RESEARCH PAPER
https://doi.org/10.1071/WF22048

International Journal of WILDLAND FIRE

International Association of Wildland Fire

Consistent, high-accuracy mapping of daily and sub-daily wildfire growth with satellite observations

Crystal D. McClure[A], Nathan R. Pavlovic[A,*], ShihMing Huang[A], Melissa Chaveste[A] and Ningxin Wang[A]

For full list of author affiliations and declarations see end of paper

*Correspondence to:
Nathan R. Pavlovic
Sonoma Technology, Inc., 1450 N. McDowell Boulevard, Suite 200, Petaluma, CA 94954, USA
Email: npavlovic@sonomatech.com

Received: 9 April 2022
Accepted: 23 February 2023
Published: 3 April 2023

Cite this:
McClure CD *et al.* (2023)
International Journal of Wildland Fire
doi:10.1071/WF22048

ABSTRACT

Background. Fire research and management applications, such as fire behaviour analysis and emissions modelling, require consistent, highly resolved spatiotemporal information on wildfire growth progression. **Aims.** We developed a new fire mapping method that uses quality-assured sub-daily active fire/thermal anomaly satellite retrievals (2003–2020 MODIS and 2012–2020 VIIRS data) to develop a high-resolution wildfire growth dataset, including growth areas, perimeters, and cross-referenced fire information from agency reports. **Methods.** Satellite fire detections were buffered using a historical pixel-to-fire size relationship, then grouped spatiotemporally into individual fire events. Sub-daily and daily growth areas and perimeters were calculated for each fire event. After assembly, fire event characteristics including location, size, and date, were merged with agency records to create a cross-referenced dataset. **Key results.** Our satellite-based total fire size shows excellent agreement with agency records for MODIS ($R^2 = 0.95$) and VIIRS ($R^2 = 0.97$) in California. VIIRS-based estimates show improvement over MODIS for fires with areas less than 4047 ha (10 000 acres). To our knowledge, this is the finest resolution quality-assured fire growth dataset available. **Conclusions and Implications.** The novel spatiotemporal resolution and methodological consistency of our dataset can enable advances in fire behaviour and fire weather research and model development efforts, smoke modelling, and near real-time fire monitoring.

Keywords: fire behaviour, fire detection, fire growth, fire history, MODIS, remote sensing, VIIRS, wildfire perimeters.

Introduction

Increasing wildfire activity in the western United States has led to unprecedented impacts on communities including evacuations, destruction of homes and businesses, loss of services such as electric power (Mitchell 2013; McLennan *et al.* 2019; Jazebi *et al.* 2020; Wong *et al.* 2020; Grajdura *et al.* 2021; Jahn *et al.* 2022), smoke that drives declining air quality outdoors and indoors (Brey *et al.* 2018; Larsen *et al.* 2018; McClure and Jaffe 2018; Messier *et al.* 2019; Shrestha *et al.* 2019; Jaffe *et al.* 2020; Li *et al.* 2020), and health impacts (Averett 2016; Reid *et al.* 2016; Reid and Maestas 2019; Reid *et al.* 2019). To address the societal, economic, and health impacts of wildfires, modelling is used to assess fire risk, predict fire growth and behaviour, forecast air quality impacts from smoke, and develop methods for smoke exposure assessments (Li *et al.* 2020). Increasingly, statistical models based on machine learning approaches are used to understand the impact of wildfires (Reid *et al.* 2015; Yao *et al.* 2018; Zou *et al.* 2019; Jain *et al.* 2020; Hung *et al.* 2021). Creation of a consistent, historical, high-resolution record of wildfire activity is crucial to developing these models that support fire impact mitigation. An ideal dataset incorporates: (1) multiple sources of wildfire information; (2) a combination of agency and satellite products; (3) a sub-daily temporal resolution to aid with wildfire growth and emissions modelling; (4) estimates of fire characteristics such as Fire Radiative Power/Fire Radiative Energy (FRP/FRE) at sub-daily and daily levels; and (5) a multi-decadal record of fire activity.

Many wildfire mapping approaches and databases exist. Agency fire records, such as the California Department of Forestry and Fire Protection (CAL FIRE) Fire and Resource

Assessment Program (FRAP) fire history dataset provide historical data for some locations. Recently, United States Forest Service (USFS) has prepared a database that combines many of these agency sources (Fire Program Analysis Fire-Occurrence Database, FPA FOD) (Short 2021). While they typically provide high quality information on total fire events, these records generally lack temporal resolution and, frequently, spatial detail. Further, these datasets do not provide consistent data characteristics over space and time due to changes in fire record methods, differences in reporting across jurisdictions, and occasionally, errors in reporting of fire location. As an alternative to agency records, fire activity datasets are available from satellite such as the Loboda and Csiszar (2007) northern Eurasia wildfire growth dataset; Coen and Schroeder (2013) high-resolution, single fire dataset; Veraverbeke *et al.* (2014), and Sá *et al.* (2017) large wildfire growth datasets; Benali *et al.* (2016), Laurent *et al.* (2018), Andela *et al.* (2019), Artés *et al.* (2019), Sayad *et al.* (2019), and Lizundia-Loiola *et al.* (2020) global wildfire datasets; Balch *et al.* (2020) continuous United States fire event database; Scaduto *et al.* (2020) northern California large wildfire dataset; and the California-specific wildfire dataset recently published by Chen *et al.* (2022). However, these datasets lack an agency or fire identification component, have short temporal coverage, and/or are based on only Moderate Resolution Imaging Spectroradiometer (MODIS) or Visible Infrared Imaging Radiometer Suite (VIIRS) retrievals. Finally, the Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation (SmartFire) ver. 2 fire information system (Raffuse *et al.* 2012) used to create the fire database for the wildland fire portion of the United States Environmental Protection Agency (EPA) National Emissions Inventory (Larkin *et al.* 2020) was developed to merge any number of fire information datasets, including agency and satellite-based datasets into a single, unified dataset. While SmartFire has been employed to create daily fire growth information, it lacks the ability to take full advantage of the sub-daily sampling frequency of satellite fire detection instruments.

We developed a satellite-based methodology that meets the need for spatially and temporally consistent, high-resolution fire growth data for statistical modelling applications. Our approach uses satellite-detected thermal anomaly data to generate spatially specific fire perimeter and growth records at a sub-daily temporal resolution. The satellite fire detection data are reviewed for quality, filtered to remove persistence and agricultural fires, and merged spatially to estimate fire growth at the time of each satellite observation. Using this time-resolved growth record, we construct: (1) an overall fire event; (2) daily fire growth; and (3) sub-daily fire activity and growth databases that can be linked with agency records, where available, to incorporate fire names, agency estimated fire size, and other values such as fire type (e.g. wildfire, prescribed fire), fire cause descriptions, and damage estimates. The method is applied to multiple satellite fire detection sources and can be used consistently across large spatial domains. Our approach is unique compared with other satellite-derived fire growth products currently detailed in the literature. Some aspects that, together, make our approach unique include automatic merging multiple sources of agency data, use of MODIS fire data to provide a longer historical time period, and screening out of persistence and agricultural fires.

Using this method, we developed fire growth data for California from 2003 using the MODIS sensor data and 2012 using the VIIRS sensor data through 2020 and investigated spatial and temporal trends in fire activities. To validate the resulting datasets, we used independent agency data to evaluate the accuracy of the MODIS and VIIRS datasets at overall event, daily, and sub-daily fire activity scales. Although we focus on this dataset through 2020 in this paper, the dataset is expected to be updated yearly to include new satellite and agency information when it becomes available.

Methods

Satellite data sources

We obtained satellite data from the MODIS sensor on-board the National Aeronautics and Space Administration (NASA) Terra and Aqua Satellites and the VIIRS sensor on-board the Suomi National Polar-orbiting Partnership (Suomi NPP) satellite. MODIS Level 3 ver. 6 Active Fire Product data (Giglio *et al.* 2016) were downloaded from the NASA Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC) data portal (NASA 2021). We downloaded all 1-km × 1-km resolution Aqua and Terra fire detection data (MOD14A1 and MYD14A1) for the western US from 2003 (the first year of complete availability) to 2020, providing four observations per day at approximately 10:30 hours, 22:30 hours, 01:30 hours and 13:30 hours local time.

NASA VIIRS Land Science Investigator Processing System (SIPS) Active Fire (VNP14IMG) Collection 1 Standard Processing Products for 2012 (beginning of availability) to 2019 were downloaded from the Fire Information for Resource Management System (FIRMS) archive (Schroeder *et al.* 2014; NASA 2020). Data from 2020 were downloaded from the University of Maryland VIIRS Active Fire data portal (https://viirsfire.geog.umd.edu/maps; unfortunately, at the time of this publication, this website and dataset are no longer available – raw VIIRS data can instead be found at the LAADS DAAC data portal referenced previously). The VIIRS I-band Active Fire Product has a nominal 375-m spatial resolution at nadir, with pixel size increasing from nadir towards the limb due to viewing angle. Observations are available at approximately 01:30 hours and 13:30 hours local time, and additional observations can occur due to overlapping VIIRS swaths over California.

To prepare the satellite fire data, satellite data were visualised in geographic information system software to

identify product anomalies. A substantial number of false detections covering most of northern California was identified in a Terra image on 11 December 2014, and the affected image was excluded from the fire database during processing. Detections were filtered for quality, persistence, and agricultural fires during the fire growth mapping, as described in the Fire Growth Mapping section.

Agency data sources

The focus of this research is wildfire occurrence and growth mapping. Therefore, we sought out all centralised agency databases that contained wildfire records for California. We did not acquire specific prescribed fire records. Agency records from CAL FIRE FRAP, Incident Status Summary (ICS-209), Geospatial Multi-Agency Coordination (GeoMAC), National Interagency Fire Center (NIFC), USFS Fire Statistics System (FIRESTAT), and FPA FOD datasets were used to enhance the satellite-based fire database by providing fire names, agency estimate fire size, and other values such as fire type, fire cause descriptions, and damage estimates. Agency fire perimeter data transitioned from GeoMAC to NIFC starting in 2020. The agency datasets used to augment MODIS and VIIRS fire event database are in Table 1. Agency datasets and QC procedures are described in the Supplementary material to this paper.

