https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Open Access Earth System Science Data Discussions

A Global Dataset of Individual Fire Events (2012-2025) derived from VIIRS Active Fire Product

Hongxuan Su[1], Yan Yu[1,2,3]

[1]Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China
[2]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China
[3]China Meteorological Administration Tornado Key Laboratory, Beijing, China

Correspondence to: Yan Yu (yuyan@pku.edu.cn)

Abstract. Accurate characterization of individual fire events is critical for understanding global fire regimes and their interactions with climate. However, existing global inventories relying on coarse-resolution data and annual-aggregation algorithm may obscure small-scale thermal anomalies and artificially fragment prolonged fires. Here, we present a global dataset of individual fire events for 2012–2025 derived from Suomi-NPP VIIRS 375 m active fire detections. To address the spatiotemporal discontinuities in conventional clustering approaches, we develop a tracking framework based on a sliding-window method that merges spatially overlapped and temporally continuous fire patches. This approach maintains the continuity of long-duration events and reduces artificial segmentation, as demonstrated by validation against ground-based and higher-resolution satellite benchmarks. Our dataset provides a range of event-level metrics, including ignition location, fire duration, daily expansion rates, burned area, and fire radiative power. This fire event inventory offers an observational foundation for tracking extreme fire behavior, calibrating fire spread models, and refining global emissions estimates.

1 Introduction

Fire is an essential disturbance process in the Earth system, influencing ecosystem structure, carbon cycling and atmospheric composition. Global fire monitoring has relied on grid-based active fire detections or burned area (BA) products (Giglio et al., 2013; Chuvieco et al., 2008). These datasets have provided valuable information on estimating global burned area trends, but they generally describe fire at the grid scale rather than at the level of individual fire events(Archibald et al., 2013; Descombes et al., 2018). In recent years, more studies have adopted event-based approaches that group fire detections into individual fire events. One example is the Global Fire Atlas(Andela et al., 2019), which enables the derivation of event-level characteristics such as fire duration, fire size and spread rate. These metrics are essential not only for advancing process-level knowledge about the life cycle of wildfire (Hantson et al., 2022; Su et al., 2025) but also for evaluating the parameterization of global fire models. For example, by providing independent observational constraints on fire duration and daily expansion rates, these metrics may help improve the representation of fire processes in Earth system models(Rabin et al., 2017).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Event-based wildfire datasets provide valuable insights into fire behaviors; however, current global inventories are primarily derived from the MODIS MCD64A1 product (500 m resolution). While the MODIS record is foundational for long-term studies, its relatively coarse spatial resolution limits the detection of small-scale fires, which are recognized as significant contributors to emissions in fragmented landscapes and tropical forests(Ramo et al., 2021; Schroeder et al., 2014). The Suomi-NPP VIIRS sensor, featuring a 375 m spatial resolution and higher sensitivity to thermal anomalies, detects up to 3–25 times more fire pixels than MODIS depending on the region and time of day(Schroeder et al., 2014). Recent regional studies, such as the Fire Event Data Suite (FEDS) for North America(Chen et al., 2022), the near-real time classifying approach for Amazon (Andela et al., 2022) and multi-ignition fire complexes for boreal ecoregions(Scholten et al., 2026), have demonstrated the utility of VIIRS for tracking fire events. However, a consistent, long-term global inventory of individual fire events derived from VIIRS 375 m data remains unavailable.

Furthermore, temporal segmentation remains a methodological challenge when delineating fire events. To manage computational complexity, existing approaches typically aggregate detections into annual layers(Andela et al., 2019; Frantz et al., 2016). While computationally efficient, this strategy can conflate distinct fires occurring in the same location across different seasons (e.g., merging a spring agricultural fire with an autumn wildfire), thereby obscuring intra-annual reburn dynamics. Conversely, monthly partitioning may fragment long-duration fires into discontinuous segments when they span across month boundaries. As highlighted by(Balch et al., 2020), establishing a flexible temporal aggregation framework is an important step towards preserving the full lifecycle of fire events while accounting for reburns.

To address these limitations, we presented the Global Dataset of Individual Fire Events (2012–2025), derived from the VIIRS S-NPP 375 m active fire product. We developed a spatiotemporal tracking framework that maps active fire detections onto the 0.005° grid. A sliding-window approach combined with cross-timestep checks was used to link detections through months. Compared with conventional annual aggregation methods, this approach distinguished separate reburning events within a year, while maintaining the continuity of long-duration wildfires based on spatiotemporal overlap. The resulting dataset provided event-level information for millions of fires, including key metrics describing fire behavior. It can support studies of fire regime variability and related environmental impacts.

2 Data and methods

2.1 VIIRS data acquisition

We developed a global dataset of individual fire events covering all terrestrial regions (from 90°N to 90°S, 180°W to 180°E), based on the Visible Infrared Imaging Radiometer Suite (VIIRS) Active Fire Product obtained by the Suomi National Polar-orbiting Partnership (S-NPP) satellite. Specifically, we utilized the VIIRS S-NPP Level-2 active fire data(Schroeder et al., 2014), which provided precise thermal anomalies, including fire spot location, brightness temperature, and Fire Radiative Power (FRP). Compared to the MODIS active fire products (~ 1 km resolution), the finer spatial resolution of VIIRS (~375

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

m) significantly enhanced the capability of detecting small-scale fires and improves the delineation of fire perimeters. While NOAA-20 and NOAA-21 provided supplemental data, we prioritized the S-NPP yearly CSV files for their continuous long-term coverage from January 20, 2012, to the present. The raw active fire records (2012–2025) were acquired from the NASA Fire Information for Resource Management System (FIRMS).

2.2 Pre-processing and gridding

Raw VIIRS active fire detections were first filtered to exclude non-vegetation thermal anomalies. This was achieved by leveraging the native classification attribute (Type) provided in the standard VIIRS active fire product(Schroeder et al., 2025). Only thermal anomalies classified as presumed vegetation fires (Type = 0) were retained. Detections flagged as active volcanoes (Type = 1), other static land sources (Type = 2), and offshore detections (Type = 3) were discarded to prevent industrial and geological artifacts from entering the processing pipeline.

To further filter persistent industrial thermal anomalies (e.g., unclassified gas flaring and factory heat signatures) that may remain in the native product classification, a secondary static source mask was applied. Specifically, 0.005° spatial grids recording more than 90 thermal anomalies per year were identified as static sources and subsequently masked out.

Following these filtration steps, the retained vegetation fire detections were temporally partitioned into monthly intervals to reduce the limitations of traditional yearly aggregating. A previous study (Balch et al., 2020) has noted that yearly aggregations can artificially merge distinct fire events occurring in the same location into a single record, thereby obscuring intra-annual reburn dynamics and underestimating fire frequency. Conversely, monthly partitioning may fragment single continuous fire events that span across month boundaries. To address these methodological challenges, we coupled a monthly time-step with a cross-timestep validation process (see 2.3). This approach mitigated the artifacts introduced by annual aggregation while preserving the temporal continuity of long-duration fires.

Then these valid fire detections were aggregated onto an equidistant global grid with a resolution of 0.005° (approximately 550 m at the equator). For grid cells containing multiple fire records, we retained the earliest burn date (Fig. 1a).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

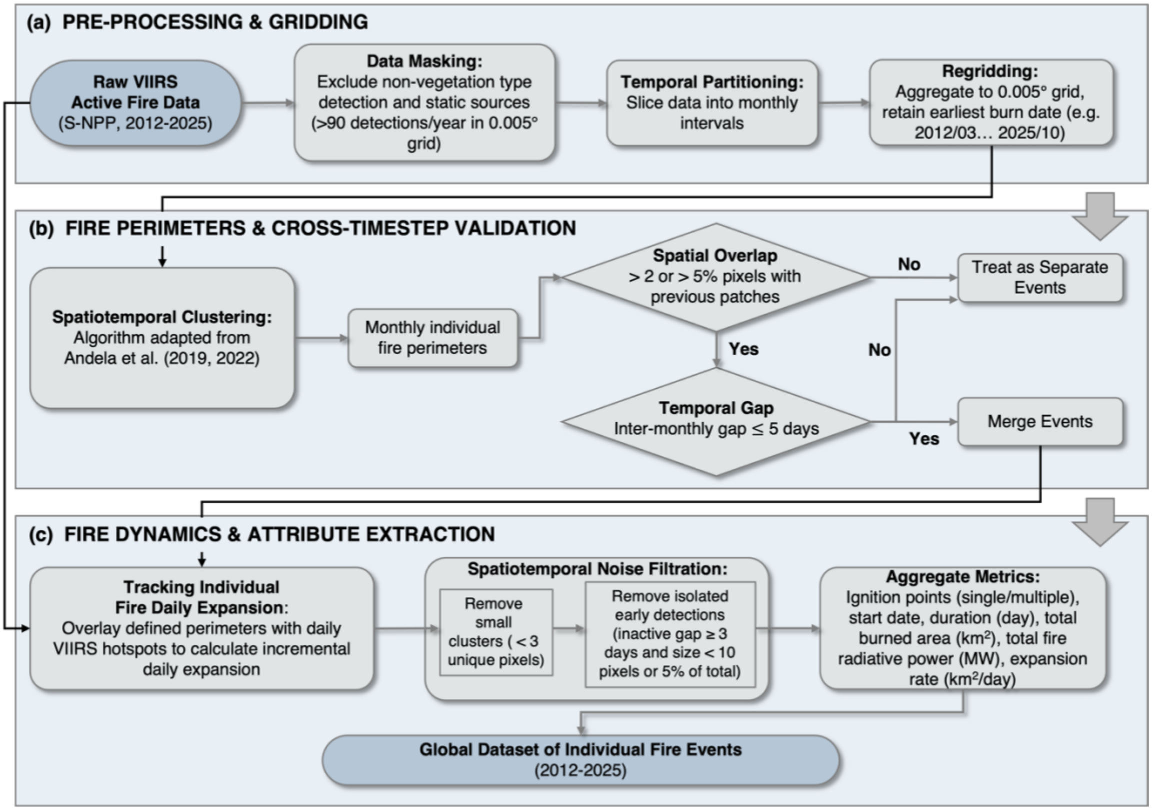


Figure 1. Methodological workflow for tracking individual wildfires and extracting fire attributes. The algorithm processed raw VIIRS active fire data through three main stages: **(a)** pre-processing and gridding to filter static noise and standardize spatial resolution to 0.005°; **(b)** fire perimeter delineation and cross-timestep validation to cluster monthly fire patches and merge cross-month events based on defined spatiotemporal thresholds; and **(c)** fire dynamics tracking and attribute extraction to reconstruct daily expansion trajectories and calculate fire metrics for the global dataset.

2.3 Fire perimeter delineation and cross-timestep merging

To delineate the spatial boundaries of individual fire events, we applied a spatiotemporal clustering algorithm adapted from previous studies(Andela et al., 2022; Andela et al., 2019). The algorithm was first applied to the active fire data to generate individual fire perimeters on a monthly basis. However, a known issue in fire tracking was the fragmentation of long-duration wildfires that extended across multiple months.

To reduce this artifact, we introduced a cross-timestep merging procedure. Fire patches identified in consecutive months were evaluated for their spatial and temporal linkage: (1) the spatial overlap between the two patches exceeded either two pixels or 5% of the patch area; (2) the time gap between consecutive active fire detections in the two patches was less than or equal to five days (Fig. 1b, Fig. A1 in Appendix A). If both conditions were met, the patches were merged into a

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

single fire event. Otherwise, they were treated as separate events (Fig. 1b). This approach helped maintain the continuity of long-duration fires while limiting the merging of nearby but temporally distinct events.

To assess the robustness of the cross-month aggregation settings, we conducted a one-factor-at-a-time sensitivity analysis over CONUS for 2020. The maximum temporal gap was varied from three to seven days while the minimum overlap was fixed at two grid cells; the minimum overlap was varied from one to three grid cells while the temporal gap was fixed at five days. The temporal gap mainly affected the number of cross-month events, whereas total event count, median duration, and median retained grid-cell count changed by less than 0.5% across the tested range (Fig. C1; Table C2). We therefore retained a five-day temporal gap and a two-grid-cell overlap as balanced baseline settings that limit artificial fragmentation without adopting a more permissive merging rule.

2.4 Daily expansion tracking and spatiotemporal noise filtration

Following the delineation of fire event perimeters, we derived dynamic attributes by overlaying the perimeters with daily VIIRS active fire detections to estimate daily fire growth (Fig. 1c).

To improve the consistency of the reconstructed fire trajectories, a two-step spatiotemporal filtration procedure was applied. First, a spatial filter was used to remove small clusters consisting of fewer than three unique pixels, as such detections were more likely to represent transient signals rather than sustained fires. Second, a temporal filter was applied to remove early isolated detections. These detections were excluded if they represented only a small fraction of the total event size (< 10 pixels or < 5%) and were separated from the main fire activity by a gap of at least 3 days. This step helped reduce the influence of short-lived signals (Fig. 1c).

For each retained fire event, we derived ignition location, start date, fire duration (days), total burned area (km^2), total fire radiative power (FRP; MW) and expansion rate (km^2/day). To assess the sensitivity of area estimation to spatial delineation, we additionally evaluated three alternative approaches: pixel accumulation, convex hull and concave hull. Pixel accumulation estimated area by summing the retained fire pixels and therefore provided a conservative estimate of the observed active-fire footprint (Fig. 2a). The convex-hull approach constructed the minimum convex polygon enclosing the fire pixels (Pateiro-L´Opez and Rodr´Iguez-Casal, 2010) and overestimated burned area by including unburned land (Fig. 2b). The concave-hull approach, using the GEOS library through Python Shapely package(Gillies et al., 2025), can better preserve complex fire-front geometry (Fig. 2c) but required parameter selection and robust geometric processing.

The released total area (km^2) variable was calculated by pixel accumulation from retained VIIRS active-fire detections. This approach was selected for the global product because it was computationally efficient, projection-independent, and consistently applicable across all regions. It provides a conservative estimate of the observed active-fire footprint. Convex-hull and concave-hull estimates were evaluated as alternative spatial delineation methods but were not included in the global event table because their robust implementation required event-specific projection handling and parameter choices. The final

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

dataset provided a global dataset of 15,377,087 individual fire events, from March 2012 to December 2025, capturing both the spatial distribution of fires and their temporal evolution.

