K.C. Sources and implications of bias and uncertainty in a century of US wildfire activity data. *Int. J. Wildland Fire* **2015**, *24*, 883–891, doi:10.1071/WF14190.
48. Laurent, P.; Mouillot, F.; Yue, C.; Ciais, P.; Moreno, M.V.; Nogueira, J.M.P. Data Descriptor: FRY, a global database of fire patch functional traits derived from space-borne burned area products. *Sci. Data* **2018**, *5*, 1–12, doi:10.1038/sdata.2018.132.
49. GeoMAC Geospatial Multi-Agency Coordination Available online: www.geomac.gov

(accessed on Oct 9, 2019).
50. St. Denis, L.A.; Mietkiewicz, N.P.; Short, K.C.; Buckland, M.; Balch, J.K. All-hazards dataset mined from the US National Incident Management System 1999–2014. *Sci. Data* **2020**, *7*, 64, doi:10.1038/s41597-020-0403-0.
51. St. Denis, L.; Hughes, A.; Diaz, J.; Solvik, K.; Joseph, M. "What I Need to Know is What I Don't Know!": Filtering Disaster Twitter Data for Information from Local Individuals. In Proceedings of the Proceedings of 17th International Conference on Information Systems for Crisis Response and Management; Blacksburg, VA USA, 2020.
52. Diaz, J.; St. Denis, L.; Joseph, M.; Solvik, K. Classifying Twitter Users for Disaster Response: A Highly Multimodal or Simple Approach? In Proceedings of the Proceedings of 17th International Conference on Information Systems for Crisis Response and Management; Blacksburg, VA USA, 2020.
53. Fusco, E.J.; Finn, J.T.; Abatzoglou, J.T.; Balch, J.K.; Dadashi, S.; Bradley, B.A. Detection rates and biases of fire observations from MODIS and agency reports in the conterminous United States. *Remote Sens. Environ.* **2019**, *220*, 30–40.