Fire growth mapping

The method used to create our fire databases was organised into Ingest, Clean, Cluster, Event, Merge, and Export data processing modules (Fig. 1). An overview of the method is provided here, with additional details in the Supplementary material.

During Ingest, MODIS and VIIRS files that contained fire detections (were not empty) were loaded into the processing stream. The fire detection location, FRP, QC codes, fire mask codes, and observation time for each fire pixel were ingested. During this step, the fire detections can be subset to the area of interest (i.e. California for this study).

Next, fire detections were cleaned by filtering based on QC (MODIS)/Type (VIIRS) and Fire Mask (MODIS)/Confidence (VIIRS) codes. We filtered out all data that had a QC or Type flag for water or coast (both day and night pixels), keeping only day/night land fire detections with nominal or high confidence. We also removed all fire detections in locations with persistent detections and agricultural land. We created an agricultural mask for California using the 2016 National Land Cover Database (NLCD 2016; Jin *et al.* 2019) data for 'cropland' and 'hay/pasture' land types. We calculated percent agricultural land cover on a fixed grid, and fire detections in locations with more than 50% agricultural land cover were removed from the dataset. For the persistent fire mask, we identified locations that showed a large number of fire detections over the full time period of the satellite record. For each location, we calculated the number of detections over the full dataset period and visually checked those grid cells with a high number (greater than 20 total detections) for possible sources of persistent fire detections, such as industrial areas. We added all screened grid cells with an identifiable persistent heat source or false detection source to the persistence mask.

The Cluster module spatially clusters data to form instantaneous fire perimeters for each satellite overpass. We used a density-based scan (dbscan) to search for and form clusters of fire pixels at each time step within 2.5 km. This spatial search radius was derived from a literature review of other fire detection clustering methodologies (Urbanski *et al.* 2009; Vilar *et al.* 2015; Benali *et al.* 2016; Ying *et al.* 2019). Once clusters were identified, each pixel was buffered using a 500-m buffer radius for MODIS detections and 300-m for VIIRS detections. The shape was then dissolved into a single fire perimeter. We performed sensitivity testing to determine the buffer radius to be used for each satellite dataset. Buffer radii between 350 and 550 m were tested for MODIS by comparing final fire perimeters to agency fire perimeters. We selected a 500-m buffer radius for MODIS pixels because it corresponded to the lowest bias between

Table 1. Summary of agency datasets used to augment MODIS and VIIRS fire event information.

Data source	Date range provided	Number of fires	Mean fire size (ha [acres])	Median fire size (ha [acres])
FPA FOD	2003–2017	116 327	35.8 [88.5]	0.10 [0.25]
ICS-209	2018–2020	4048	680 [1679]	0.04 [0.10]
GeoMAC	2018–2019	180	4653 [11 499]	97.5 [241]
NIFC	2020	113	16 303 [40 285]	1235 [3052]
CAL FIRE FRAP	2018–2019	13 050	116 [287]	0.04 [0.10]
CAL FIRE FRAP Online	2020	239	4269 [10 548]	63.5 [157]
FIRESTAT	2018–2020	1517	131 [323]	0.04 [0.10]

The FPA-FOD incorporates all agency datasets, so data from individual agencies were only used in years when FPA-FOD was unavailable (2018–2020). 2020 agency datasets were not fully finalised when this analysis was completed, therefore, some datasets were pulled from different sources for pre-2020 and 2020 fires (e.g. CAL FIRE FRAP and CAL FIRE FRAP Online).

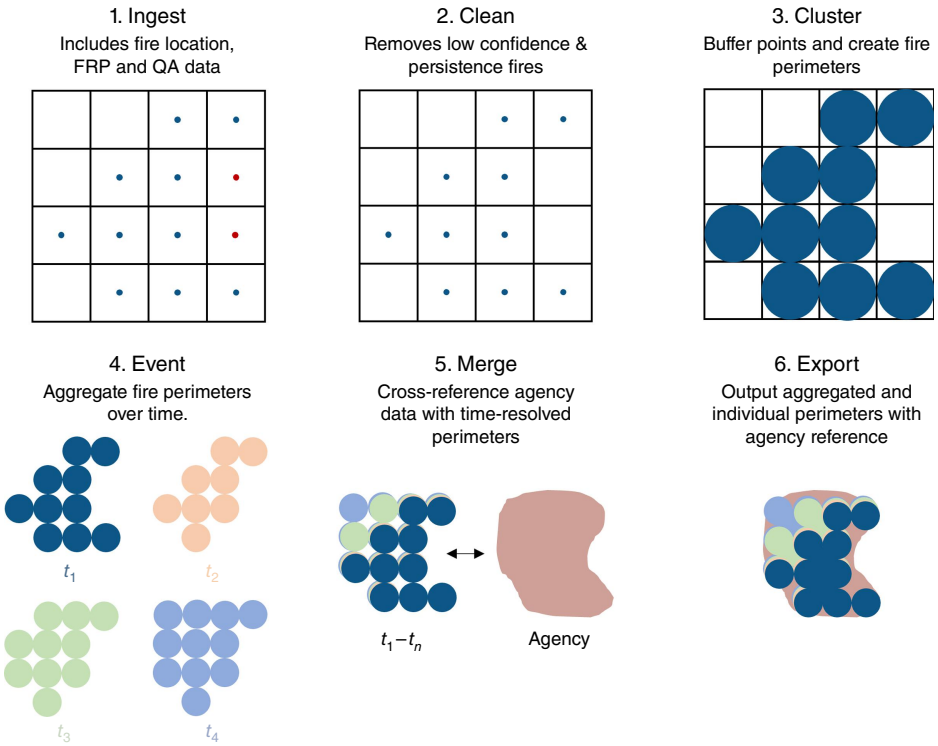

Fig. 1. Satellite fire data processing modules. Ingest adds all satellite fire detection pixel data to the dataset base, which includes fire location, FRP, and quality parameters. The Clean step removes low confidence and persistence detections. Cluster searches for nearby fire detections to aggregate, buffer, and create fire perimeters for each satellite overpass. Event searches spatially and temporally based on the results of Cluster to create a fire event record. The Merge step cross-references agency information with the satellite fire events. Export combines all information about a fire event and outputs total event, daily, and sub-daily fire perimeters with merged attributes.

the data sources. Similarly, we performed sensitivity tests for VIIRS by buffering perimeters between 187.5 and 400 m and selected a 300-m buffer radius. The sensitivity studies and buffers chosen take into account the different resolutions and geolocation accuracies for each satellite product. Once the clusters were complete, the Cluster module calculated the perimeter area, number of pixels in each cluster, total FRP, and centroid.

The Event module spatiotemporally aggregates fire perimeters into complete fire event records. Based on previous studies (Environmental Protection Agency 2015; Larkin *et al.* 2020), we used a 5-day (120-h) temporal search and a 4-km spatial search to associate fire perimeters in space and time for MODIS and VIIRS data. Sensitivity testing was also performed to derive these values, which is expanded on in the Supplementary material. Each unique fire event is given a unique identifier and associated with all growth data that occurred (information from each satellite overpass) and total fire information including final perimeter size, area, start/end date, FRP, and ignition location. Supplementary Fig. S1 and associated text provide an example of how the Event module performs the spatiotemporal aggregation in unique cases.

Spatiotemporal merging of satellite-based and agency reported fire data is commonly used to provide additional information about satellite-detected fires (Soja *et al.* 2009; Larkin *et al.* 2020). Within the optional Merge module, we use a spatiotemporal search and decision tree approach to merge agency records with satellite data. For each agency record, we identify which satellite-based fire to merge with by applying a 4-km spatial search radius and 48-h temporal search window to identify potential matches and selecting a match from all candidates using a priority-based system. After merging the agency records with fire perimeters, there can be more than one agency record associated with a single fire event. We consolidate the associated agency fire records to create a single agency record, while including information from each agency data source. The details of this agency information reconciliation step can be found in the Supplementary material.

The resulting fire activity database has three exports: (1) overall fire events; (2) daily fire growth; and (3) sub-daily fire activity and growth files. Within the Export module, holes between neighbouring pixels can be filled to make a more uniform fire perimeter (see Fig. 1 and the Supplementary material). With the inclusion of agency records in the exported products, consolidated agency name, type, cause, and ID of a fire, agency start/end date and time, agency ignition location, agency reported fire area, and all reported agency names for a particular fire event are recorded. Suspect data flags are also added to fire events when there is an area discrepancy between satellite and agency data (more flagging details are provided in the Supplementary material).

Result evaluation

We evaluated each of the fire database export products (fire events, daily fire growth, and sub-daily fire perimeters). Fire areas evaluated in this study are derived from the fire perimeter polygons. The validation was performed using

independent observations of fire activities at the fire event, daily fire growth, and sub-daily fire perimeter levels. Annual totals were also evaluated. Overall fire event area was validated using agency-reported final fire size on a per-event basis. To evaluate the total area reported, all fire event areas were summed for each year and compared with fire area totals reported by NIFC, CAL FIRE, and FPA FOD. We also calculated statistics including fire area per year and number of daily records per day. These summaries were used to numerically validate the amount of fire area burned against yearly reported fire data and to review the seasonality of fires.

For the daily fire growth files, we validated our results using National Infrared Operation (NIROPS) and Incident Information System (InciWeb) daily fire growth and cumulative fire size. The USFS NIROPS Unit uses infrared instrumentation aboard aircrafts to image select fires each year and provides very high-resolution imaging (3 m at nadir) of fire perimeters as well as hot spots within those fire perimeters throughout the lifetime of a fire. Data were obtained from the public FTP site (National Infrared Operations 2021) and show near-daily sampling of fires with the exception of operational and weather downtime. InciWeb data was only available for 2020 fires at the time of this publication.