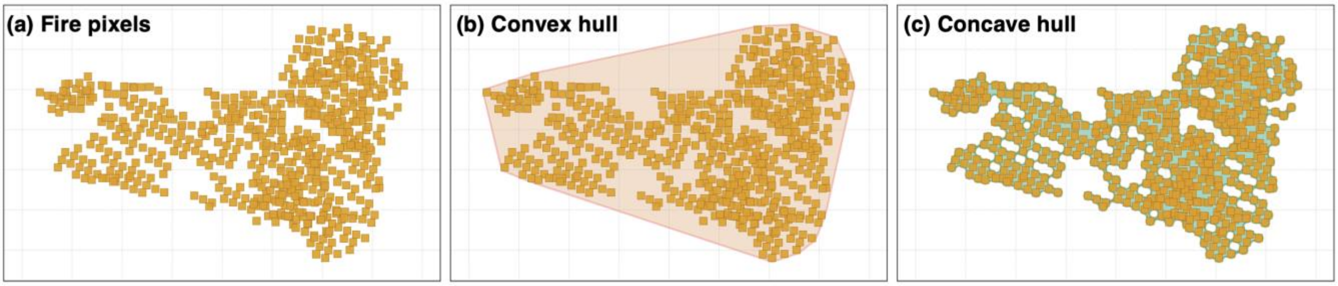


Figure 2. Conceptual illustration of burned area estimation methods from discrete VIIRS active fire detections. (a) Raw VIIRS active fire pixels (375 m resolution) exhibited a spatial discontinuity. **(b)** The convex hull approach encompassed unburned regions, leading to burned area overestimation. **(c)** The concave hull approach, which bridged spatial gaps while preserving the complex fire fronts.

2.5 Accuracy assessment

To evaluate the accuracy of our VIIRS-derived individual fire dataset, we conducted fine-scale progression analysis of two wildfire case studies in the United States and statistical evaluations across three diverse regions.

2.5.1 Case study

Daily incident perimeters from the California Department of Forestry and Fire Protection (CAL FIRE) (California Department of Forestry and Fire Protection, 2026) and the Monitoring Trends in Burn Severity (Eidenshink et al., 2007) were acquired to demonstrate the capability of our dataset in capturing fine-scale fire behavior. Moreover, we obtained the daily accumulated area data from official reports for the 2013 Rim Fire (Mymotherlode.Com, 2013) and the 2025 Palisades Fire (California Department of Forestry and Fire Protection, 2025).

Independent reference data describing daily fire progression were not consistently available across global fire regimes. We therefore conducted a case-based assessment of daily fire-growth dynamics only for events with time-resolved official cumulative-area reports. For each pair of consecutive official reports, DER (daily expansion rate) was calculated as the reported cumulative-area increment divided by the reporting interval. VIIRS DER was calculated as the spatially de-duplicated newly detected VIIRS active-fire pixel area aggregated over the identical interval and divided by its duration. We evaluated temporal correspondence using Spearman’s rank correlation coefficient, mean absolute error, root mean square error, mean bias, and peak-timing offset (Table C2).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

2.5.2 Statistical validation across three regions

We utilized five independent official wildfire datasets across three fire-prone regions, encompassing the United States, Canada and Australia (Table 1, Fig. B7). These datasets were selected for their long-term consistency and authority to represent diverse fire regions.

The Fire Program Analysis Fire-Occurrence Database (FPA-FOD) (Short, 2014; Short, 2022) was a comprehensive dataset that integrates interagency human-reported wildfire records, providing point-based fire locations (i.e., ignition coordinates), temporal metrics (e.g., discovery dates and control dates) and total burned area. The Monitoring Trends in Burn Severity (MTBS) dataset (Eidenshink et al., 2007) offered reliable polygon-based burn perimeters derived from 30 m Landsat multispectral imagery. The Canadian National Fire Database (CNFDB) (Canadian Forest Service, 2026) compiled by the Canadian Forest Service provided point-based ignition coordinates of fire occurrences and their final burned area. We also acquired official polygon-based fire history perimeters from two state agencies, the New South Wales National Parks and Wildfire Service (NPWS) (State Government of Nsw and Department of Planning Industry and Environment, 2026) and the Victoria Department of Energy, Environment and Climate Action (Department of Energy Environment and Climate Action, 2026). These datasets provided overall burn perimeters and total burned area.

Table 1. Overview of the authoritative wildfire datasets for validation. The table detailed the geographic region, dataset name, geometry type (point-based or polygon-based), evaluated metrics, temporal coverage, reference and source.

Region	Validation Dataset	Geometry Type	Evaluated Metrics	Temporal Span	Reference / Source
United States	Fire Program Analysis Fire-Occurrence Database (FPA-FOD)	Point	Duration Burned Area	2012–2020	Short (2022)
United States	Monitoring Trends in Burn Severity (MTBS)	Polygon	Burned Area	2012–2020	Eidenshink et al. (2007)
Canada	The Canadian National Fire Database (CNFDB)	Point	Burned Area	2012–2020	Canadian Forest Service (2023)
Australia	NPWS Fire History - Wildfires and Prescribed Burns (NPWS)	Polygon	Burned Area	2012–2020	State Government of NSW and Department of Planning, Industry and Environment (2026)
Australia	Fire History Records of Fires across Victoria	Polygon	Burned Area	2012-2020	Department of Energy, Environment and Climate Action (2026)

Then we developed two different methods to match these validation datasets with our VIIRS-based inventory. For point-based databases (i.e., the US FPA-FOD and the Canadian CNFDB), the geographic coordinates represented the reported ignition location. We applied a 14-day ignition gap and four-kilometer spatial threshold to match the VIIRS-derived fire event with FPA-FOD record. Specifically, a valid spatial match required the official FPA-FOD reporting point to fall either entirely within the VIIRS-derived fire boundary or at a radial distance of no more than four kilometers from the

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

polygon's outer perimeter (Fig. 3a). Meanwhile, the absolute gap between the satellite-detected start date and the official discovery date must be ≤ 14 days. In cases of the multiple pairs (e.g., multiple proximal ground records), we selected the ground records with the minimum temporal gap and spatial distance. The attributes (e.g., active-fire duration and burned area) of these successfully matched cases from both VIIRS-derived and official records were extracted for further accuracy assessment.

For polygon-based databases, including the US MTBS, the official fire history records from Australia NPWS and Victoria, we matched the datasets using a maximum ignition time gap of 14 days and a minimum spatial overlap of 0.1 km^2. Spatially, a valid match was established if the VIIRS-derived boundary physically intersected with the official perimeter. Then a minimum overlapping area threshold of 0.1 km^2 was enforced under the same ≤ 14 days temporal constraints (Fig. 3b). If an official record intersected with multiple VIIRS-derived events, the pairing with the maximum spatial overlap was selected.

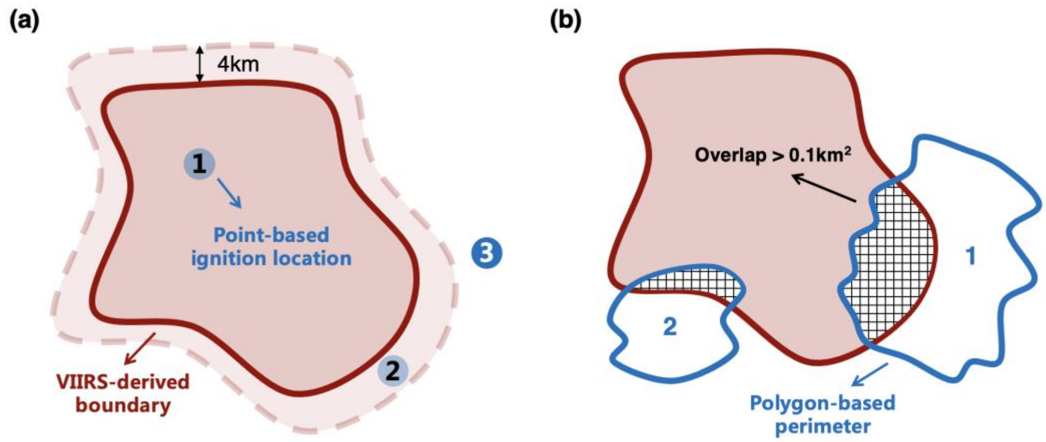


Figure 3. Conceptual diagrams of the spatial matching method for validation. (a) The point-based matching strategy, where a valid match required the official reported ignition location (blue points) to fall either entirely within the VIIRS-derived fire boundary (solid red line) or its four-km radial buffer zone (dashed pink area). **(b)** The polygon-based matching strategy, which required topological intersection between the VIIRS-derived boundary and the official reported perimeter (blue outline), enforcing a minimum overlapping area threshold of 0.1 km^2. Both matching methods operated under a concurrent temporal constraint, requiring an absolute difference of ≤ 14 days between the VIIRS-detected and official fire start dates.

2.5.3 Accuracy metrics

We employed the Precision-Recall analysis framework to evaluate the detection performance of our VIIRS-derived inventory. The Precision-Recall evaluation relied on the matching results based on the spatiotemporal criteria (≤ 14 days ignition gap and specific spatial threshold). A VIIRS-derived fire event that successfully matched with an official record was defined as true positives (TP) and a VIIRS-derived fire event that lacked any corresponding official record was defined as

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

false positives (FP). While an official record that our algorithm failed to detect was defined as false negatives (FN). Precision rate and recall rate were calculated as:

$$\text{Precision rate} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall rate} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Precision date indicated the probability that a VIIRS-derived fire was actually reported by agencies, while recall measures the proportion of actual fires successfully captured by VIIRS. To assess the performance of detection accuracy, we calculated both the precision and recall rate based on varying fire size thresholds.

Following the detection accuracy, we further evaluated the performance of our VIIRS-derived burned area and active-fire duration using statistical indicators, including the coefficient of determination (R2) evaluated in log-log space, root mean square error (RMSE), and median absolute error (MedAE). To filter out transient anomalies, these evaluations were only conducted in a matched subset with burned area > 1 km^2 and active-fire duration > 3 days.

2.6 Dataset variables, recommended use, and quality information

Each record in the released dataset represents one reconstructed individual fire event. The core event-level variables include a unique identifier (fire_id), VIIRS-detected ignition location (ign_lat and ign_lon), start and end dates (start_date and end_date), fire duration (duration_days), observed detection days (active_days), pixel-accumulated area (area_km2), daily expansion metrics (mean_der, peak_der, and explosive_idx), cumulative fire radiative power (frp_mw), and a dominant GFED5 16-class fire-type code (Chen et al., 2023). Table 2 summarizes the core variables, their units, and recommended interpretation. A complete variable-level data dictionary, including calculation procedures and quality-information definitions, is provided in Table C1 and the repository README.

The released global event dataset contains one primary area variable, area_km2, calculated by pixel accumulation from unique retained VIIRS active-fire detections. It provides a conservative and globally internally consistent estimate of the observed active-fire footprint and is the area metric used throughout this study unless otherwise stated. Convex-hull and concave-hull approaches are evaluated separately to assess methodological sensitivity to spatial delineation, but are not included in the released global event table because their robust global implementation requires event-specific projection handling and parameter choices.

The dataset provides event-level quality information describing observation continuity and algorithmic conditions. The variable n_pixels report the number of unique retained fire-pixel locations associated with an event, and max_gap_days reports the largest number of consecutive days without a retained VIIRS detection between two observed event days. The variable cross_month indicates whether an event spans more than one calendar month. The temporal_flag is set to 1 when max_gap_days is four days or longer and to 0 otherwise.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

The quality_flag combines predefined warning conditions using a bit-mask convention. A value of 1 indicates limited spatial support (n_pixels < 5), a value of 2 indicates a high temporal observation gap (max_gap_days ≥ 4), and a value of 4 indicates an event spanning more than one calendar month (cross_month = 1). These codes are combined by addition. For example, a value of 6 indicates both a high temporal-gap warning and a cross-month event. A value of 0 indicates that no predefined algorithmic warning was triggered; it does not imply zero measurement error.

In addition, the dataset provides regression-derived reference-equivalent estimates and 90 % prediction intervals for pixel-accumulated area (area_ref_km2, area_p90_lo, and area_p90_hi) and calendar-span duration (duration_ref_days, duration_p90_lo, and duration_p90_hi). The fields area_extrap, duration_extrap, and reg_flag identify whether a prediction is within the corresponding calibration range. These regression-derived fields quantify agreement with the available MTBS and FPA-FOD validation samples and are conditional on their United States coverage, reference definitions, matching procedure, and sampling characteristics. They should not be interpreted as globally uniform estimates of true event-level error.

Table 2. Principal variables in the global individual-fire-event dataset and their recommended interpretation.

Variable group	Variable(s)	Unit	Description and recommended interpretation
Event identification	fire_id	–	Unique identifier for each reconstructed individual fire event.
Event timing and ignition location	start_date; end_date; ign_lat; ign_lon	date; °N; °E	VIIRS-detected start and end dates and the centroid of fire pixels detected on the first day of the event, representing the initial VIIRS-detected ignition location.
Fire duration and observation continuity	duration_days; active_days; max_gap_days	days	Calendar duration of the reconstructed event, number of days with active-fire detections, and maximum gap between successive detection days.
Burned area	area_km2	km²	VIIRS-derived burned area calculated from unique fire pixels. This metric provides a conservative estimate of the area with observed thermal activity.
Fire expansion	mean_der; peak_der; explosive_idx	$km^2\ d^{-1}$; ratio	Mean and peak daily expansion rates and the proportion of total expansion occurring on the day of maximum growth.
Fire radiative power	frp_mw	MW	Total fire radiative power (FRP) recorded for all retained VIIRS active-fire detections within the event.
Fire type	gfed5_type	categorical code	Dominant fire type among event pixels based on the 16 fire types defined by GFED5.
Quality information	N_pixels; cross_month; temporal_flag; quality_flag	pixels; categorical	Indicators of spatial support, temporal continuity, and whether the reconstructed event extended across calendar months.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Variable group	Variable(s)	Unit	Description and recommended interpretation
Validation-derived uncertainty	area_ref_km2; area_p90_lo; area_p90_hi; area_extrap; duration_ref_days; duration_p90_lo; duration_p90_hi; duration_extrap; reg_flag	km²; days; categorical	Regression-derived estimates, 90 % prediction intervals, and extrapolation flags based on matched US MTBS and FPA-FOD records.

3 Results

3.1 Case-based quantitative evaluation of daily fire-growth dynamics

To explore the capacity of our VIIRS-derived dataset in capturing fine-scale spatiotemporal dynamics, we presented diagnostic analyses of two impactful fire events in California: the 2013 Rim Fire and the 2025 Palisades Fire (Fig. 4).