We selected five case study wildfire events for evaluation of daily growth: (1) Red Salmon Complex-2020; (2) August Complex-2020; (3) Hirz-Delta Fire-2018; (4) Whaleback Fire-2018; and (5) Chips Fire-2012. We selected these case studies to balance limited NIROPS plane deployment with the full scale of our fire activity dataset (2012–2020). Due to the operational nature of the data source, NIROPS data are not collected for smaller fires. For the August Complex, we only compare data for the first month of the fire when most fire activity occurred. After this period, NIROPS data for this fire is broken into two zones and is difficult to compare with our full fire perimeters. We compared the NIROPS 'Heat Perimeter' values to daily growth from the satellite data products. We also evaluated cumulative daily fire growth using both NIROPS and InciWeb data. To compare cumulative daily fire size without introducing undue bias towards large fires, we calculated and compared a cumulative fire size ratio by dividing cumulative daily size by the total size for each fire.

We also used the five case study fires to validate the sub-daily fire perimeter data using NIROPS observations. The NIROPS 'Intense Heat' perimeters were compared with the individual, sub-daily fire perimeters from our fire database. To account for spatial resolution mismatch, we applied a uniform grid to the MODIS + NIROPS data (at 0.01°) and the VIIRS + NIROPS data (at 0.005°). Using the uniform grid, we evaluated the precision, recall, and *F*-score (i.e. a 0–1 est of accuracy calculated as the harmonic mean of the precision and recall values for each satellite-NIROPS comparison) of the fire dataset using the NIROPS product as the true observation. For every NIROPS flight, we linked the closest MODIS and VIIRS sub-daily fire perimeter both forward and back in time from the flight time. From those two choices, an analyst observed both and picked the most closely related satellite-derived fire perimeters. The reason for this choice was to minimise discrepancies caused by temporal lag between the NIROPS and satellite observations.

Results

Overall fire events

The MODIS 2003–2020 and VIIRS 2012–2020 fire event records created by our method indicate significant fire activities throughout California over the study period (Fig. 2). Compared with MODIS, VIIRS shows similar but more detailed fire perimeters. Total annual burned area from VIIRS and MODIS show similar patterns of interannual variability as NIFC, CAL FIRE, and FPA FOD-reported burned area data (National Interagency Fire Center 2020; Short 2021; CAL FIRE 2020) (Fig. 3). All data sources showed total burned area was highest in 2020 and near the lowest in 2019 as well as in 2010 and 2011. MODIS shows a typical overestimation in fire area burned (Urbanski *et al.* 2009), while VIIRS shows a more consistent fire area per year in California when compared with agency reports.

Across all fire sizes with slopes between 0.80 and 0.98 and R^2 values between 0.95 and 0.97, both MODIS and VIIRS-based exports show very strong agreement with agency-reported fire area, where agency data are available (Fig. 4). The August Complex Fire in 2020 (the largest fire in California history) heavily weights both regressions. Removing this fire shows similar statistics, with the MODIS slope increasing to 1.1 and the VIIRS slope increasing to 0.85 (R^2 of 0.94 for both). For smaller fires (less than 4047 ha/10 000 acres), VIIRS shows a better comparison with agency records than MODIS, which frequently over-estimates fire sizes for small fires.

Many fires in both satellite-based datasets did not match with any agency-reported fire, and vice versa. The mean and median of unmatched agency-reported data was lower than matched data for both MODIS and VIIRS (Fig. 4). Overall, for the MODIS-derived dataset between 2003 and 2020, 20% of the fire perimeters are matched with at least one agency record, accounting for 72% of the total fire area matched. For all agency records, 24% of the fire records are unmatched, and unmatched agency records have a median fire size of 0.08 ha (0.2 acres). Manual review of unmatched fire cases indicated many were caused by agency reporting of individual fires separately that are subsequently combined into a single fire complex or duplicate fire reports within an individual agency dataset. For the VIIRS-derived dataset between 2012 and 2020, 15% of the fire perimeters are matched, representing 82% of the total fire area in

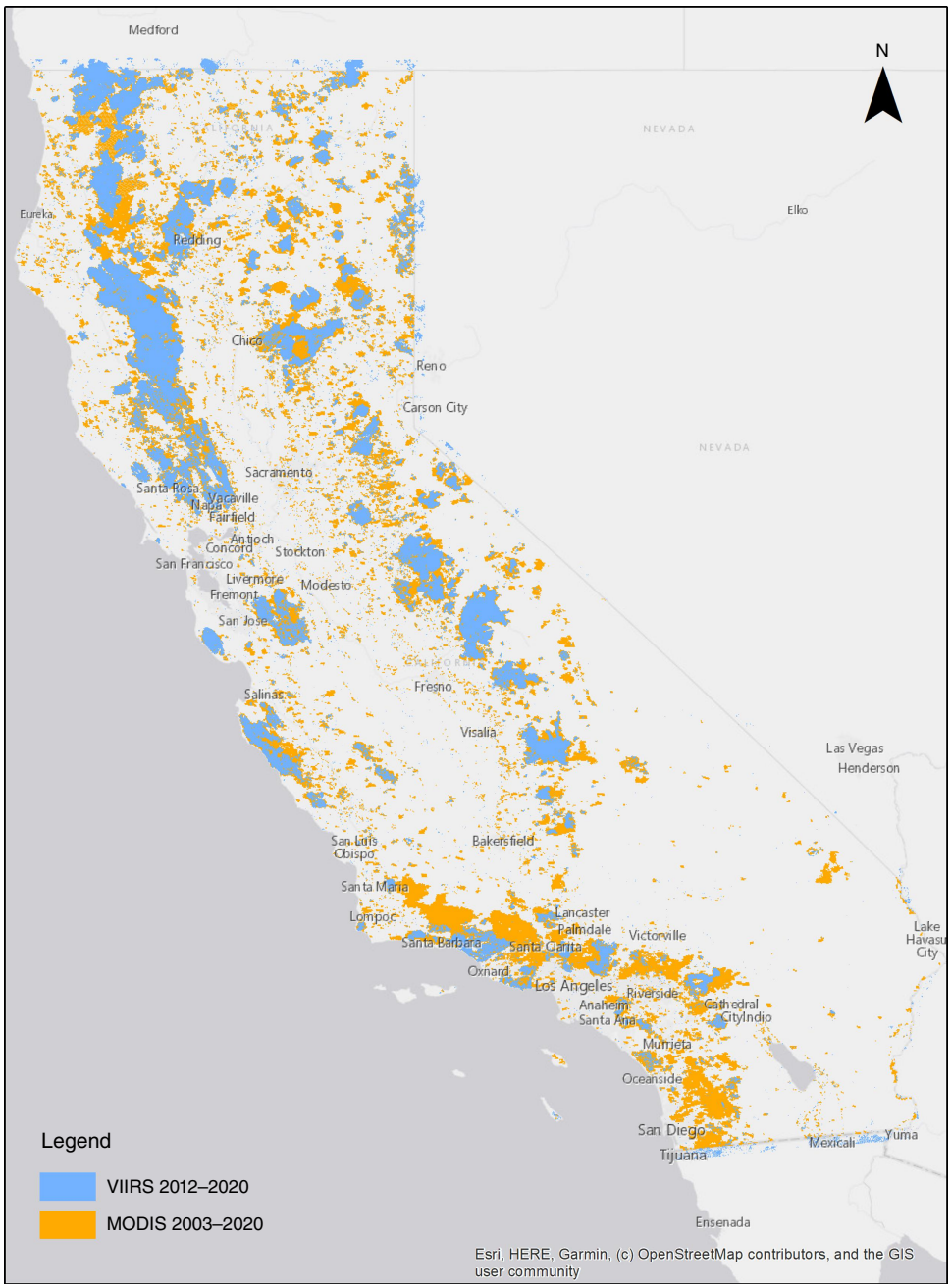

Fig. 2. MODIS (2003–2020 in orange) and VIIRS (2012–2020 in blue) final fire event perimeters are shown for California.

California matched with at least one agency record. For all agency records, this also accounts for 24% of the fire record unmatched with a median fire size of 0.04 ha (0.1 acres) and the same caveat as above. This indicates that the large fires (and the primary concern for risk and wildfire growth modelling) are well captured and matched with an agency record, while very small fires are less likely to have a successful match with agency data. Large fires (area > 4047 ha [10 000 acres]) that did not match with an agency record were reviewed manually and most frequently, the agency-reported location was found to be a significant distance from the actual fire perimeter.