As illustrated in the spatial pattern (Fig. 4a and 4c), our fire dataset generally reconstructed the daily expansion of both events. Spatially, our dataset captured the progression of burn date, radiating from the initial VIIRS-detected ignitions (cyan stars). It also provided a reasonable burn perimeter that aligned closely with the authoritative spatial boundaries provided by MTBS and CAL FIRE. Temporally, the dataset exhibited the capability to track active burning area (Fig. 4b and 4d) and identify rapid daily expansion. For example, the 2013 Rim Fire expanded rapidly within a week after ignition and the newly burned area from August 21 to 22 exceeded 40 km^2 (yellow bars). Indeed, the expansion during the first week (August 17 and 23) accounted for nearly half of the total burned area. While the 2025 Palisades Fire experienced a rapid expansion within the first two days. The newly burned area exhibited an approximate seven-fold increase within 24 hours from January 7 to 8.

We compared three area estimation results, including pure pixel (solid blue line), convex hull (dotted orange line) and concave hull (dashed green line). Overall, all three VIIRS-derived area metrics successfully captured the spreading dynamics of both fires (Fig. 4b and 4d). Their performances diverged depending on the fire characteristics. These three methods remained relatively close during the initial stages of the fire while showed divergence as the fire progressed. The convex hull approach overestimated the burned area, whereas both the pure pixel and concave hull methods provided more conservative estimates. Notably, the pure pixel approach yielded the lowest overall burned area in both two cases.

The case study also revealed differences between satellite observations and human reports. During the 2013 Rim Fire, the official report indicated that fire expansion ceased at September 14 (Fig. 4b, black lines with stars). However, VIIRS continuously detected active thermal anomalies for an additional week, terminating on September 21. This extension highlighted the VIIRS sensor's ability to monitor residual smoldering and unextinguished interior hotspots long after the primary fire front has been administratively claimed controlled. Conversely, the Palisades Fire exhibited a prolonged plateau in the officially reported area weeks after VIIRS-detected expansion ceased (Fig. 4d, black lines with starts), reflecting the administrative containment. These contrasting scenarios indicated that the VIIRS-derived metrics offered a representation of active thermal events, complementing administrative reporting records.

The case-based comparison indicated that the VIIRS reconstruction captured the temporal ordering of fire-growth activity in the selected events (Table C2). For the Rim Fire, the rank correlation between VIIRS-derived and official DER was 0.714 across 28 aligned reporting intervals. For the Palisades Fire, the corresponding correlation was 1.000 across 22 aligned intervals. The maximum VIIRS-derived and official DER occurred within the same reporting interval for both cases (peak timing offset is zero). However, the non-zero MAE, RMSE, and mean bias values indicated differences in the estimated magnitude of daily expansion. For the Rim Fire, VIIRS-derived peak DER was 110.531 km^2 d^{-1}, compared with 324.125 km^2 d^{-1} in the official series. For the Palisades Fire, the corresponding values were 39.375 and 64.636 km^2 d^{-1}, respectively.

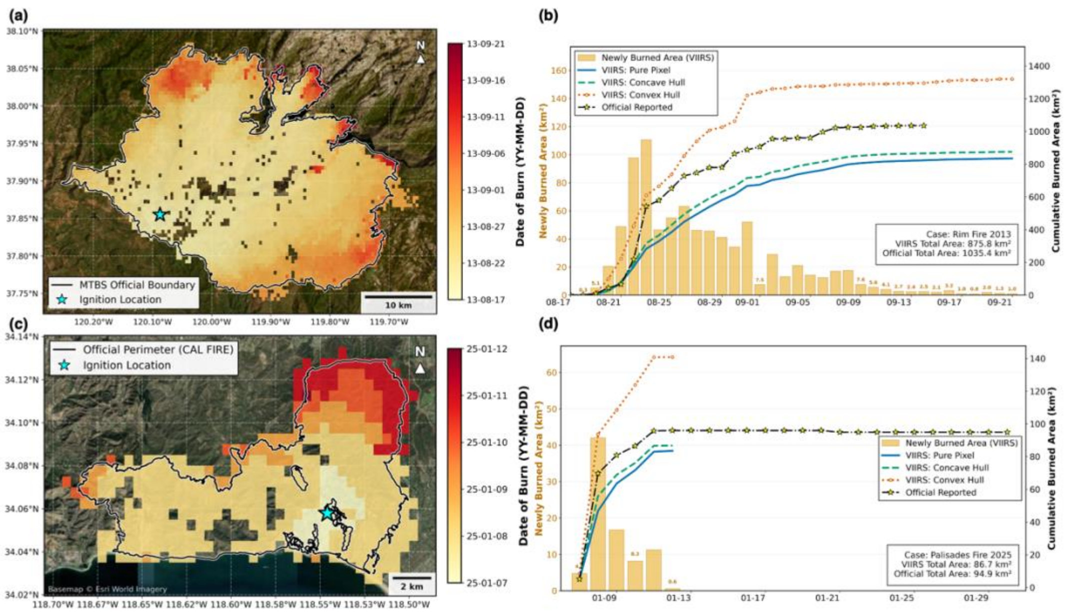


Figure 4. Spatiotemporal validation of the VIIRS-derived reconstruction against two independent data records of two large fires. Cases were the 2013 Rim Fire (top row) and the 2025 Palisades Fire (bottom row). **(a) (c),** VIIRS-derived spatial progression overlaid on Esri World Imagery. Pixel colours indicated the localized date of burn, progressing from initial outbreaks (light yellow) to later stages (dark red). Cyan stars denoted VIIRS-detected ignitions, and solid black lines represented authoritative boundaries from MTBS (a) and CAL FIRE (c). The basemaps were designed and developed by Esri | Powered by Esri. **(b) (d),** temporal evolution of burning area. Yellow bars (left y-axis) illustrated the VIIRS-derived daily active burned area based on fire pixels. Lines (right y-axis) contrasted the cumulative burned area derived from three VIIRS-derived metrics, including pure pixel (solid blue), concave hull (dashed green), and convex hull (dotted orange), in comparison with the officially reported cumulative area (black line with stars).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

3.2 Accuracy assessment

3.2.1 Detection accuracy

The evaluation of our data's detection accuracy against four official fire datasets revealed a size-dependent performance curve (Fig. 5). When we included small fires (e.g., < 4 km^2), the precision is relatively low, starting near 0.1. Meanwhile, the recall remained relatively higher even at small thresholds. It suggested that the VIIRS-derived inventory contained more small burning events than official records, as supported by the contrasted sample sizes within the VIIRS-derived and official inventories. Indeed, the sampling bias was physically and administratively justifiable. At its 375 m nadir resolution, the VIIRS sensor captured a vast array of minor thermal anomalies. However, the accuracy of official records was limited by various reasons. For example, the MTBS program intentionally applied certain inclusion criteria, mapping only sustained large wildfires (typically ≥ 1000 acres in the western US and ≥ 500 acres in the eastern US)(Eidenshink et al., 2007). Although our pre-processing (section 2.2) and spatiotemporal filtering (section 2.4) attempted to minimize the contamination of agricultural and industrial burning, complete exclusion of small-scale noises from the VIIS-based inventory especially at the 375-m resolution remained practically challenging. Consequently, our dataset identified thousands of thermal events that were excluded by official databases, resulting in a low precision at these small scales.

Conversely, as the fire size threshold increased to encompass large fire events, the detection accuracy improved. In response to elevated size threshold and resultant elimination of small thermal anomalies, the precision curve climbed steadily and reached around 0.55 for FPA-FOD (Fig. 5a), 0.8 for CNFDB (Fig. 5c), 0.75 for Australia (Fig, 5d) for fires larger than 10 km^2. Meanwhile, the recall surged alongside the size threshold, ultimately exceeding 0.95 for these mega-fires in MTBS (Fig. 5b). This high recall indicated that our VIIRS-derived dataset successfully captured virtually every single major wildfire documented by the MTBS program.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

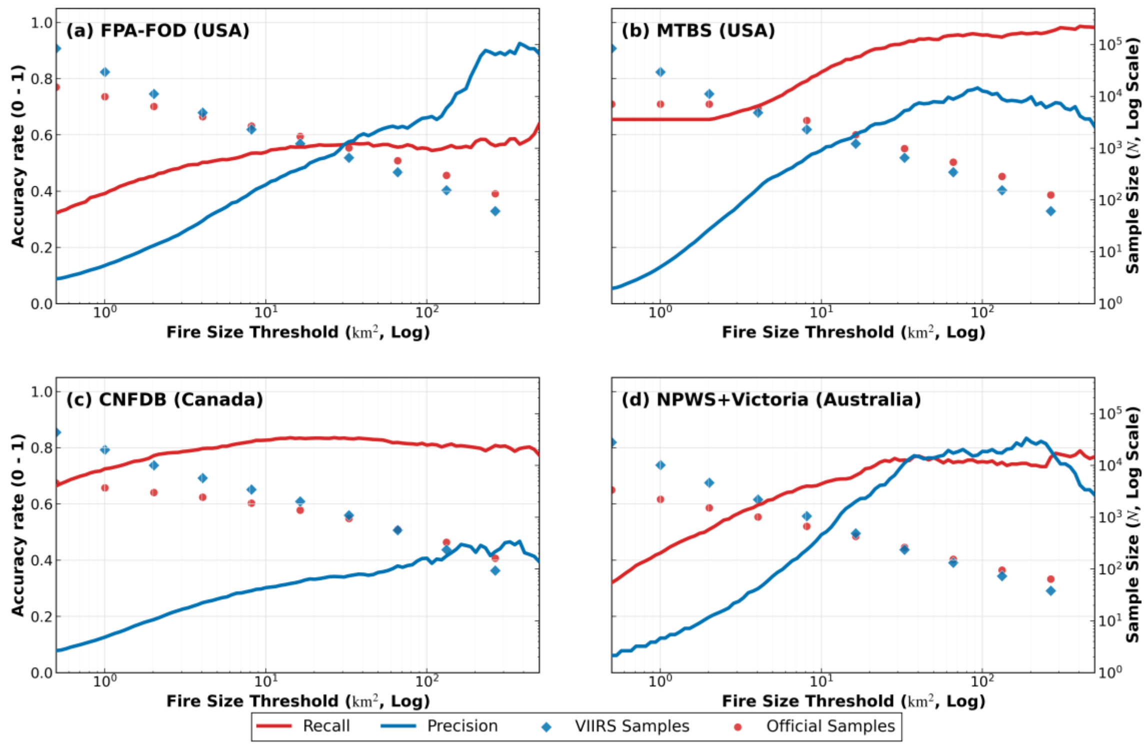


Figure 5. Validation of the VIIRS-derived fire dataset against four regional records across varying lower threshold of size. The panels illustrated the variation in detection accuracy metrics across (a) FPA-FOD, (b) MTBS, (c) CNFDB, (d) NPWS and Victoria. Solid lines demonstrated the recall (red lines) and precision (blue lines), referring to the left y axis. Scatter dots presented the retained sample size of the VIIRS-derived events (blue diamonds) and officially reported events (red circles), referring to the right y axis (log scale), at different size thresholds (x-axis, log scale).

3.2.2 Burned area accuracy

Overall, the VIIRS-derived burned areas were consistent with official records across the evaluated regions (Fig. 6). The correlation was strongest (log-R^2=0.781) with the MTBS dataset in the United States across the 1,392 matched wildfires (Fig. 6b), and remained significant for FPA-FOD (log-R^2=0.635, N =1,461, Fig. 6a), CNFDB (log-R^2=0.500, N =737, Fig. 6c) and Australia (log-R^2=0.219, N =707, Fig. 6d).

Generally, both the median and mean biases were negative across all four regional databases, indicating that VIIRS-derived burned area tended to underestimate the officially recorded burn area. (Fig. 6). The impact of a few extreme mega-fires caused the mean bias and RMSE to a large value. For example, the mean bias inflated to -32.42 km^2 and RMSE to 127.89 km^2 for MTBS (Fig. 6b). While the median bias and median absolute error (MedAE) showed -4.23 km^2 and 5.49 km^2, respectively, confirming that the baseline deviation for the vast majority of fire events remained constrained.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Notably, the VIIRS-derived dataset underestimated burned area especially for large fires (>1000 km^2) in Australia (Fig. 6d). It was likely because multiple fire events matched with one officially recorded event. In case of official fire complexes that grouped distinct fire patches, we summed the areas of multiple VIIRS-derived events that matched with a single official perimeter. Following the aggregation, the correlation (log-R^2) remained consistent for the datasets in the USA and Canada, whereas it exhibited a large improvement for the Australian region (log-R^2 = 0.655, Fig. B1). For example, the case identified as ID 32933 in the NPWS dataset corresponded to 16 fire events in VIIRS-derived dataset (Fig. B2). The aggregated area was about 1700 km^2, still smaller than the officially recorded area 2884.9 km^2.