Daily fire growth

Satellite daily growth compared with NIROPS daily growth shows fair agreement (R^2 values between 0.58 and 0.68) for both satellite products (Fig. 5), but both show an underestimation with a large spread at less than 4047 ha/10 000 acres of growth per day. We also compared the

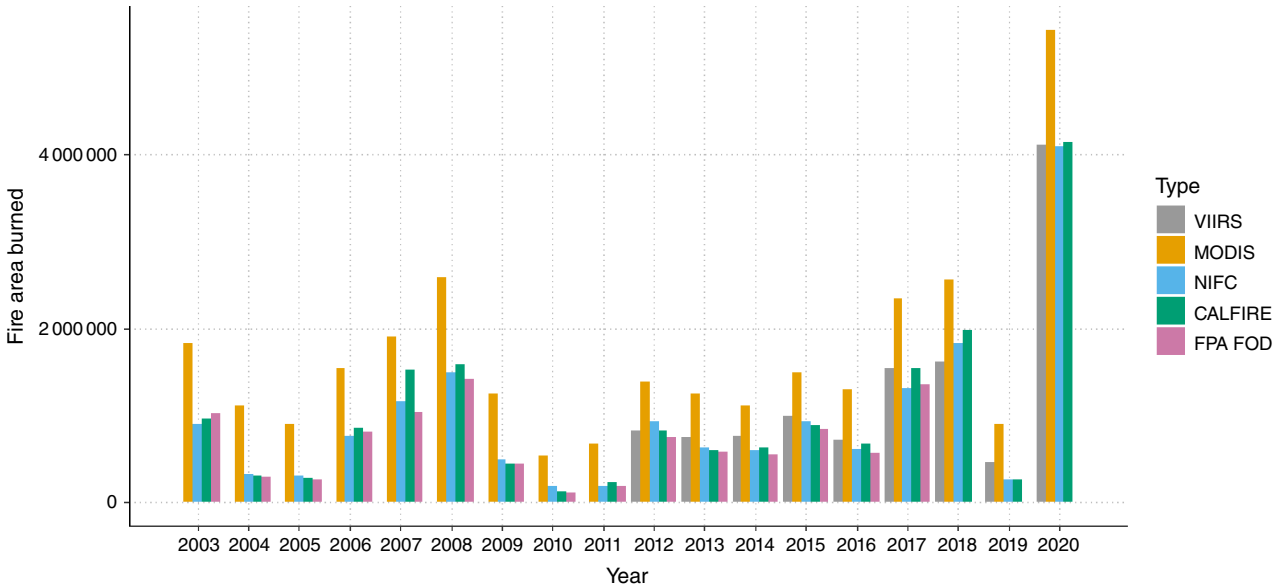

Fig. 3. Fire area burned in California between 2003 and 2020 as determined by the fire mapping method from MODIS fire detection data (orange) and VIIRS fire detection data (grey). The NIFC (blue), CALFIRE (green), and FPA FOD (pink) fire area burned in California by year is shown for reference.

agency-estimated daily growth (which is calculated by taking the fraction of fire growth each day and multiplying by the agency-reported fire area) with NIROPS area. The slope between these two datasets is nearer to unity with higher R^2 values (0.61–0.71), but still shows the same spread at the high and low end of the spectrum.

Both satellite products show strong comparison of daily cumulative area burned versus the NIROPS and InciWeb estimates, with slopes near unity and high R^2 values (0.93–0.98). However, they also show an uptick for multiple fires near the 0.5 NIROPS/InciWeb cumulative fire size ratio. For the NIROPS case, these points are near the end of the first month of the August Complex and near the end of imaging for the Hirz-Delta Fire when flights were more sporadic and may not have picked up the total fire area and growth. Additionally, for the August Complex, a large run of fire growth occurred on 8–10 September 2020. The sub-daily satellite data is more likely to pick up a quick increase in fire size compared with daily aircraft flights. For those fires that are sampled until near-complete containment (Whaleback and Red Salmon), these show very strong comparison between the satellite and NIROPS daily cumulative fire size data. For the comparison with InciWeb data, the same issue occurs with the August Complex, where a jump in satellite data occurs during a quick run of the fire near the end of the first month.

Sub-daily fire activity

Sub-daily fire perimeters generated by our method show detailed growth of fires over time (Fig. 6). Precision, recall, and *F*-score metrics are used to evaluate data quality. All metrics use a 0–1 scale, with a higher value meaning better performance. Precision evaluates the positive predict value between our fire perimeters and the NIROPS data, recall evaluates the sensitivity, and the *F*-score provides a combined metric of precision and recall. In a representative evaluation example from the August Complex (Fig. 7), VIIRS shows a precision of 0.37 and recall of 1.00 (*F*-score = 0.54), and MODIS shows a precision of 0.40 and recall of 0.55 (*F*-score = 0.46). VIIRS shows a better comparison with the NIROPS data likely due to higher resolution imaging and the satellite overpass occurring during the NIROPS flight. It should be noted that MODIS and VIIRS are able to pick up additional detections near 39.9°N and 122.7–122.9°W where NIROPS does not scan. Across all fires reviewed, VIIRS showed high recall values (median ranging between 0.75 and 0.90) with MODIS showing lower recall values (medians between 0.40 and 0.85) (Fig. S5). The precision of both satellite products compared with NIROPS observations show a median of around 0.5 across the five events we assessed, reflecting limited scan areas for NIROPS. In the case of MODIS, coarse resolution pixels that sometimes report fire activity directly adjacent but outside NIROPS detections. We provide more detailed analysis of these events in the Supplementary material.

Observed daily and sub-daily fire behaviour

Based on the daily and sub-daily fire growth, we assess the typical fire behaviour characterised in the MODIS and VIIRS-derived datasets. In the top row of Fig. 8, we examine the percentage of total area burned by date since first detect for individual fires that lasted 5 days or longer. Percent burned area is highest for the earliest time steps and decreases over time. For 89% of fires from MODIS and 85% of fires from

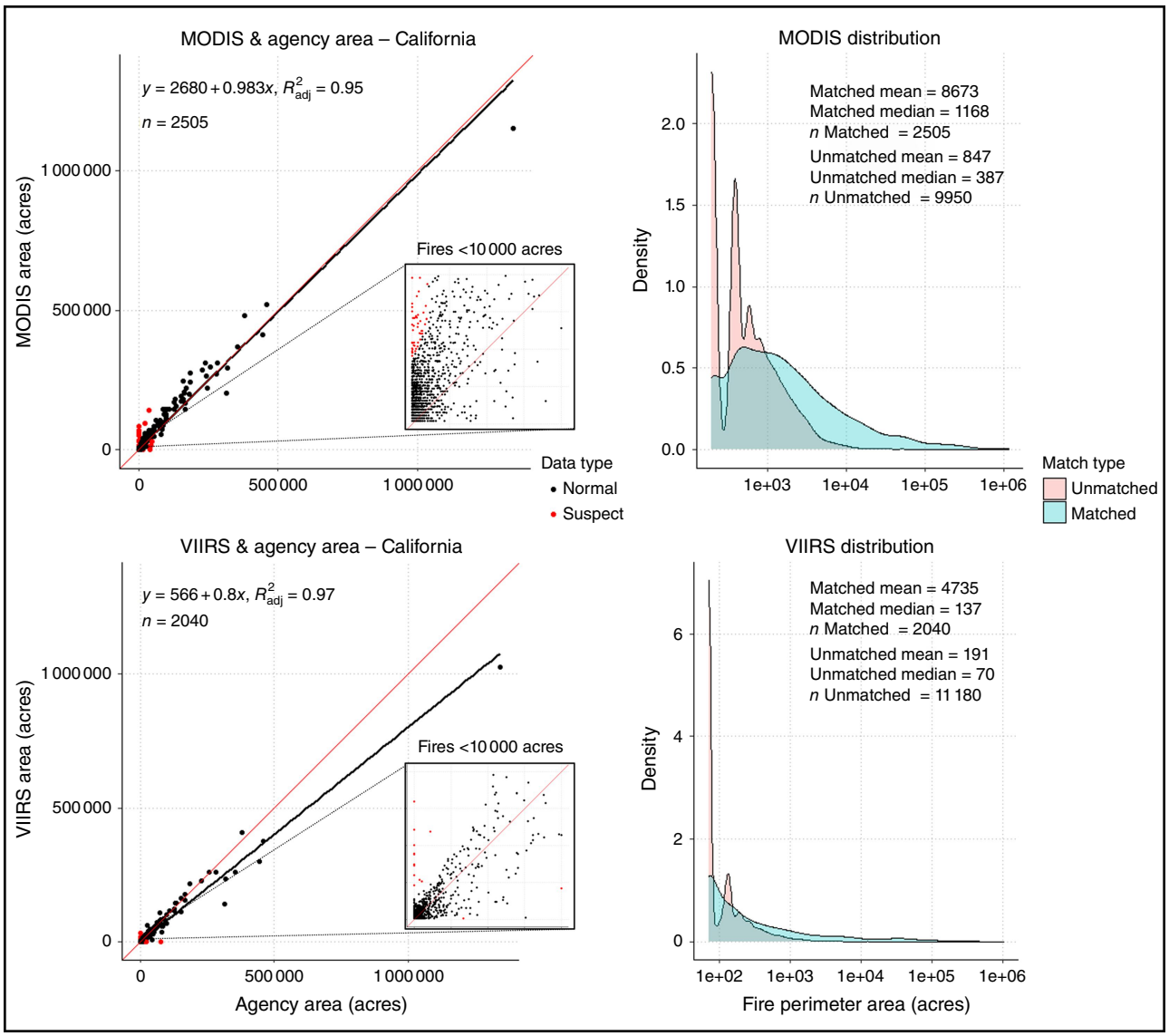

Fig. 4. MODIS (top) and VIIRS (bottom) fire area compared to matched agency records (left) and satellite matched/unmatched record distributions (right) are shown. In the scatter plots on the left, the red line indicates the 1:1 line, while the bold black line is the linear regression line with the regression equation, R^2, and count of points (*n*) in the top left of each plot. The insets show satellite compared to matched agency fire size for fires smaller than 4047 ha/10 000 acres. 'Suspect' points are shown in red. On the right, the unmatched satellite-based fire data distribution is shown in red, and the matched data distribution is shown in blue. Summary statistics about the distributions for matched and unmatched MODIS and VIIRS-based data are also shown.

VIIRS, 25% of the total burned area occurred within the first 2 weeks from ignition. In the bottom row of Fig. 8, the percentage growth by detection of the total area burned is shown. For 74% of fires from MODIS and 71% of fires from VIIRS that lasted at least 5 days, 25% of the total burned area occurred within the first 10 detections from ignition. The percentage burned per day shows more day-to-day variability than the consecutive sub-daily detections. Because only fires that lasted 5 days or longer were included in this analysis, the average fire size is 5151 ha (with an average duration of 11.5 days) for VIIRS data and 9492 ha (with an average duration of 11.7 days) for MODIS data. These results demonstrate the utility of our datasets in advancing the understanding of fire behaviour at a daily and sub-daily level by studying specific events or on a regional scale.