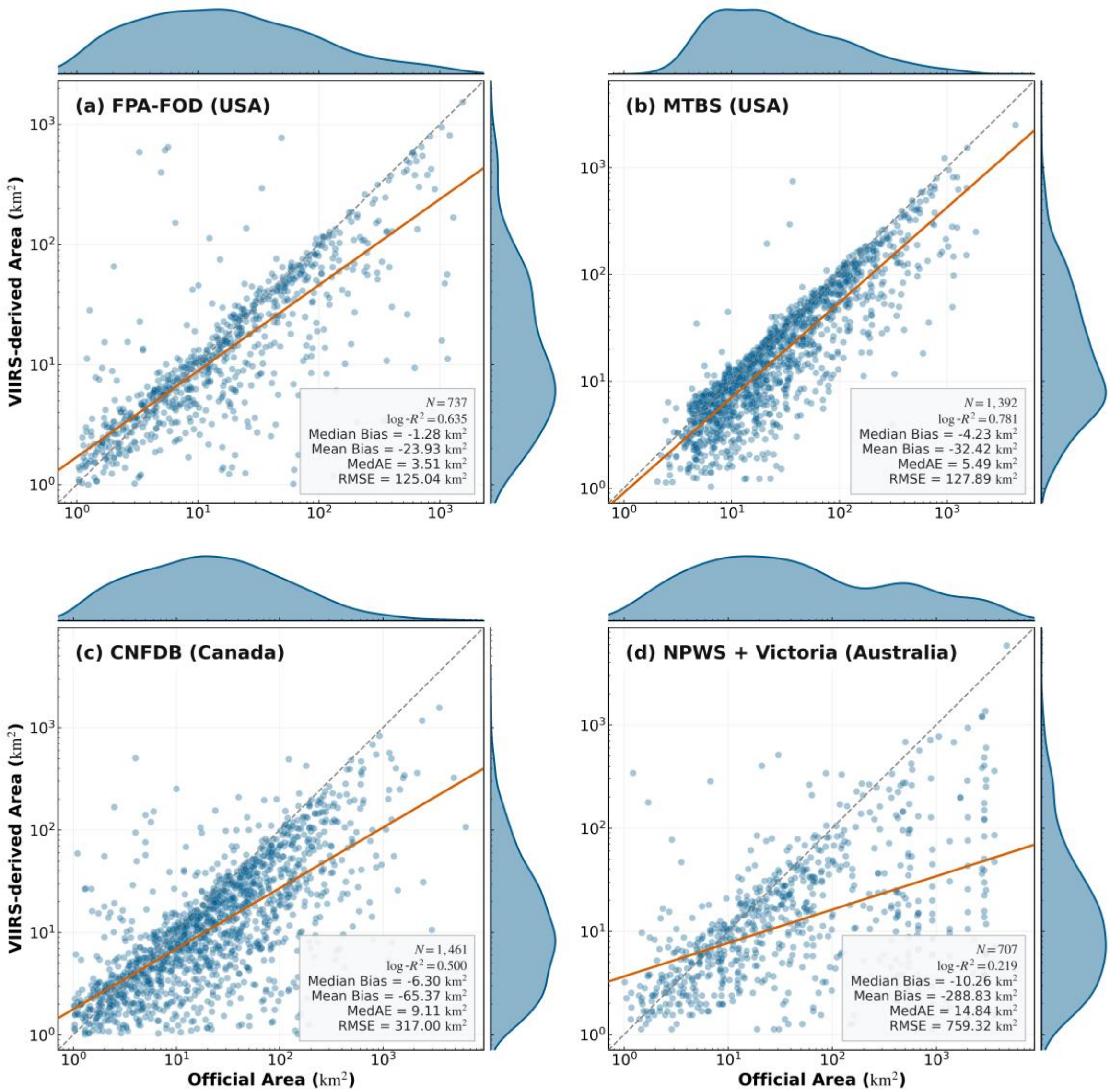


Figure 6. Validation of VIIRS-derived burned area against official burned area from four regional databases. The panels displayed scatter plots for **(a)** FPA-FOD (US), **(b)** MTBS (US), **(c)** CNFDB (Canada), and **(d)** NPWS and Victoria (Australia). Both axes were

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

presented on a logarithmic scale. The dashed grey line was the 1:1 reference line and the solid orange line indicated the power-law fit. Marginal density plots (blue shades) showed the data distribution along each axis. Text box in the bottom right corner of each panel provided statistical metrics, including sample size (N), log-R^2, median bias, mean bias, median absolute error (MedAE), and root mean square error (RMSE).

3.2.3 Duration accuracy

Overall, a positive relationship was observed between the two datasets, with a log-R^2 of 0.29 (Fig. 7). However, the vast majority of the scatter points fell below the 1:1 reference line, reflecting a systematic underestimation of fire duration by the VIIR-derived dataset. The mean bias was -30.95 days and the median bias was -24 days.

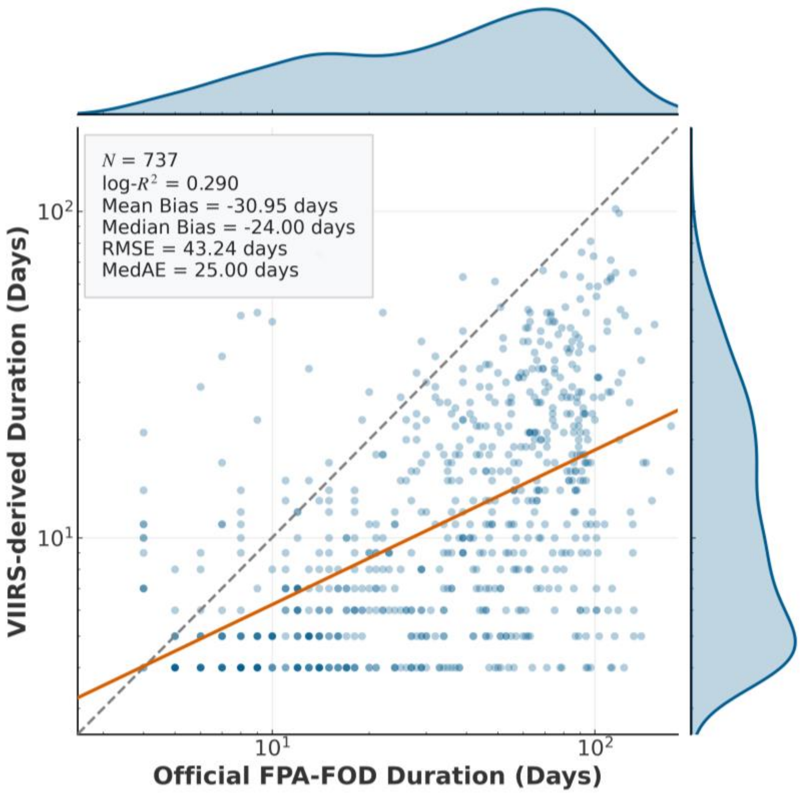


Figure 7. Validation of the VIIRS-derived duration against the point-based FPA-FOD dataset for the United States (2012-2020). The analysis identified a matched subset of 737 wildfire events using the criteria of area > 1.0 km^2 and duration > 3 days. Both the x-axis and y-axis employed logarithmic scales. The scatter plot was complemented by marginal kernel density distributions (blue shaded areas) along the top and right axes to illustrate the overall distribution from both datasets. The dashed black line represented the 1:1 ideal reference, while the solid orange line indicated the power-law (log-log) regression fit. Performance metrics (log-R^2, median bias, MedAE, mean bias, and RMSE) were detailed in the inset text box.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

The underestimation of fire duration came primarily from the end data rather than the start date (Fig. 8). For fire start dates, the differences were tightly centered around zero, with a small positive mean bias of +0.4 days. It indicated an overall good agreement and only a slight delay in VIIRS detection relative to official records. In contrast, fire end dates exhibited a broader distribution and a negative mean bias of -6.1 days, suggesting that VIIRS-derived fire events ended earlier than their counterparts as reported in the official dataset. Overall, these results indicated that while VIIRS reliably captured fire ignition timing, it tended to underestimate fire persistence, especially during later stages.

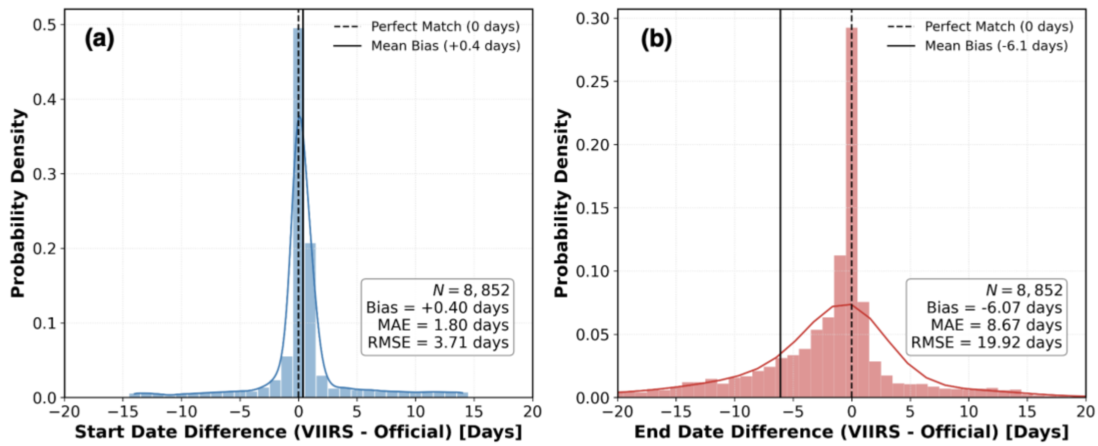


Figure 8. Distribution of differences between VIIRS-derived and officially reported fire dates. (a) Start date differences and **(b)** end date differences, defined as VIIRS estimates minus official records (in days). Histograms showed the probability density of the differences, with overlaid kernel density lines. The dashed vertical line indicated perfect agreement (0 days), and the solid vertical line denoted the mean bias. Summary statistics, including sample size (N), mean bias, mean absolute error (MAE), and root mean square error (RMSE), were provided in each panel.

3.3 Global spatiotemporal patterns of fire regimes

Our VIIRS-derived dataset was an inventory of 15,377,087 wildfire events from 2012 to 2025. Here we showed global characteristics, including frequency, burned area, expansion rate, and fire duration of individual wildfire events within 0.25° × 0.25° grids (Fig. 9). And also analysed in different fire type regions defined in Global Fire Emissions Database (GFED5) (Fig. 9 insets; Fig. B3; Table B1) (Chen et al., 2023).

Fire occurrence (Fig. 9a) was concentrated in tropical savanna regions of Africa and northern Australia, as well as parts of South America, reflecting frequent and widespread burning in these ecoregions. Additional hotspots were evident in boreal regions of North America and Eurasia, although with lower event density compared to tropical regions. The mean daily expansion rate (DER; Fig. 9b) showed a different spatial pattern, with higher values predominantly observed in boreal regions (with grid-mean > 2.7 km^2 day^{-1}), particularly in northern North America and Siberia. It indicated that, although fires

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

occurred less frequently in these regions, they tended to spread more rapidly once ignited. In contrast, tropical regions exhibited relatively lower expansion rates despite high fire occurrence. Mean burned area per event (Fig. 9c) further highlighted regional contrasts. Larger fire sizes were more common in boreal (with grid-mean > 9.4 km^2) and some temperate regions. Similarly, fire duration (Fig. 9d) tended to be longer in boreal ecosystems and shorter in frequently burning regions such as savannas. This extended period of high-latitude burning was probably even longer than the VIIRS-derived dataset suggested, because of the persistent smoldering of the deep organic soils, peatlands, and heavy duff layers abundant in high-latitude biomes with thermal anomalies below the detection threshold of VIIRS(Scholten et al., 2021; Turetsky et al., 2014).

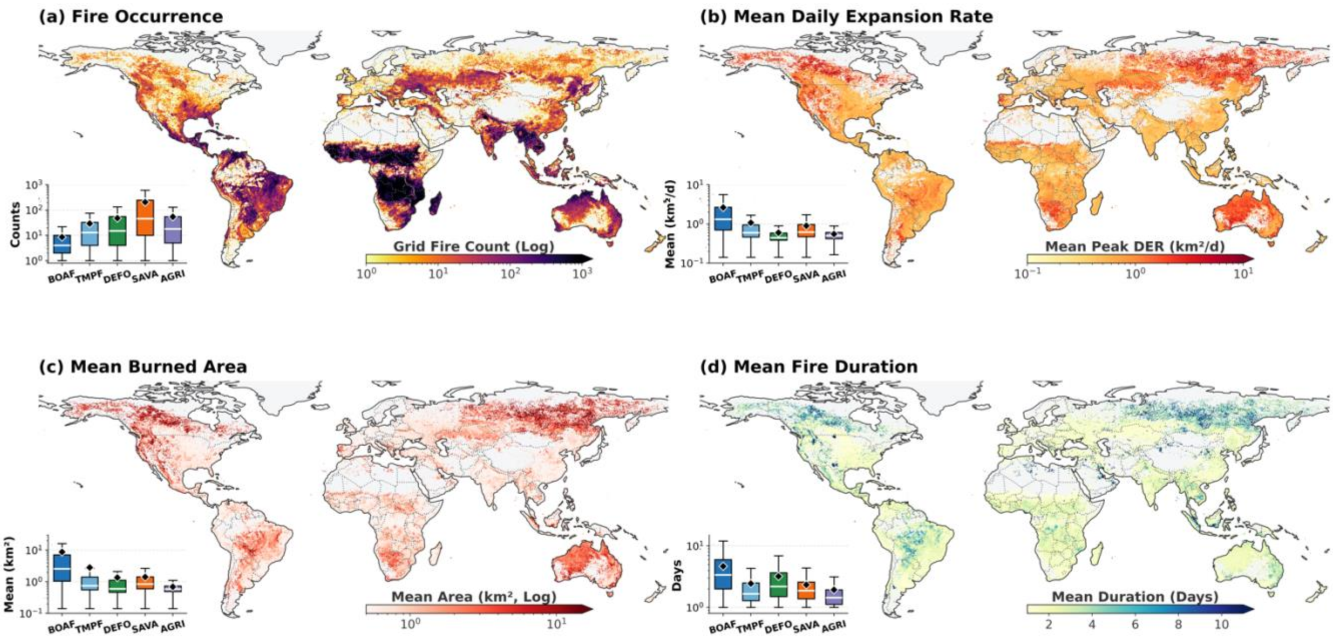


Figure 9. Global patterns of wildfire characteristics derived from VIIRS (2012–2025). (a) Fire occurrence, expressed as the total number of fire events per grid cell (log-scaled). **(b)** Mean daily expansion rate (DER; km^2 d^{-1}). **(c)** Mean burned area per event (km^2; log-scaled). **(d)** Mean fire duration (days). Insets in each panel showed the boxplot of corresponding metrics across major biome types, including boreal forest (BOAF), temperate forest (TMPF), deforestation regions (DEFO), savanna (SAVA), and agricultural lands (AGRI). In the inset boxplots, the black dot indicated the mean, the horizontal white line represented the median, the box spanned the interquartile range (IQR), and the whiskers extended to the minimum and maximum non-outlier values. The All maps were aggregated to a 0.25° × 0.25° grid.

The seasonality of wildfire activity exhibited spatial and biome-dependent patterns across different fire metrics (Fig. 10). In many regions, fire occurrence, expansion rate, burned area and fire duration showed similar seasonal peaks, while differences were also observed in some fire-prone areas. For example, in western coastal North America, fire occurrence mainly peaked in late autumn and early winter, whereas expansion rate, burned area and fire duration tended to peak earlier.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

This indicated that periods with more frequent fire ignitions did not always correspond to periods with the largest or longest fires. In boreal North America and Siberia, the four metrics generally showed similar summer peaks, reflecting the strong seasonal cycle of fire activity. In tropical savanna regions, fire occurrence was distributed across the dry season, while larger burned area and longer fire duration were more concentrated within shorter periods. From the biome level, savanna regions accounted for a large proportion of global fire activity across several seasons (spring/autumn/winter), while boreal fires became more important during summer.