Discussion

Evaluation discussion

The evaluation of our new method for fire mapping with satellite data demonstrates high spatial and temporal

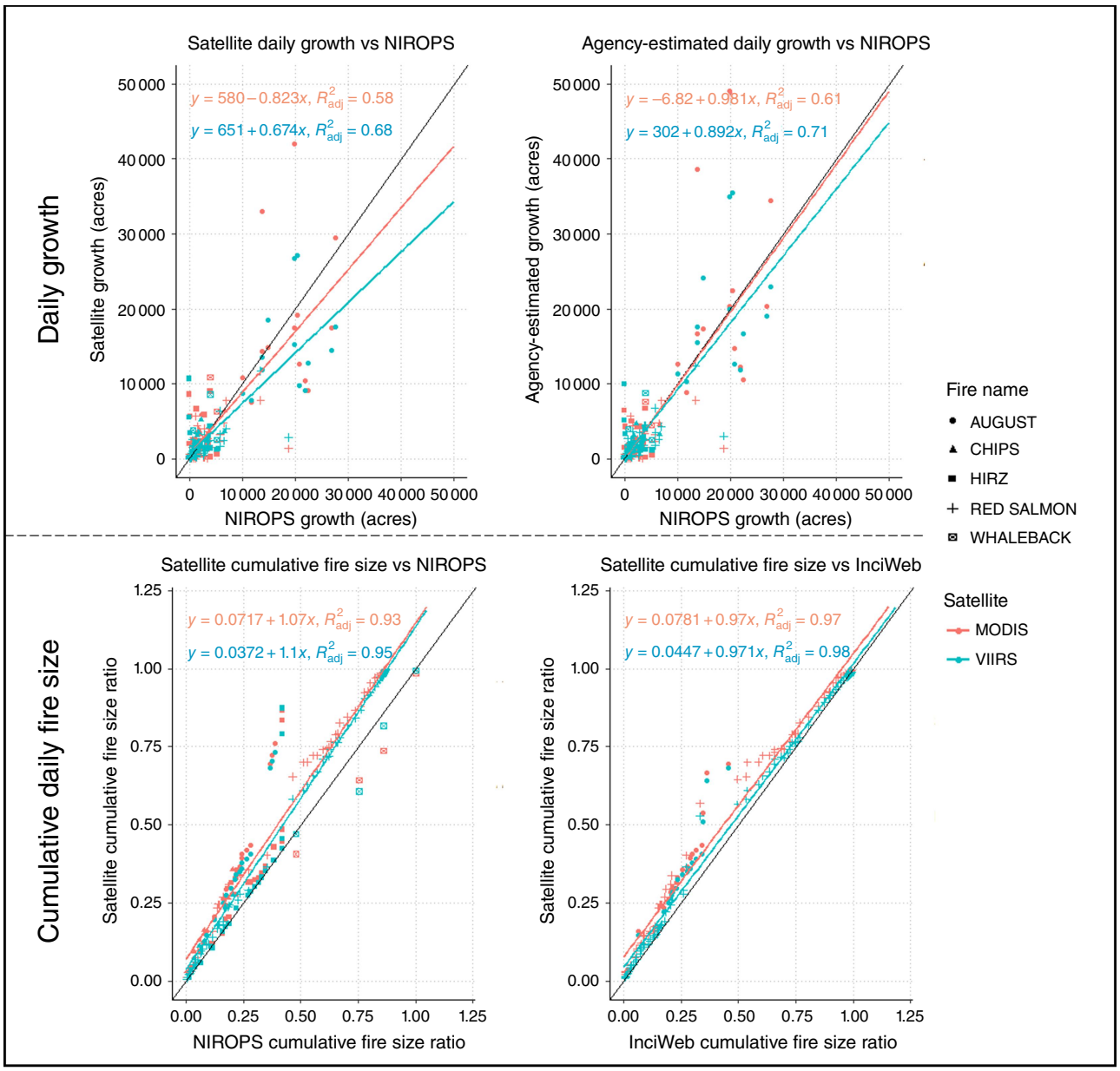

Fig. 5. Daily growth and cumulative daily fire size derived from MODIS and VIIRS fire detections are compared with NIROPS and InciWeb estimates. Daily growth comparisons with NIROPS data are shown in the top panel. Satellite-based daily growth estimates are shown in the top left and agency-estimated daily growth is shown in the top right. Cumulative daily fire size ratio comparisons with NIROPS (left) and InciWeb (right) data are shown in the bottom panel. Note that InciWeb only has data for 2020 fires, so the bottom-right figure shows only data for the August and Red Salmon Complexes.

consistency between our satellite-derived and independent agency and aircraft fire records at the sub-daily, daily, and overall event levels. For overall fire event information, we see consistent total fire area burned and capture year-to-year variations in area burned when compared with agency records. Additionally, we can match a large portion of the fire area burned with agency records (72% for MODIS and 82% for VIIRS). For the daily fire growth and cumulative daily fire area burned, we show that cumulative daily fire area burned is highly comparable with high resolution aircraft and agency data. Daily growth shows consistent results, but with more variability due to operational constraints of NIROPS, such as omitted portions of the fire and spatiotemporal mismatch. Sub-daily fire perimeters show high recall values for both satellite-derived products with lower precision values at least partially due to detecting fire areas not scanned by the aircraft.

Although agreement between satellite-based data and agency-reported data at the event-level was similar for MODIS and VIIRS, we observed important differences in

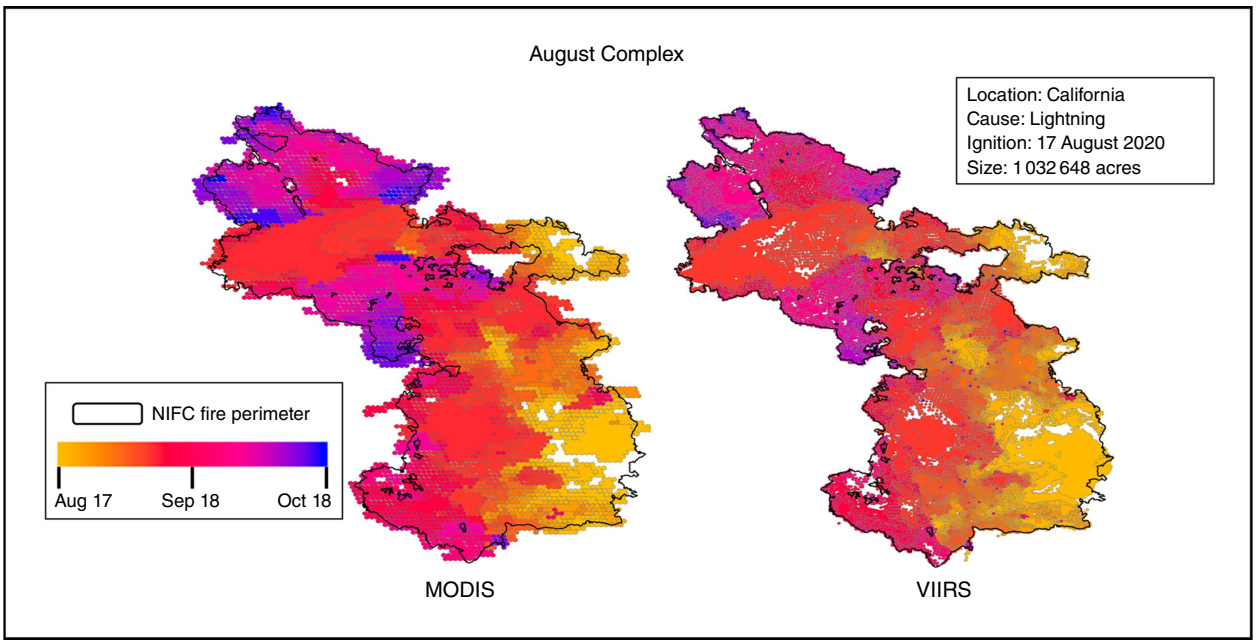

Fig. 6. Sub-daily growth from MODIS (left) and VIIRS (right) fire detections for the 2020 August Complex, colour-coded by date. The NIFC fire perimeter for the August Complex is shown as a black outline overlaid on the satellite data. Metadata on the August Complex is shown on the top right.