Ultimately, the diverse spatial patterns of global wildfire dynamics highlighted the need for further investigation into their underlying drivers to improve future simulations. By providing a multi-dimensional characterization of fire behavior (size, duration, and expansion rate), this dataset aimed to offer useful observational constraints for global carbon cycle and Earth system modeling.

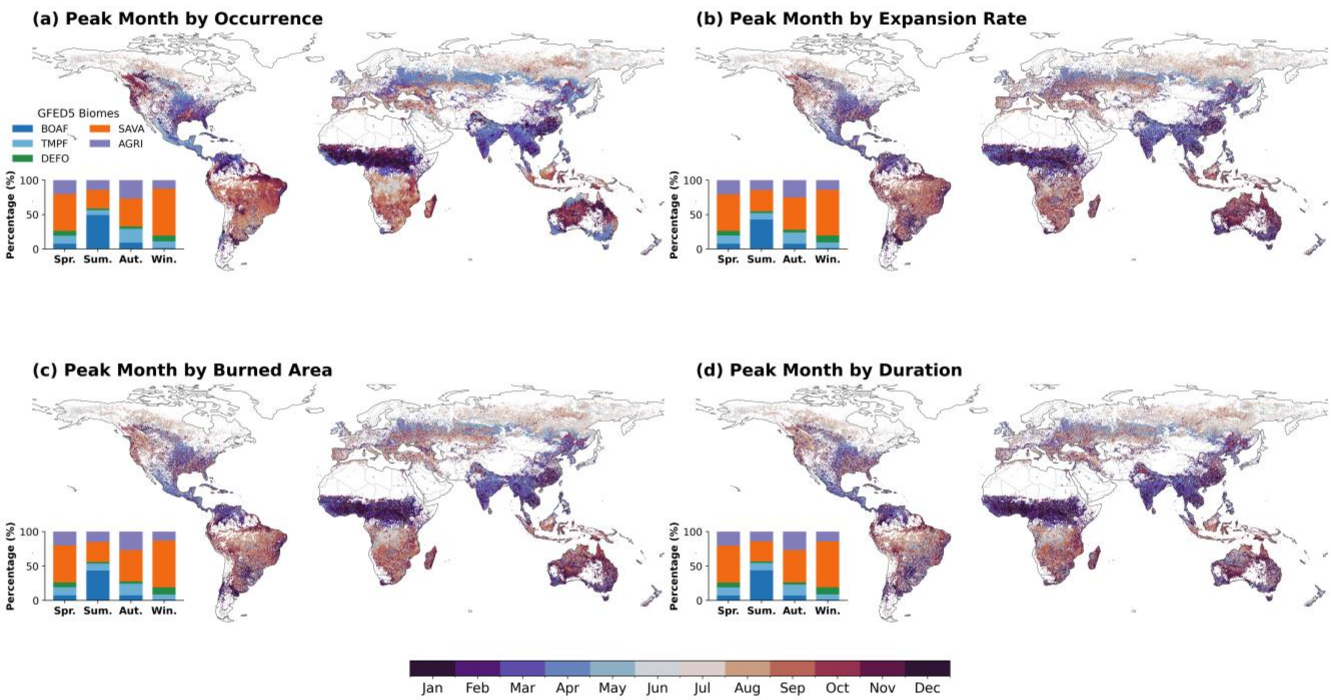


Figure 10. Global distribution of fire seasonality derived from four fire metrics. (a) total fire occurrence, **(b)** mean daily expansion rate (DER), **(c)** mean burned area, and **(d)** mean fire duration. Colours indicated the month of peak activity, defined as the calendar month (January-December) with the maximum value of each metric for a given 0.25° grid cell. Insets in each panel showed the seasonal distribution of peak months aggregated by GFED5 fire types, including boreal forest (BOAF), temperate forest (TEMF), deforestation (DEFO), savanna (SAVA), and agricultural regions (AGRI). Bar charts represented the percentage contribution of each biome to seasonal peaks (spring, summer, autumn, winter). Local seasons accounted for hemispheric inversion (e.g., "Summer" represented JJA in the Northern Hemisphere and DJF in the Southern Hemisphere).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

4 Discussion

4.1 Advantages and disadvantages over parallel datasets

While pioneering global datasets such as the Global Fire Atlas (GFA) (Andela et al., 2019)provided foundational insights into fire dynamics, their reliance on the 500 m MODIS Burned Area product (MCD64A1) imposed some limitations in spatiotemporal details. Burned area approaches were based on post-fire spectral changes, which could introduce delays in burn date estimates and may smooth the progression of fire fronts. Our spatial comparison with official U.S. MTBS perimeters showed that the pixel-aggregation approach adopted by GFA tended to underestimate the burned area of individual events, with a median bias of about -5.7 km^2 (Fig. B6), to a larger extent than the VIIRS-derived mean bias of -4.2 km^2 (Fig. 6b). When validated against FPA-FOD dataset in USA, GFA tended to underestimate the burned area with a median bias of about -5.8 km^2 (Fig. B5). This difference is partly related to the coarser spatial resolution of the MODIS product, which could have difficulty capturing fine-scale burn boundaries and unburned patched within fire perimeters. In comparison, the VIIRS-derived dataset used 375 m active fire detections, which provided more detailed observations of thermal activity and may better capture smaller or more intense fire events.

A known challenge in Global Fire Atlas (GFA) was the representation of long-duration, large fire events. In the case of the 2013 Rim Fire (Fig. 4a), the GFA dataset separated the single continuous wildfire into multiple fire events (Fig. B4). This type of segmentation has been documented and attributed to periods of dense smoke or persistent cloud cover that interrupted optical observations for several consecutive days(Andela et al., 2019; Balch et al., 2020).

In our dataset, fire events were tracked using a sliding-window approach, which helped maintain continuity for some long-duration events. As a result, individual large fires appeared less fragmented, and the representation of fire extent was more consistent with continuous event development in these cases. Yet, the temporal coverage of our dataset (2012–2025) was limited by the availability of VIIRS observations. In contrast, MODIS-based products such as the Global Fire Atlas (GFA) provided a longer time series, which remained valuable for studies focusing on long-term variability. These differences reflected the complementary characteristics of the two datasets in terms of temporal coverage and event representation.

4.2 Interpreting validation differences: definitions of active fire and reposted events

Our validation against multiple regional datasets, including FPA-FOD (USA), MTBS (USA), CNFDB (Canada), and the Australian fire records, showed consistent differences between VIIRS-derived fire metrics and official reports. These differences were mainly related to how fire events were defined and recorded.

For spatial extent, the VIIRS-derived burned area generally agreed with official reports but maintained a conservative area estimation. This official reported burned area represented overall burn extent. As a result, their perimeters may include

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

unburned areas. In contrast, our approach was based on active fire detections and outlined areas with observed thermal activity, thus providing a more reliable estimation of actual burned area.

Differences were also evident for small fires and fire duration. For small fires (e.g., < 1 km^2), lower detection accuracy was found across several datasets. This was partly because VIIRS observations can detect small or short-lived thermal signals that were not always included in regional fire inventories. In addition, VIIRS-derived fire durations were generally shorter than reported durations. Official reports often recorded the full management period of a fire, from initial discovery to final containment; whereas the VIIRS-derived method reflected periods with detectable thermal activity.

Overall, these differences were consistent across regions and datasets, and they mainly reflected differences in observation methods and reporting definitions rather than inconsistencies in the data themselves.

4.3 Inherent limitations and algorithmic uncertainties

Despite the overall validation performance, some uncertainties remained in our dataset. One limitation came from the optical sensitivity of the VIIRS sensor. Persistent cloud cover can interrupt the detection of active fires, even with the temporal buffering used in the algorithm. This may lead to breaks in the tracking of long-duration events. Additionally, understory fires that were fully covered by dense forest canopies may fall below the thermal detection threshold, leading to omissions of surface and underground fires.

Independent event-level observations of daily fire progression were not consistently available across global fire regimes. The available case studies therefore provided evidence that the dataset reproduced aspects of the temporal evolution of fire growth where time-resolved independent reports were available. However, they did not constitute a globally representative independent validation of daily expansion rate (DER). Future evaluation should extend this assessment using event-level fire-progression reference datasets spanning multiple fire regimes as such data becomes available.

5 Conclusion

In this study, we presented a global wildfire event dataset derived from the VIIRS active fire detections (375 m) covering the period from 2012 to 2025. We developed a spatiotemporal tracking algorithm to reconstruct daily evolution of individual fire events, estimated burned area based on multiple methods (pure pixel, convex and concave hull) and extracted other fire metrics.

Validations against authoritative regional databases (FPA-FOD, MTBS, CNFDB, NPWS and Victoria) demonstrated high detection capability (recall exceeding 0.95 for large wildfires). Multi-region statistical analyses supported the reconstruction of event detection, pixel-accumulated fire burned area, and duration. Case-based comparisons further showed that the temporal evolution of fire growth could be reproduced where time-resolved official reports were available, while the global representativeness of DER evaluation remained limited by the availability of independent daily fire-progression reference data.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Case studies and statistical analyses revealed that our VIIRS-derived fire dynamical metrics captured the lifespan and boundaries of thermal anomalies, complementing administrative reporting.

The global analysis of our dataset revealed clear spatial contrasts between fire frequency and severeness across different biomes. This dataset provided an observational foundation for examining the fire life cycle, including ignition, spread and extinction. It can also support an improved representation of fire processes in models. In this way, the dataset offered a useful basis for future studies of wildfire dynamics under a changing climate.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Earth System Science Data Discussions
Open Access

Appendix A: Supporting material for the methods

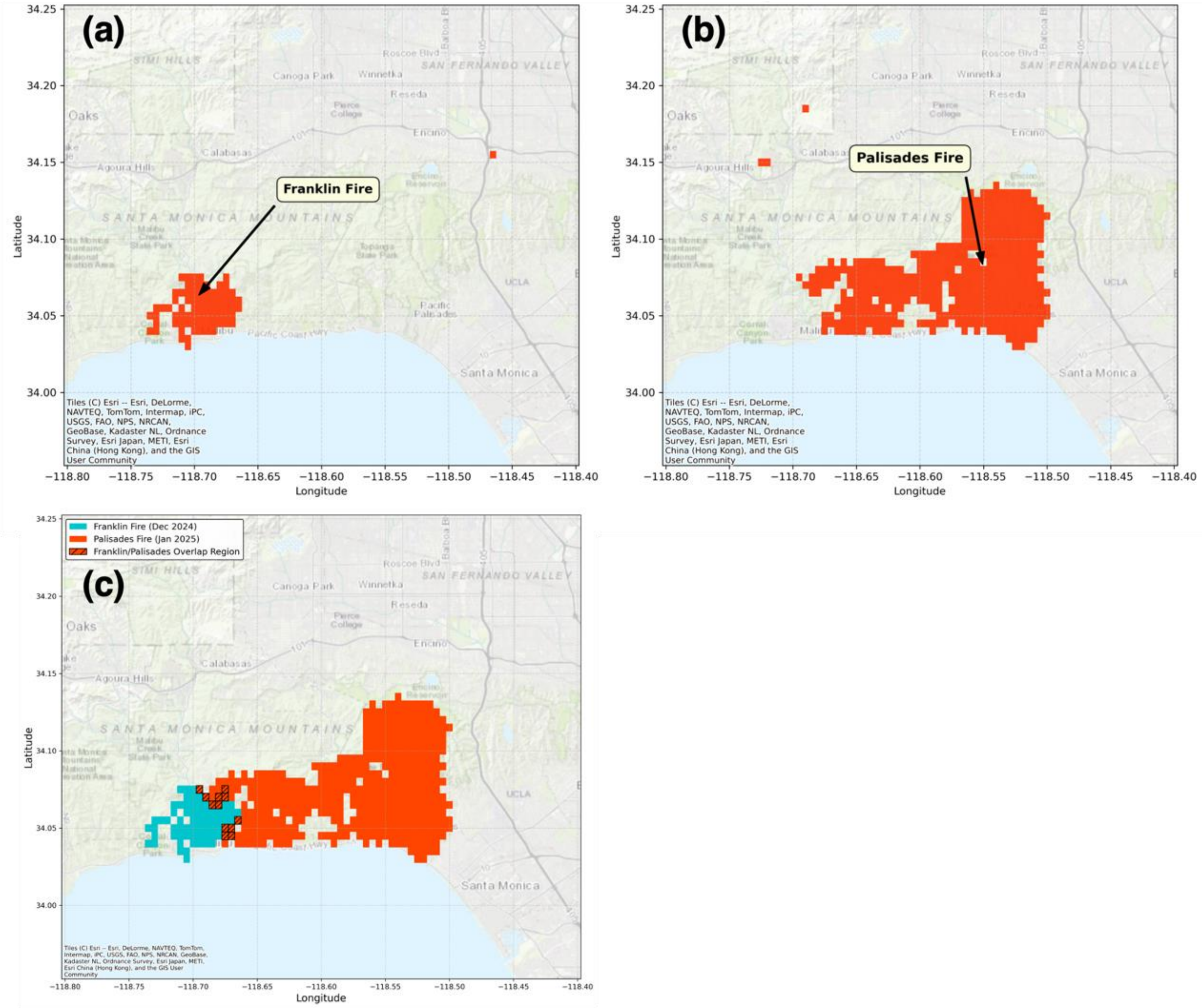
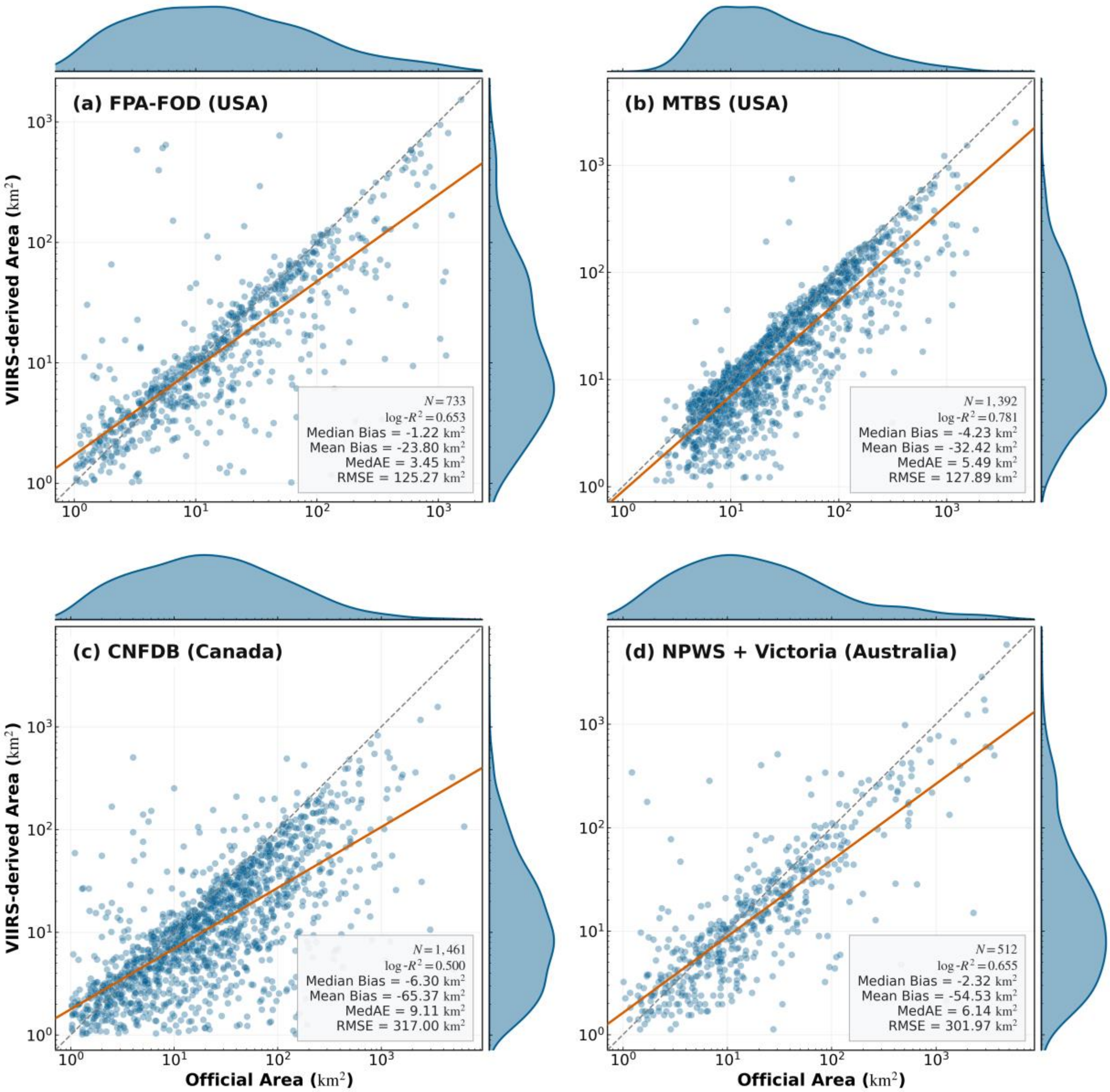