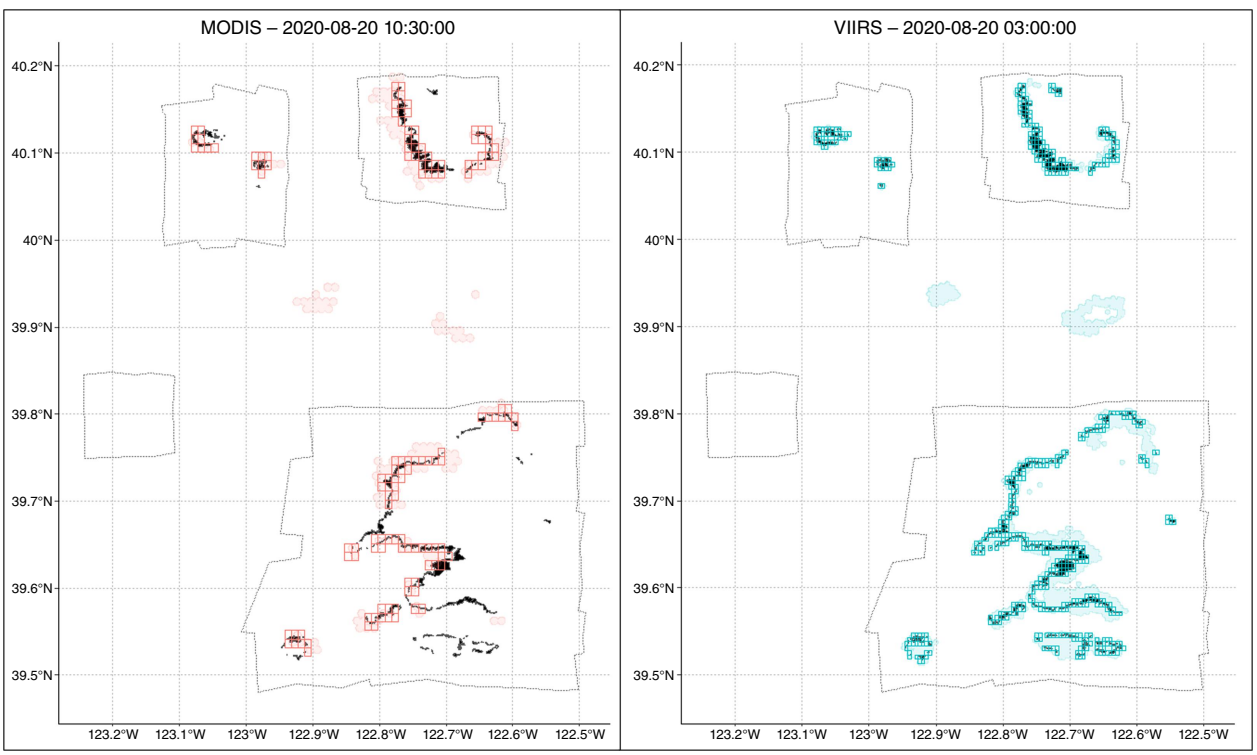

Fig. 7. MODIS- and VIIRS-derived sub-daily fire perimeters compared with the NIROPS scan closest in time of the August Complex on 20 August 2020. In both figures, the black area represents the NIROPS 'Intense Heat' fire perimeter taken during the 20 August 2020, 02:30–03:20 hours PDT flight. The red (left)- and blue (right)-filled area represents the MODIS and VIIRS sub-daily fire perimeters for the August Complex at the time stamps listed at the top of each figure. The dark red (left) and dark blue (right) grid boxes represent the grid cells where NIROPS and MODIS or VIIRS data overlap. The dotted grey line indicates the NIROPS scan area for the 20 August flight.

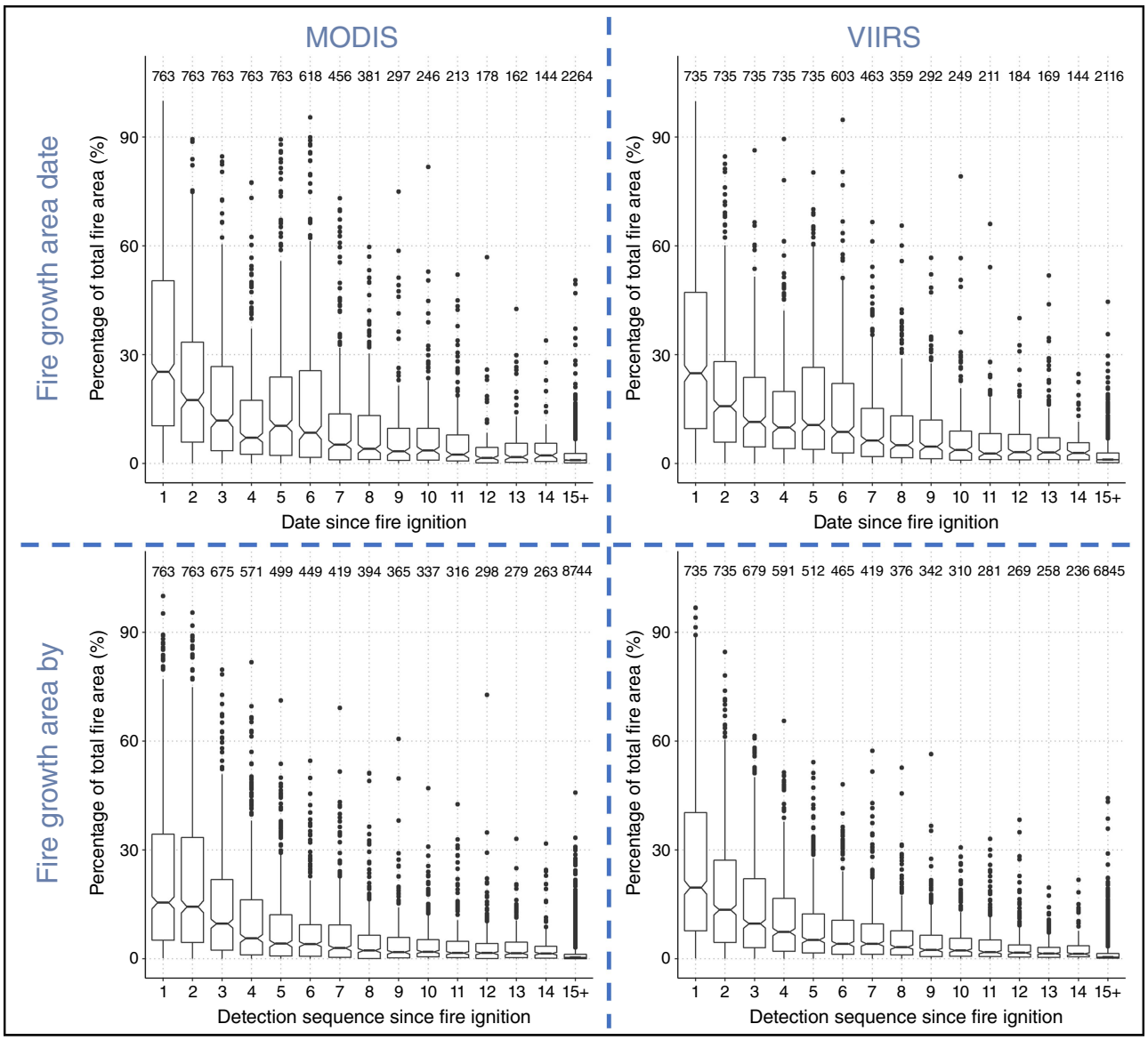

Fig. 8. Distribution of fire growth area over time since first detect. Daily burned area as a percentage of individual fire total area burned is shown in the top row, while the growth area per detection after ignition as a percentage of individual fire total area burned is shown in the bottom row. The top row uses daily fire growth data, and the bottom row uses sub-daily detection perimeters. MODIS data (2003–2020) are shown on the left and VIIRS data (2012–2020) are shown on the right. The plots include growth data from fires that burned for 5 days or longer. The number of data points per boxplot is shown above each box.

the characteristics and performance of the MODIS and VIIRS satellite products. For small fires, VIIRS event perimeters show better agreement with agency records, while MODIS events show lower agreement with agency data (Fig. 4). This lower agreement is likely due to the lower spatial resolution of MODIS observations and the fact that MODIS is less likely to detect small-scale fire activity (Hawbaker *et al.* 2008; Fusco *et al.* 2019). For MODIS fires, the minimum fire area (as single pixel fire) is 79 ha (194 acres), whereas the VIIRS minimum fire area is 28 ha (70 acres), which contributes to significantly more noise in estimating the size of small fires (these single pixel areas are derived directly from the fire dataset and are directly related to the 500-m and 300-m buffer for MODIS and VIIRS fire pixels). This effect likely also contributed to the over-estimation of annual burned area by MODIS. VIIRS showed close agreement with agency records at the annual scale. Overall, these results indicate that both MODIS and VIIRS are capable of mapping major fire perimeters with high accuracy, and that VIIRS-based fire event data shows better performance for small fires and greater spatial details. MODIS has the advantage of providing a longer-term fire record (2003–present) compared with VIIRS (2012–present).

Although the validation shows good results for matched fires between satellite fire detections and agency records,

it should be noted that many satellite-observed fires could not be validated due to a lack of matching agency data. Small fires in particular were frequently unmatched. Several factors contribute to this lack of matching. Small events, and especially prescribed fires, may not be included in the agency records used, especially since we targeted the agency fire data acquisition effort on wildfire records, not prescribed fires. In addition, small fires are more difficult to detect and match with specific agency record in space and time. For example, a fire may only be observed during a single satellite overpass and the agency record could show a significantly different start or end time. Or typically, agency records provide the location of the nearest road to the fire, but a satellite fire pixel may be too far from the road to be matched. We attempt to account for these types of discrepancies in the Merge module by using spatial and temporal buffers, but despite the sensitivity tests we used to determine our buffers, matching every fire perimeter with an agency record is not feasible. For events that were successfully matched, we see very strong agreement between datasets.

The validation of our results at the daily and sub-daily scales indicates that the satellite data products perform well at these scales. A number of data characteristics contribute to this lower agreement relative to event-level comparisons. The single greatest factor is that the NIROPS flights do not always sample the whole fire, which has shown a significant discrepancy between our satellite product and the NIROPS-derived daily fire growth. Additional factors that contribute to this lower agreement include temporal mismatches between observation times for satellite data and independent validation data, fast growing portions of fires that may burn out between satellite observations, or cloud/smoke plume interference in satellite retrievals.

While our validation results indicate that we have selected values that result in low bias overall in fire area, it may be possible to improve our method using parameters that vary based on environmental factors. For the spatial search radius, the fire spread rate is an important selection consideration, and the rate may vary based on biome, fuel loading, fuel conditions, meteorology, slope and aspect, and other factors (Linn *et al.* 2010; Carmo *et al.* 2011; Ager *et al.* 2014; Meng *et al.* 2015; Povak *et al.* 2018; Liang *et al.* 2019; Marshall *et al.* 2020; Vitolo *et al.* 2020; Chen *et al.* 2022). In conditions with rapid fire spread, our search radius may be too short to merge detections from consecutive detections. However, during visual review of our dataset, we did not find any instances of this. Our inability to find examples of this may be due to reviewing mostly fires that were greater than 400–800 ha. In future work, the potential to develop parameters that vary with conditions could be investigated to improve results and support the extensibility of the approach to new locations. Our evaluation results show excellent agreement between our satellite-based dataset and the independent observations for the study time period and domain.