Figure A1. Two distinct, spatially overlapping fire events demonstrating the necessity of temporal event delineation. (a) Franklin Fire 2024, **(b)** Palisades Fire 2025, and **(c)** the overlapped patch. The basemaps were designed and developed by Esri | Powered by Esri. Although these two fires overlapped (black slashed pixels) spatially, a distinct temporal gap existed between the final detection of the first event (green) in December 2024 and the initial detection of the second event (red) in January 2025. This case study demonstrated a critical challenge in satellite fire tracking: wildfires reoccurring in the identical geographic location require rigorous temporal clustering rules to differentiate discrete events.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Appendix B: Supporting material for the results



495 **Figure B1.** Same as Figure 6 but aggregating the VIIRS-derived fire events if one official event matched multiple VIIRS events.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Open Access Earth System Science Data Discussions

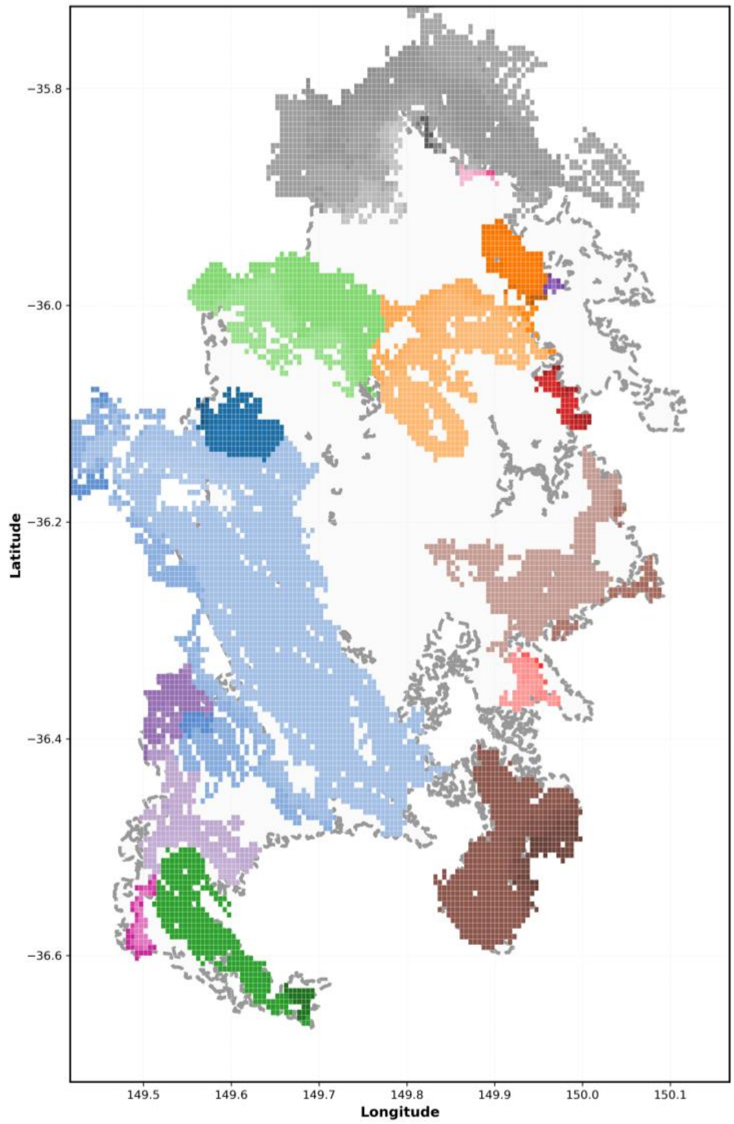
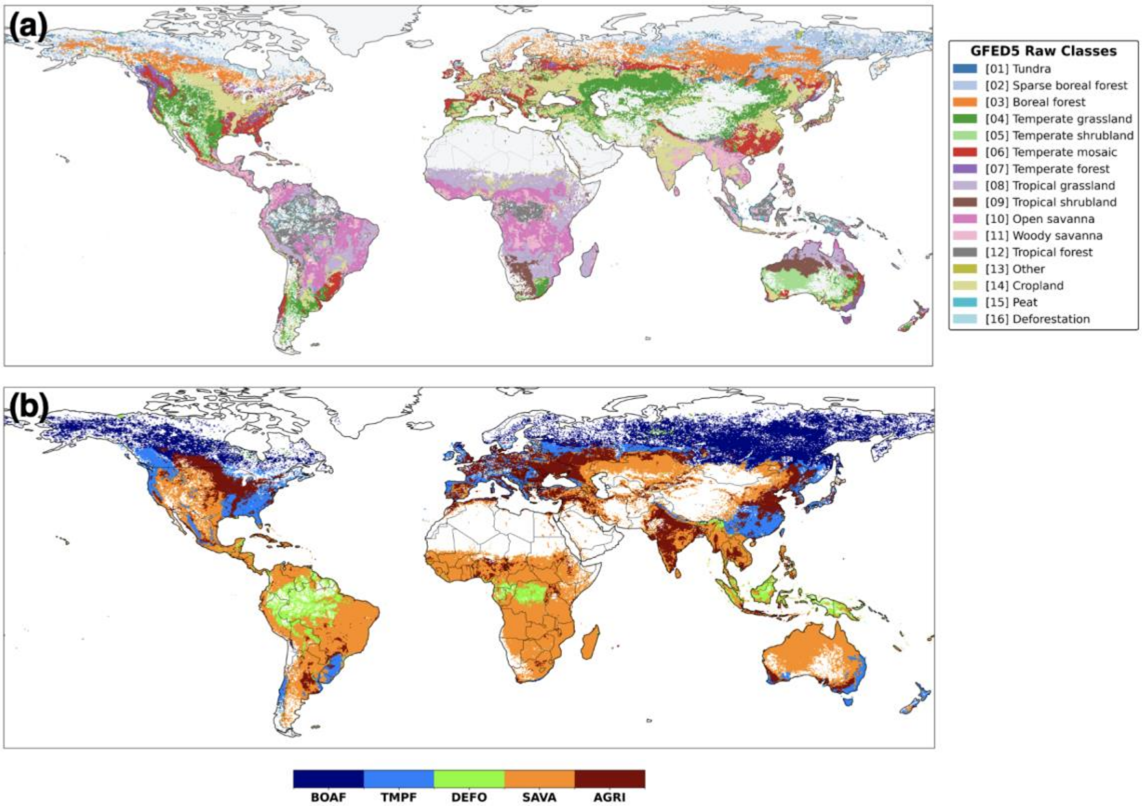


Figure B2. Reconstruction of the NSW_32933 mega-fire complex in the Australia database using a 0.005° gridded rasterization approach.
500 The map illustrated the spatial distribution of 16 individual VIIRS-derived fire events (color-coded by their unique IDs) within the official fire management perimeter (grey outline). Different colours corresponded different ID for each VIIRS-derived event and darker shades indicated later burning times.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Open Access Earth System Science Data Discussions



505 **Figure B3.** Dominant fire types from Global Fire Emissions Database (GFED5). **(a)** Raw 16 fire types defined by GFED5 and **(b)** summarized to five types including boreal forests (BOAF), temperate forests (TMPF), deforestations (DEFO), savannas (SAVA) and agriculture (AGRI).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Table B1. Fire characteristics across 16 fire types defined by Global Fire Emissions Database (GFED5) during 2012-2022. 510 We showed the mean values of burned area (km^2), duration (days), expansion rate (km^2 day^{-1}) for individual fires and the burned-area-weighted mean values in parentheses. For occurrence, we showed the total number and percentage contribution to global total occurrence in parentheses.

GFED fire type	Occurrence	Burned Area (km^2)	Duration (d)	Expansion Rate (km^2 d^{-1})
All	15,337,129	1.1 (62.2)	1.9 (7.6)	0.7 (9.9)
Tundra	13,810 (0.1%)	2.4 (68.6)	2.5 (10.6)	1.1 (13.4)
Boreal forest	95,954 (0.6%)	8.3 (366.9)	3.7 (19.4)	2.5 (62.7)
Sparse boreal forest	88,231 (0.6%)	7.2 (291.3)	3.5 (17.0)	2.3 (50.4)
Temperate forest	67,712 (0.4%)	3.2 (649.4)	2.2 (24.9)	1.2 (86.6)
Temperate grassland	369,232 (2.4%)	1.3 (33.5)	1.8 (6.9)	0.9 (8.7)
Temperate shrubland	61,438 (0.4%)	2.5 (27.9)	2.5 (6.9)	1.5 (12.3)
Temperate mosaic	358,244 (2.3%)	1.6 (176.6)	1.9 (12.5)	0.9 (22.6)
Tropical forest	335,692 (2.2%)	1.3 (224.3)	2.4 (13.9)	0.6 (19.2)
Tropical grassland	3,454,864 (22.5%)	1.0 (10.6)	1.7 (4.4)	0.7 (2.8)
Tropical shrubland	169,906 (1.1%)	2.6 (40.7)	1.9 (5.3)	1.6 (12.1)
Deforestation	25,511 (0.2%)	3.4 (83.5)	4.3 (21.8)	0.9 (8.1)
Woody savanna	2,906,586 (19.0%)	1.0 (12.8)	2.1 (6.7)	0.6 (2.3)
Open savanna	5,952,781 (38.8%)	1.0 (16.5)	1.9 (5.6)	0.6 (2.7)
Peat	50,099 (0.3%)	2.1 (208.2)	3.2 (26.0)	0.7 (12.5)
Cropland	1,374,114 (9.0%)	0.8 (4.3)	1.8 (5.1)	0.5 (1.5)
Other	12,955 (0.1%)	1.3 (73.3)	3.3 (29.0)	0.6 (8.2)

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

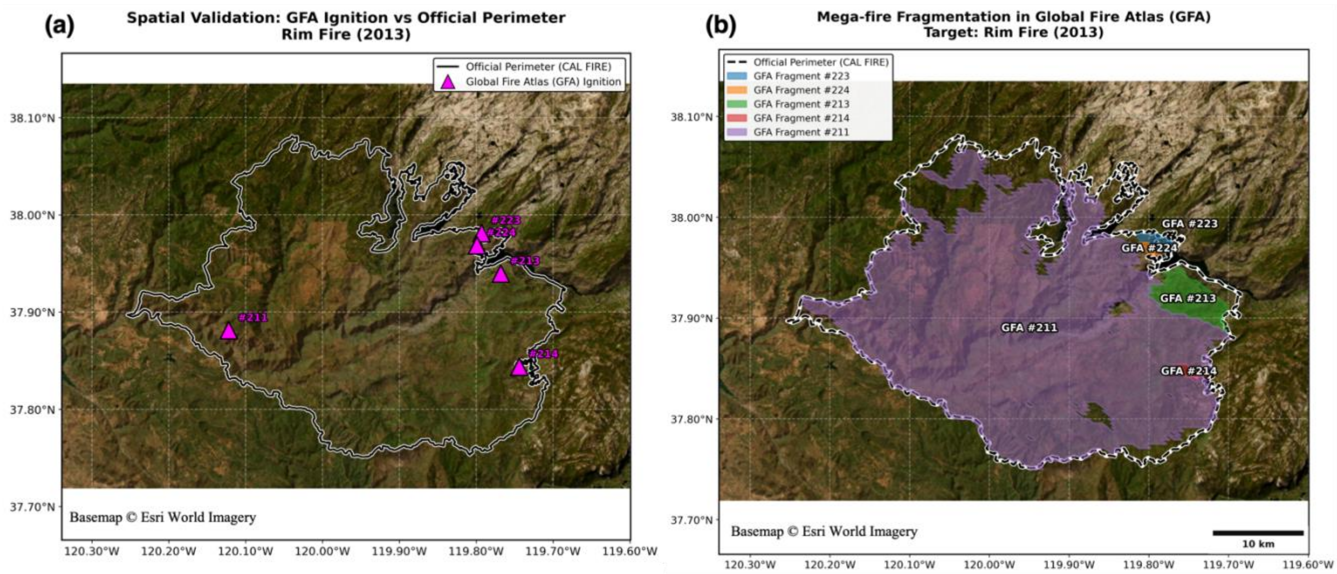
515


Figure B4. Algorithmic fragmentation of the 2013 Rim Fire in Global Fire Atlas (GFA). The basemaps were designed and developed by Esri | Powered by Esri. While the official CAL FIRE perimeter confirmed a single mega-fire (Fig 4a), the GFA algorithm artificially fragmented the event into five distinct fire events (coloured polygons and corresponding ignition points). This fragmentation resulted from temporary satellite obscuration (e.g., thick smoke plumes) that broke the temporal continuity.