Given the sub-daily sampling of fire detections by both satellites, this method and resulting datasets can provide highly spatiotemporally resolved fire information and consistently track the growth of a wildfire when agency and aircraft data cannot. While using satellite fire detection data increases the frequency of fire observations, there are inherent drawbacks to using this data. MODIS frequently underestimates or misses small fires (Hawbaker *et al.* 2008; Fusco *et al.* 2019). The VIIRS fire detection data used is not on a regular grid and therefore suffers from the 'bow-tie' effect (U.S. Department of Commerce 2017). This can cause issues for creating fire perimeters. Additionally, fire pixels can be missed or removed due to clouds or bright smoke plumes. Another limitation arises when multiple fire events that are tracked separately on the ground are associated into one event in our dataset. In these cases, the ignitions (first fire detections) and growth of adjacent events are merged and represented as a single fire event in the database. Another caveat is that agency records may merge fires that are disparate, but within the same fire management unit's domain. For example, in some agency fire area records, the 2020 Walbridge Fire in Sonoma County is listed with the 2020 LNU Fire Complex (including the Hennessey Fire), even though these two fires are far apart (around 30 km). This makes for a more complicated comparison of agency fire area and satellite-derived fire area. These effects were reduced by the sensitivity tests we performed, and we find that our Event and Merge spatial and temporal settings reduce these issues significantly.

Applications

These datasets can be applied to fire behaviour and fire weather research and model development efforts, smoke modelling, and near real-time fire monitoring. In this work, we present the results of our analysis of the percent of total area burned relative to temporal fire progression. Our results indicate that the greatest fire growth most commonly occurs during the earliest portion of the fire. This finding underscores the importance of efforts to detect fires early to support appropriate rapid response. Our results can be further used to assess a variety of characteristics of fire behaviour and growth, and our results have applicability to improve modelling efforts to represent and predict fire growth and smoke emissions.

Conclusions

Despite noted limitations, the fire activity dataset created through this method shows a very strong spatiotemporal comparison with agency records that were matched to each unique fire record and high-resolution aircraft data. This overall, daily, and sub-daily fire activity dataset provides a significant improvement over current fire datasets that usually have a 1–2-year lag in availability, are missing

an aggregated spatiotemporal fire record, and/or do not provide consistent sub-daily fire information (including fire location, area, FRP, and perimeter data). In particular, our evaluation finds that the VIIRS-based fire activity dataset provides excellent fire growth information across fine spatial and temporal scales, and MODIS also provides very good fire growth information at a coarser spatial scale but with an additional 10 years of coverage.

In addition to the attributes of the dataset itself, the method we have developed to generate this dataset has several key attributes. The strengths of this method include spatiotemporal consistency for all fire activity datasets produced as well as: (1) many configurable parameters at each data processing step (e.g. search radius, temporal search window); (2) flexibility to adapt to fire data from other satellites; and (3) ability to merge information from multiple data sources including satellites and agency/ground reports (i.e. integrating the best aspects of each satellite product into a single more accurate and comprehensive dataset). The limitations of the method and fire activity datasets produced are tied to the inherent biases and limitations of the satellite fire detection products themselves. Since this method can be applied in the event that new satellite or agency fire data is available, the fire activity data output can be continually updated. We have recently extended the resulting fire database to incorporate the western United States. The method can also be adapted in an operational setting to automatically track the growth of active fire events by accessing near real-time MODIS and VIIRS fire pixel information.

Wildfire season in the western United States has significantly increased in length and area burned in the past decade, culminating in the largest western U.S. wildfire season on record in 2020. Specifically in California, 18 of the top 20 largest wildfires have occurred since 2000, and nine of the top 10 have occurred in the last 10 years. Better understanding of the evolving fire regime is ever more important. The method we have developed provides high-resolution and consistent historical records of fire activity and growth, which can be used to study the factors driving fire behaviour (e.g. rate of spread), including weather, fuels, and topography, in great spatial and temporal detail on a regional and climatological scale. The method can also be applied in real-time to provide burned area and fire spread estimates, support smoke modelling, and improve emergency response. Other future expansions of this work include evaluating satellite-derived FRP and FRE estimates, calculating spread vector (rate and direction), and incorporating geostationary satellite fire data at the hourly level to fill in the gaps between MODIS and VIIRS detections (i.e. creating an hourly fire dataset since at least 2018).

Supplementary material

Supplementary material is available online.

References

Ager AA, Day MA, Finney MA, Vance-Borland K, Vaillant NM (2014) Analyzing the transmission of wildfire exposure on a fire-prone landscape in Oregon, USA. *Forest Ecology and Management* **334**, 377–390. doi:10.1016/j.foreco.2014.09.017

Andela N, Morton DC, Giglio L, Paugam R, Chen Y, Hantson S, van der Werf GR, Randerson JT (2019) The Global Fire Atlas of individual fire size, duration, speed and direction. *Earth System Science Data* **11**(2), 529–552. doi:10.5194/essd-11-529-2019

Artés T, Oom D, de Rigo D, Durrant TH, Maianti P, Libertà G, San-Miguel-Ayanz J (2019) A global wildfire dataset for the analysis of fire regimes and fire behaviour. *Scientific Data* **6**, 296. doi:10.1038/s41597-019-0312-2

Averett N (2016) Smoke signals: teasing out adverse health effects of wildfire emissions. *Environmental Health Perspectives* **124**(9), A166. doi:10.1289/ehp.124-A166

Balch JK, St. Denis LA, Mahood AL, Mietkiewicz NP, Williams TM, McGlinchy J, Cook MC (2020) 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**(21), 3498. doi:10.3390/rs12213498

Benali A, Russo A, Sá ACL, Pinto RMS, Price O, Koutsias N, Pereira JMC (2016) Determining fire dates and locating ignition points with satellite data. *Remote Sensing* **8**(4), 326. doi:10.3390/rs8040326

Brey SJ, Barnes EA, Pierce JR, Wiedinmyer C, Fischer EV (2018) Environmental conditions, ignition type, and air quality impacts of wildfires in the southeastern and western United States. *Earth's Future* **6**, 1442–1456. doi:10.1029/2018EF000972

CAL FIRE (2020) GIS Data. Available at https://frap.fire.ca.gov/mapping/gis-data/ [accessed 2020]

Carmo M, Moreira F, Casimiro P, Vaz P (2011) Land use and topography influences on wildfire occurrences in northern Portugal. *Landscape and Urban Planning* **100**(1–2), 169–176. doi:10.1016/j.landurbplan.2010.11.017

Chen Y, Hantson S, Andela N, Coffield SR, Graff CA, Morton DC, Ott LE, Foufoula-Georgiou E, Smyth P, Goulden ML, Randerson JT (2022) California wildfire spread derived using VIIRS satellite observations and an object-based tracking system. *Scientific Data* **9**, 249. doi:10.1038/s41597-022-01343-0

Coen JL, Schroeder W (2013) Use of spatially refined satellite remote sensing fire detection data to initialize and evaluate coupled weather-wildfire growth model simulations. *Geophysical Research Letters* **40**(20), 5536–5541. doi:10.1002/2013GL057868

Environmental Protection Agency (2015) EPA's method for the wildland fire portion of the 2014 NEI (2015 EIC). Covers EPA's methods for the Wildland Fire Portion of the 2014 NEI, with a focus on stated provided data. Available at https://www.epa.gov/air-emissions-inventories/epas-method-wildland-fire-portion-2014nei-2015-eic [accessed 2020]

Fusco EJ, Finn JT, Abatzoglou JT, Balch JK, Dadashi S, Bradley BA (2019) Detection rates and biases of fire observations from MODIS and agency reports in the conterminous United States. *Remote Sensing of Environment* **220**, 30–40. doi:10.1016/j.rse.2018.10.028

Giglio L, Schroeder W, Justice CO (2016) The Collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment* **178**, 31–41. doi:10.1016/j.rse.2016.02.054

Grajdura S, Qian X, Niemeier D (2021) Awareness, departure, and preparation time in no-notice wildfire evacuations. *Safety Science* **139**, 105258. doi:10.1016/j.ssci.2021.105258

Hawbaker TJ, Radeloff VC, Syphard AD, Zhu Z, Stewart SI (2008) Detection rates of the MODIS active fire product in the United States. *Remote Sensing of Environment* **112**(5), 2656–2664. doi:10.1016/j.rse.2007.12.008

Hung W-T, Lu C-HS, Alessandrini S, Kumar R, Lin C-A (2021) The impacts of transported wildfire smoke aerosols on surface air quality in New York State: a multi-year study using machine learning. *Atmospheric Environment* **259**, 118513. doi:10.1016/j.atmosenv.2021.118513

Jaffe DA, O'Neill SM, Larkin NK, Holder AL, Peterson DL, Halofsky JE, Rappold AG (2020) Wildfire and prescribed burning impacts on air quality in the United States. *Journal of the Air & Waste Management Association* **70**(6), 583–615. doi:10.1080/10962247.2020.1749731