520

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

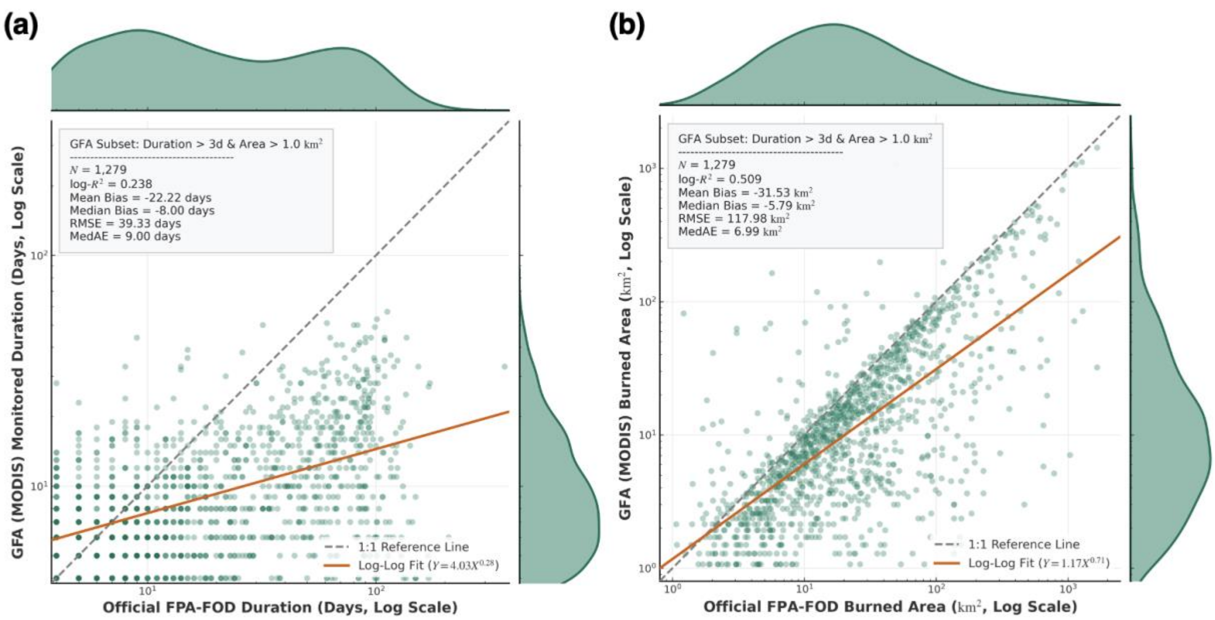
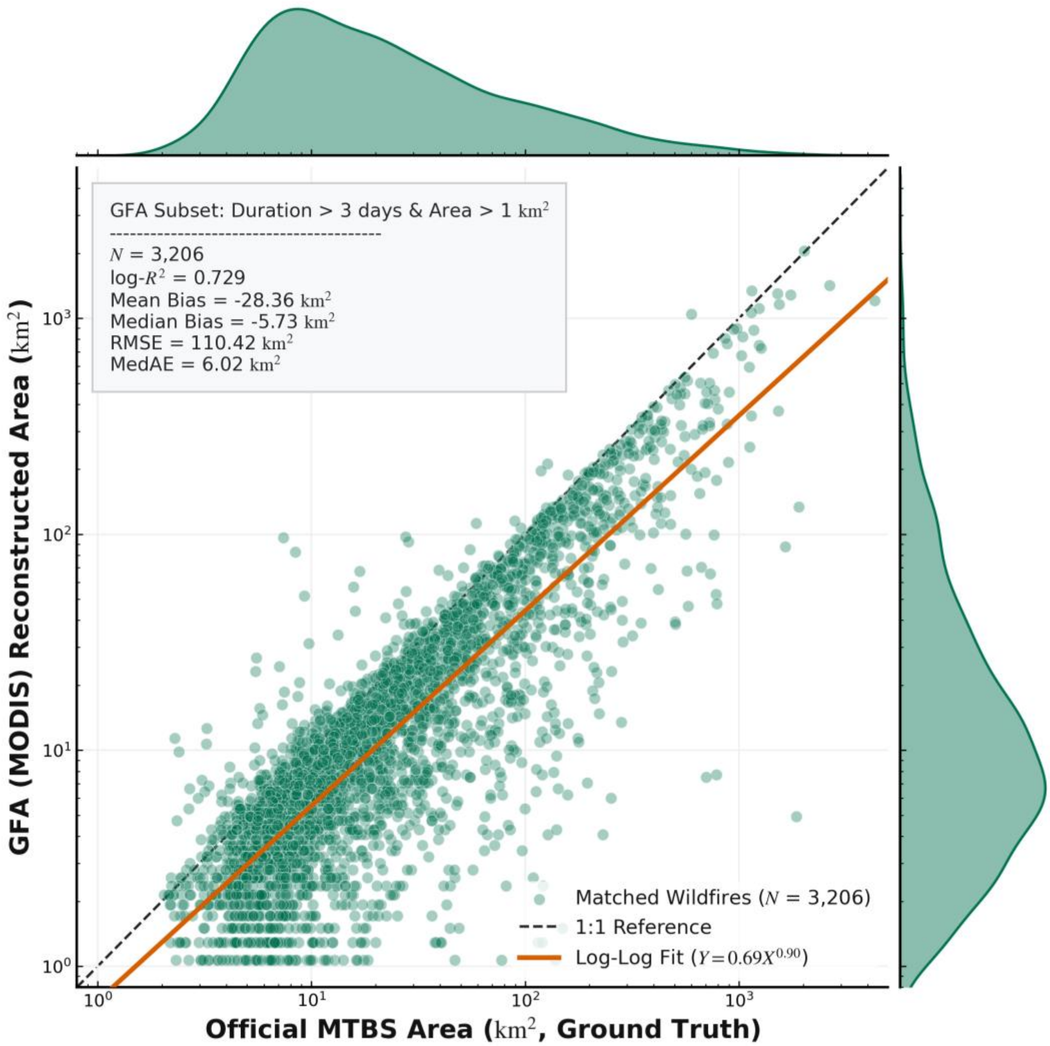


Figure B5. Spatiotemporal benchmarking of the Global Fire Atlas (GFA) dataset against official FPA-FOD incident records. Validation of the MODIS-derived GFA reconstructed duration and burned area against the point-based FPA-FOD dataset for the United States (2012-2020). The analysis identified a matched subset of 1,279 wildfire events using the criteria of area > 1.0 km^2 and duration > 3 days. The comparison assessed **(a)** fire duration and **(b)** total burned area. Both the x-axis and y-axis employed logarithmic scales. The scatter plot was complemented by marginal kernel density distributions (green shaded areas) along the top and right axes to illustrate the overall distribution from both datasets. The dashed black line represented the 1:1 ideal reference, while the solid orange line indicated the power-law (log-log) regression fit. Performance metrics (log-R^2, median bias, MedAE, mean bias, and RMSE) were detailed in the inset text box.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Open Access Earth System Science Data Discussions



530 **Figure B6.** Spatial benchmarking of the Global Fire Atlas (GFA) burned area against official MTBS perimeters. Validation of the MODIS-derived GFA reconstructed burned area against the polygon-based Monitoring Trends in Burned Severity (MTBS) dataset for the United States (2012-2020). The analysis identified a matched subset of 3,206 wildfire events using the criteria of area > 1.0 km^2 and duration > 3 days. Both the x-axis (official MTBS area) and y-axis (GFA reconstructed area) employed logarithmic scales. The scatter plot was complemented by marginal kernel density distributions (green
535 shaded areas) along the top and right axes to illustrate the overall distribution of burned area from both datasets. The dashed black line represented the 1:1 ideal reference, while the solid orange line indicated the power-law (log-log) regression fit. Performance metrics (log-R^2, median bias, MedAE, mean bias, and RMSE) were detailed in the inset text box.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

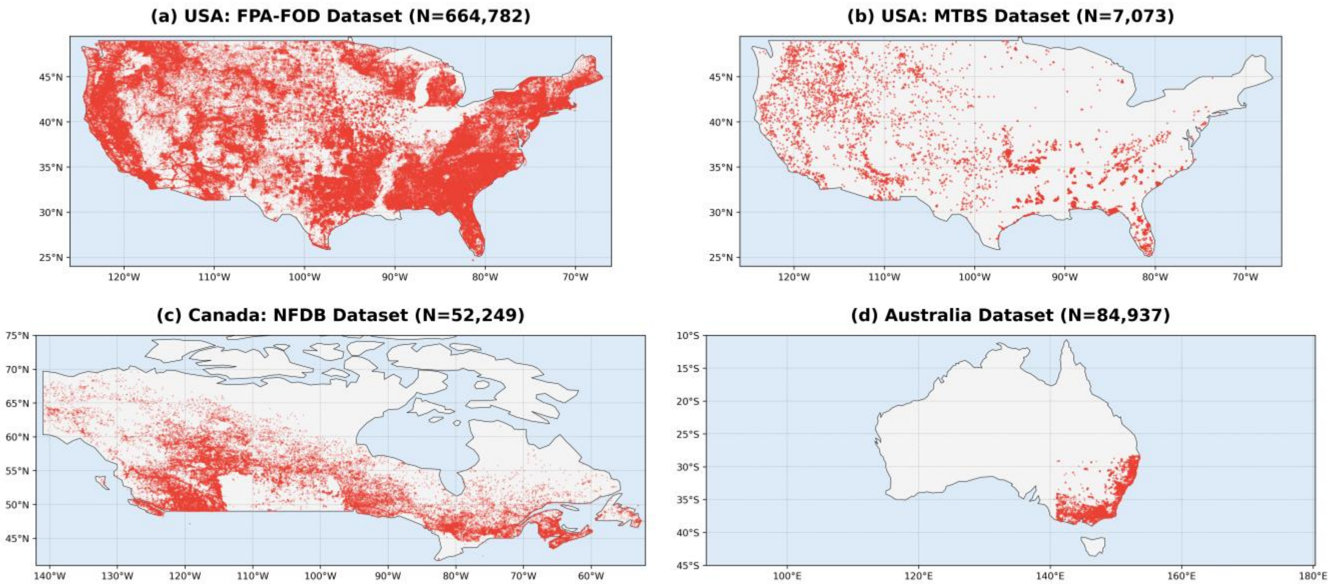


Figure B7. Distribution of the official datasets in USA, Canada and Australia for the period during 2012-2020. Red points denote the locations for fire events. (a) The Fire Program Analysis Fire-Occurrence Database (FPA-FOD) in the contiguous United States (CONUS). (b) The Monitoring Trends in Burn Severity (MTBS) dataset in CONUS, representing large wildfire events. (c) The National Fire Database (NFDB) point dataset for Canada. (d) State-level official fire history records from Australia, including NSW and Victoria. The total number of fire records (N) was indicated above each panel.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Open Access Earth System Science Data Discussions

Appendix C: Dataset documentation and sensitivity analysis

Table C1. Complete data dictionary for the global individual-fire-event dataset.

Variable	Unit	Definition and calculation	Recommended interpretation
fire_id	–	Identifier constructed from the source year, source month, and monthly event identifier.	Unique identifier for each reconstructed individual fire event.
start_date	YYYY-MM-DD	First date of the reconstructed VIIRS fire event.	VIIRS-detected fire start date.
end_date	YYYY-MM-DD	Last date of the reconstructed VIIRS fire event.	VIIRS-detected fire end date.
ign_lat	°N	Latitude of the centroid of unique VIIRS fire pixels detected on the first day of the event.	Represents the initial VIIRS-detected ignition location.
ign_lon	°E	Longitude of the centroid of unique VIIRS fire pixels detected on the first day of the event.	Represents the initial VIIRS-detected ignition location; negative values indicate western longitudes.
duration_days	days	Number of calendar days from start_date to end_date.	VIIRS-derived fire duration.
active_days	days	Number of days with at least one retained VIIRS active-fire detection.	Describes the number of VIIRS-observed active-fire days within the event duration.
n_pixels	pixels	Number of unique VIIRS fire pixels assigned to the event; repeated detections of the same pixel are counted once.	Number of pixels used to calculate the VIIRS-derived burned area.
max_gap_days	days	Maximum number of intervening days without detections between successive active-fire detection days.	Larger values indicate lower temporal continuity in the VIIRS observations.
area_km2	km^2	VIIRS-derived burned area calculated as n_pixels × 0.140625 km^2 for nominal 375 m × 375 m pixels.	Conservative estimate of the area with observed thermal activity rather than the complete fire-perimeter area.
mean_der	$km^2\ d^{-1}$	Mean daily expansion rate calculated as area_km2 / duration_days.	Average expansion rate over the reconstructed fire duration.
peak_der	$km^2\ d^{-1}$	Maximum daily area of pixels newly detected during the reconstructed event.	Maximum VIIRS-derived daily expansion rate.
explosive_idx	ratio (0–1)	Ratio of peak_der to area_km2.	Proportion of total VIIRS-derived burned area detected on the day of maximum expansion; larger values indicate more concentrated growth.
frp_mw	MW	Sum of fire radiative power (FRP) from all retained VIIRS detections assigned to the event.	Total detected radiative power of the event.
gfed5_type	categorical code	Most frequent valid GFED5 16-class fire type among the unique event pixels.	Dominant GFED5 fire type associated with the reconstructed event.
cross_month	0/1	Indicates whether start_date and end_date fall in different calendar months.	0 = same calendar month; 1 = spans more than one calendar month.
temporal_flag	0/1	High-gap warning derived from max_gap_days.	0 = max_gap_days < 4; 1 = max_gap_days ≥ 4.
quality_flag	0–7	1 = n_pixels < 5; 2 = max_gap_days ≥ 4; 4 = cross_month = 1.	Codes may be combined by addition. A value of 0 means no predefined warning was triggered; it does not imply zero measurement error.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Variable	Unit	Definition and calculation	Recommended interpretation
area_ref_km2	km²	Burned area predicted from the relationship between VIIRS-derived area and MTBS area in the matched US validation sample.	Regression-adjusted estimate provided for uncertainty assessment; it should not be interpreted as a direct observation or a replacement for the primary VIIRS-derived burned area (area_km2).
area_p90_lo	km²	Lower bound of the 90 % prediction interval for area_ref_km2.	Prediction uncertainty estimated from the matched MTBS validation samples.
area_p90_hi	km²	Upper bound of the 90 % prediction interval for area_ref_km2.	Prediction uncertainty estimated from the matched MTBS validation samples.
area_extrap	0/1	Indicates whether area_km2 lies outside the VIIRS-area range represented by the MTBS calibration sample.	0 = within the calibration range; 1 = extrapolated or invalid input.
duration_ref_days	days	Fire duration predicted from the relationship between VIIRS-derived duration and FPA-FOD duration in the matched US validation sample.	Regression-adjusted estimate provided for uncertainty assessment; it should not be interpreted as a direct observation or a replacement for the primary VIIRS-derived fire duration (duration_days).
duration_p90_lo	days	Lower bound of the 90 % prediction interval for duration_ref_days.	Prediction uncertainty estimated from the matched FPA-FOD validation samples.
duration_p90_hi	days	Upper bound of the 90 % prediction interval for duration_ref_days.	Prediction uncertainty estimated from the matched FPA-FOD validation samples.
duration_extrap	0/1	Indicates whether duration_days lies outside the VIIRS-duration range represented by the FPA-FOD calibration sample.	0 = within the calibration range; 1 = extrapolated or invalid input.
reg_flag	0–3	Combined regression-applicability flag: area_extrap + 2 × duration_extrap.	0 = both within range; 1 = area only; 2 = duration only; 3 = both extrapolated. This is not an event-quality score.

Note: Prediction intervals are regression-derived and conditional on matched US validation samples. They quantify reference agreement and should not be interpreted as globally validated observation-level uncertainty or as direct measurements of true error for individual events.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

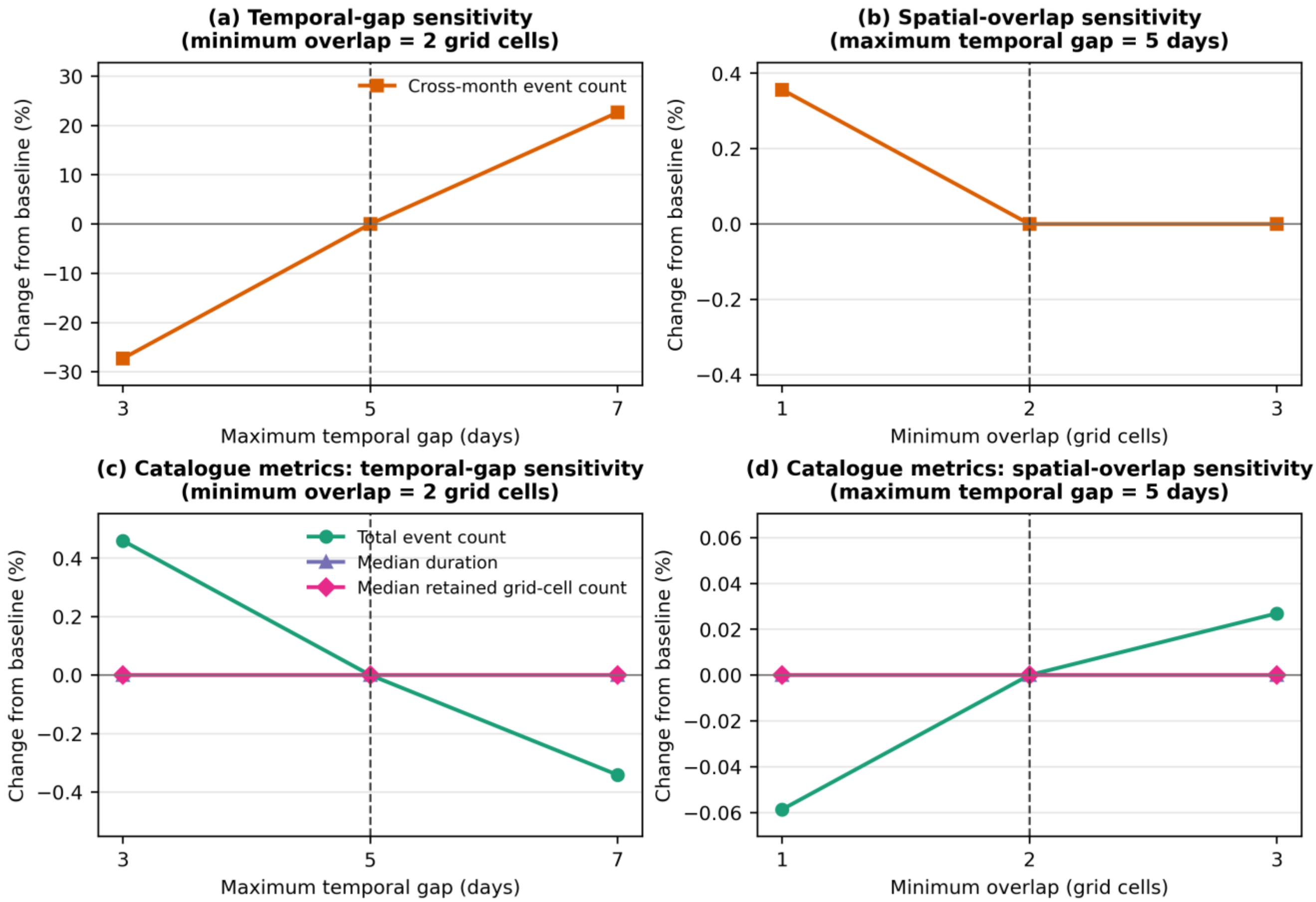


Figure C1. One-factor-at-a-time sensitivity analysis of cross-month fire-event aggregation thresholds over CONUS in 2020.
555 Panels (a–b) showed changes in the number of cross-month events, and panels (c–d) showed corresponding changes in catalogue metrics. Values were expressed as percentage changes relative to the baseline configuration of a five-day temporal gap and a two-grid-cell overlap, indicated by dashed vertical lines.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Table C2. Case-based quantitative comparison of VIIRS-observed daily expansion rate (DER) and official report-derived DER.

Case	N aligned intervals	Spearman rho	MAE ($km^2\ d^{-1}$)	RMSE ($km^2\ d^{-1}$)	Mean bias ($km^2\ d^{-1}$)	Peak timing offset (d)
2013 Rim Fire	28	0.714	23.159	45.539	−7.535	0
2025 Palisades Fire	22	1.000	1.220	5.390	−1.183	0

Data availability

The dataset is freely available as both CSV and Apache Parquet files at https://zenodo.org/records/21807914 (Su, 2026) and will be updated as the corresponding VIIRS active fire records are released. The CSV files support broad accessibility, whereas the Parquet files provide efficient columnar access for large-scale analysis. The event-level files include Global_Fire_ID, VIIRS-detected start location and date, fire duration, pixel-accumulated area (Total_Area_km2), mean and peak daily expansion rate, explosive index, total FRP, GFED5 fire-type information, and per-event quality-information fields. A complete data dictionary is provided in the repository README and Table C1.

Author contributions

YY designed the study. HS carried out the processing and analysis. HS and YY wrote the manuscript.

Competing interests

The authors declare no conflict of interest.

Acknowledgements

We thank Rebecca Scholten for helpful discussions. Computation is supported by High-performance Computing Platform of Peking University.

Financial support

This work was funded by National Key R&D Program of China (2022YFF0801303) (Y.Y.) and the Peking University – BHP Carbon and Climate Wei-Ming PhD Scholars Program (Program Number: WM202502) (H.S.).

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

References

Andela, N., Morton, D. C., Schroeder, W., Chen, Y., Brando, P. M., and Randerson, J. T.: Tracking and classifying Amazon fire events in near real time, Science advances, 2022.

Andela, N., Morton, D. C., Giglio, L., Paugam, R., Chen, Y., Hantson, S., van der Werf, G. R., and Randerson, J. T.: The Global Fire Atlas of individual fire size, duration, speed and direction, Earth System Science Data, 11, 529-552, 10.5194/essd-11-529-2019, 2019.

Archibald, S., Lehmann, C. E., Gomez-Dans, J. L., and Bradstock, R. A.: Defining pyromes and global syndromes of fire regimes, Proc Natl Acad Sci U S A, 110, 6442-6447, 10.1073/pnas.1211466110, 2013.

Balch, J. K., St. Denis, L. A., Mahood, A. L., Mietkiewicz, N. P., Williams, T. M., McGlinchy, J., and Cook, M. C.: FIRED (Fire Events Delineation): An Open, Flexible Algorithm and Database of US Fire Events Derived from the MODIS Burned Area Product (2001–2019), Remote Sensing, 12, 10.3390/rs12213498, 2020.

Palisades Fire Incident Updates: https://www.fire.ca.gov/incidents/2025/1/7/palisades-fire/updates, last access: March 18, 2026.

California Department of Forestry and Fire Protection: Fire and Resource Assessment Program (FRAP) Fire Perimeters Database [dataset], 2026.

Canadian Forest Service: Canadian National Fire Database (CNFDB) – Agency Fire Data [dataset], 2026.

Chen, Y., Hall, J., van Wees, D., Andela, N., Hantson, S., Giglio, L., van der Werf, G. R., Morton, D. C., and Randerson, J. T.: Multi-decadal trends and variability in burned area from the fifth version of the Global Fire Emissions Database (GFED5), Earth System Science Data, 15, 5227-5259, 10.5194/essd-15-5227-2023, 2023.

Chen, Y., Hantson, S., Andela, N., Coffield, S. R., Graff, C. A., Morton, D. C., Ott, L. E., Foufoula-Georgiou, E., Smyth, P., Goulden, M. L., and Randerson, J. T.: California wildfire spread derived using VIIRS satellite observations and an object-based tracking system, Sci Data, 9, 249, 10.1038/s41597-022-01343-0, 2022.

Chuvieco, E., Giglio, L., and Justice, C.: Global characterization of fire activity: toward defining fire regimes from Earth observation data, Global Change Biology, 14, 1488-1502, 10.1111/j.1365-2486.2008.01585.x, 2008.

Department of Energy Environment and Climate Action: Fire History Records of Fires across Victoria [dataset], 2026.

Descombes, P., Gaboriau, T., Albouy, C., Heine, C., Leprieur, F., and Pellissier, L.: Linking species diversification to palaeo-environmental changes: A process-based modelling approach, Global Ecology and Biogeography, 27, 233-244, https://doi.org/10.1111/geb.12683, 2018.

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., and Howard, S.: A project for monitoring trends in burn severity, Fire Ecology, 3, 3-21, 2007.

Frantz, D., Stellmes, M., Röder, A., and Hill, J.: Fire spread from MODIS burned area data: obtaining fire dynamics information for every single fire, International Journal of Wildland Fire, 25, 1228-1237, 10.1071/wf16003, 2016.

Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), Journal of Geophysical Research: Biogeosciences, 118, 317-328, 10.1002/jgrg.20042, 2013.

Gillies, S., van der Wel, C., Van den Bossche, J., Taves, M. W., Arnott, J., Ward, B. C., and others: Shapely, GitHub repository [code], 2025.

Hantson, S., Andela, N., Goulden, M. L., and Randerson, J. T.: Human-ignited fires result in more extreme fire behavior and ecosystem impacts, Nat Commun, 13, 2717, 10.1038/s41467-022-30030-2, 2022.

Rim Fire Summary Timeline: https://mymotherlode.com/community/fire/rim-fire-summary-timeline, last access: April 21, 2026.

Pateiro-L´opez, B. and Rodr´ıguez-Casal, A.: Generalizing the convex hull of a sample: the R package alphahull, Journal of Statistical Software, 34, 2010.

Rabin, S. S., Melton, J. R., Lasslop, G., Bachelet, D., Forrest, M., Hantson, S., Kaplan, J. O., Li, F., Mangeon, S., Ward, D. S., Yue, C., Arora, V. K., Hickler, T., Kloster, S., Knorr, W., Nieradzik, L., Spessa, A., Folberth, G. A., Sheehan, T., Voulgarakis, A., Kelley, D. I., Prentice, I. C., Sitch, S., Harrison, S., and Arneth, A.: The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions, Geoscientific Model Development, 10, 1175-1197, 10.5194/gmd-10-1175-2017, 2017.

https://doi.org/10.5194/essd-2026-387
Preprint. Discussion started: 12 August 2026

Ramo, R., Roteta, E., Bistinas, I., van Wees, D., Bastarrika, A., Chuvieco, E., and van der Werf, G. R.: African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data, Proc Natl Acad Sci U S A, 118, 10.1073/pnas.2011160118, 2021.

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

Scholten, R. C., Banerjee, T., Chen, Y., Delgado, A., Desai, A., Ke, Z., Liu, T., Morton, D. C., Peterson, D. A., Tang, Q., Veraverbeke, S. S., Zhang, J., and Randerson, J. T.: Multi-ignition fire complexes drive extreme fire years and impacts, Science advances, 12, 2026.

Schroeder, W., Giglio, L., and Hall, J.: Collection 2 Visible Infrared Imaging Radiometer Suite (VIIRS) 375-m Active Fire Product User's Guide, 2025.

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment, Remote Sensing of Environment, 143, 85-96, 10.1016/j.rse.2013.12.008, 2014.

Short, K. C.: A spatial database of wildfires in the United States, 1992-2011, Earth System Science Data, 6, 1-27, 10.5194/essd-6-1-2014, 2014.

Short, K. C.: Spatial wildfire occurence data for the United States, 1992-2020, 2022.

State Government of NSW and Department of Planning Industry and Environment: NPWS Fire History - Wildfires and Prescribed Burns [dataset], 2026.

Su, H.: A Global Dataset of Individual Fire Events (2012-2025) derived from VIIRS Active Fire Product, Zenodo [dataset], 10.5281/zenodo.21807914, 2026.

Su, H., Yu, Y., Guo, W., and Mao, J.: Convective potential and fuel availability complementnear-surface weather in regulating globalwildfire activity, Science Advances, 2025.

Turetsky, M. R., Benscoter, B., Page, S., Rein, G., van der Werf, G. R., and Watts, A.: Global vulnerability of peatlands to fire and carbon loss, Nature Geoscience, 8, 11-14, 10.1038/ngeo2325, 2014.