Jahn W, Urban JL, Rein G (2022) Powerlines and wildfires: overview, perspectives, and climate change: could there be more electricity

blackouts in the future? *IEEE Power and Energy Magazine* **20**(1), 16–27. doi:10.1109/MPE.2021.3122755
Jain P, Coogan SCP, Subramanian SG, Crowley M, Taylor S, Flannigan MD (2020) A review of machine learning applications in wildfire science and management. *Environmental Reviews* **28**(4), 478–505. doi:10.1139/er-2020-0019
Jazebi S, de León F, Nelson A (2020) Review of wildfire management techniques—part I: causes, prevention, detection, suppression, and data analytics. *IEEE Transactions on Power Delivery* **35**(1), 430–439. doi:10.1109/TPWRD.2019.2930055
Jin S, Homer C, Yang L, Danielson P, Dewitz J, Li C, Zhu Z, Xian G, Howard D (2019) Overall methodology design for the United States National Land Cover Database 2016 Products. *Remote Sensing* **11**(24), 2971. doi:10.3390/rs11242971
Larkin NK, Raffuse SM, Huang S, Pavlovic N, Lahm P, Rao V (2020) The Comprehensive Fire Information Reconciled Emissions (CFIRE) inventory: wildland fire emissions developed for the 2011 and 2014 U.S. National Emissions Inventory. *Journal of the Air & Waste Management Association* **70**(11), 1165–1185. doi:10.1080/10962247.2020.1802365
Larsen AE, Reich BJ, Ruminski M, Rappold AG (2018) Impacts of fire smoke plumes on regional air quality, 2006–2013. *Journal of Exposure Science & Environmental Epidemiology* **28**(4), 319–327. doi:10.1038/s41370-017-0013-x
Laurent P, Mouillot F, Yue C, Ciais P, Moreno MV, Nogueira JMP (2018) FRY, a global database of fire patch functional traits derived from space-borne burned area products. *Scientific Data* **5**, 180132. doi:10.1038/sdata.2018.132
Li L, Girguis M, Lurmann F, Pavlovic N, McClure C, Franklin M, Wu J, Oman LD, Breton C, Gilliland F, Habre R (2020) Ensemble-based deep learning for estimating $PM_{2.5}$ over California with multisource big data including wildfire smoke. *Environment International* **145**, 106143. doi:10.1016/j.envint.2020.106143
Liang H, Zhang M, Wang H (2019) A neural network model for wildfire scale prediction using meteorological factors. *IEEE Access* **7**, 176746–176755. doi:10.1109/ACCESS.2019.2957837
Linn RR, Winterkamp JL, Weise DR, Edminster C (2010) A numerical study of slope and fuel structure effects on coupled wildfire behaviour. *International Journal of Wildland Fire* **19**, 179–201. doi:10.1071/WF07120
Lizundia-Loiola J, Otón G, Ramo R, Chuvieco E (2020) A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data. *Remote Sensing of Environment* **236**, 111493. doi:10.1016/j.rse.2019.111493
Loboda TV, Csiszar IA (2007) Reconstruction of fire spread within wildland fire events in Northern Eurasia from the MODIS active fire product. *Global and Planetary Change* **56**(3–4), 258–273. doi:10.1016/j.gloplacha.2006.07.015
Marshall G, Thompson DK, Anderson K, Simpson B, Linn R, Schroeder D (2020) The impact of fuel treatments on wildfire behavior in Northern American boreal fuels: a simulation study using FIRETEC. *Fire* **3**, 18. doi:10.3390/fire3020018
McClure CD, Jaffe DA (2018) US particulate matter air quality improves except in wildfire-prone areas. *Proceedings of the National Academy of Sciences* **115**(31), 7901–7906. doi:10.1073/pnas.1804353115
McLennan J, Ryan B, Bearman C, Toh K (2019) Should we leave now? Behavioral factors in evacuation under wildfire threat. *Fire Technology* **55**, 487–516. doi:10.1007/s10694-018-0753-8
Meng Y, Deng Y, Shi P (2015) Mapping forest wildfire risk of the world. In 'World Atlas of natural disaster risk'. (Eds P Shi, R Kasperson) pp. 261–275. (Springer: Berlin, Heidelberg) doi:10.1007/978-3-662-45430-5_14
Messier KP, Tidwell LG, Ghetu CC, Rohlman D, Scott RP, Bramer LM, Dixon HM, Waters KM, Anderson KA (2019) Indoor versus outdoor air quality during wildfires. *Environmental Science & Technology Letters* **6**(12), 696–701. doi:10.1021/acs.estlett.9b00599
Mitchell JW (2013) Power line failures and catastrophic wildfires under extreme weather conditions. *Engineering Failure Analysis* **35**, 726–735. doi:10.1016/j.engfailanal.2013.07.006
NASA (2020) Archive Download. Fire Information for Resource Management System. Available at https://firms.modaps.eosdis.nasa.gov/download/ [accessed 2020]
NASA (2021) LAADS DAAC. Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center. Available at https://ladsweb.modaps.eosdis.nasa.gov/ [accessed 2020]
National Infrared Operations (2021) About us. Available at https://fsapps.nwcg.gov/nirops/pages/about [accessed 2020]
National Interagency Fire Center (2020) Statistics. Available at https://www.nifc.gov/fire-information/statistics [accessed 2020]
National Land Cover Database (NLCD) (2016) NLCD 2016 Land Cover (CONUS). Available at https://www.mrlc.gov/data/nlcd-2016-land-cover-conus [accessed 2020]
Povak NA, Hessburg PF, Salter RB (2018) Evidence for scale-dependent topographic controls on wildfire spread. *Ecosphere* **9**(10), e02443. doi:10.1002/ecs2.2443
Raffuse S, Du Y, Larkin S, Lahm P (2012) Development of the 2008 Wildland Fire National Emissions Inventory. In 'Paper presented at the 20th international emissions inventory conference'. 13–16 August, Tampa, FL, USA. STI-912012-4340. (Sonoma Technology, Inc.: Petaluma, CA, USA) Available at https://www3.epa.gov/ttnchie1/conference/ei20/session2/sraffuse_pres.pdf
Reid CE, Maestas MM (2019) Wildfire smoke exposure under climate change. *Current Opinion in Pulmonary Medicine* **25**(2), 179–187. doi:10.1097/MCP.0000000000000552
Reid CE, Jerrett M, Petersen ML, Pfister GG, Morefield PE, Tager IB, Raffuse SM, Balmes JR (2015) Spatiotemporal prediction of fine particulate matter during the 2008 Northern California wildfires using machine learning. *Environmental Science & Technology* **49**(6), 3887–3896. doi:10.1021/es505846r
Reid CE, Brauer M, Johnston FH, Jerrett M, Balmes JR, Elliott CT (2016) Critical review of health impacts of wildfire smoke exposure. *Environmental Health Perspectives* **124**(9), 1334–1343. doi:10.1289/ehp.1409277
Reid CE, Considine EM, Watson GL, Telesca D, Pfister GG, Jerrett M (2019) Associations between respiratory health and ozone and fine particulate matter during a wildfire event. *Environment International* **129**, 291–298. doi:10.1016/j.envint.2019.04.033
Sá ACL, Benali A, Fernandes PM, Pinto RMS, Trigo RM, Salis M, Russo A, Jerez S, Soares PMM, Schroeder W, Pereira JMC (2017) Evaluating fire growth simulations using satellite active fire data. *Remote Sensing of Environment* **190**, 302–317. doi:10.1016/j.rse.2016.12.023
Sayad YO, Mousannif H, Al Moatassime H (2019) Predictive modeling of wildfires: a new dataset and machine learning approach. *Fire Safety Journal* **104**, 130–146. doi:10.1016/j.firesaf.2019.01.006
Scaduto E, Chen B, Jin Y (2020) Satellite-based fire progression mapping: a comprehensive assessment for large fires in northern California. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **13**, 5102–5114. doi:10.1109/JSTARS.2020.3019261
Schroeder W, Oliva P, Giglio L, Csiszar IA (2014) The new VIIRS 375 m active fire detection data product: algorithm description and initial assessment. *Remote Sensing of Environment* **143**, 85–96. doi:10.1016/j.rse.2013.12.008
Short KC (2021) Spatial wildfire occurrence data for the United States, 1992-2018 [FPA_FOD_20210617]. Forest Service Research Data Archive, Fort Collins, CO, USA. doi:10.2737/RDS-2013-0009.5
Shrestha PM, Humphrey JL, Carlton EJ, Adgate JL, Barton KE, Root ED, Miller SL (2019) Impact of outdoor air pollution on indoor air quality in low-income homes during wildfire seasons. *International Journal of Environmental Research and Public Health* **16**(19), 3535. doi:10.3390/ijerph16193535
Soja AJ, Al-Saadi J, Giglio L, Randall D, Kittaka C, Pouliot G, Kordzi JJ, Raffuse SM, Pace T, Pierce TE, Moore T, Roy B, Pierce RB, Szykman JJ (2009) Assessing satellite-based fire data for use in the National Emissions Inventory. *Journal of Applied Remote Sensing* **3**, 031504. doi:10.1117/1.3148859
Urbanski SP, Salmon JM, Nordgren BL, Hao WM (2009) A MODIS direct broadcast algorithm for mapping wildfire burned area in the western United States. *Remote Sensing of Environment* **113**, 2511–2526. doi:10.1016/j.rse.2009.07.007
U.S. Department of Commerce (2017) Visible infrared imaging radiometer suite (VIIRS) sensor data record (SDR) user's guide. Technical Report prepared for National Oceanic and Atmospheric Administration,

Washington, DC, NESDIS 142, March. Available at https://ncc.nesdis.noaa.gov/documents/documentation/viirs-users-guide-tech-report-142a-v1.3.pdf [accessed 2020]

Veraverbeke S, Sedano F, Hook SJ, Randerson JT, Jin Y, Rogers BM (2014) Mapping the daily progression of large wildland fires using MODIS active fire data. *International Journal of Wildland Fire* **23**(5), 655–667. doi:10.1071/WF13015

Vilar L, Camia A, San-Miguel-Ayanz J (2015) A comparison of remote sensing products and forest fire statistics for improving fire information in Mediterranean Europe. *European Journal of Remote Sensing* **48**(1), 345–364. doi:10.5721/EuJRS20154820

Vitolo C, Di Giuseppe F, Barnard C, Coughlan R, San-Miguel-Ayanz J, Libertá G, Krzeminski B (2020) ERA5-based global meteorological wildfire danger maps. *Scientific Data* **7**, 216. doi:10.1038/s41597-020-0554-z

Wong SD, Broader JC, Shaheen SA (2020) Review of California wildfire evacuations from 2017 to 2019. UC Office of the President: University of California Institute of Transportation Studies. doi:10.7922/G29G5K2R. Available at https://escholarship.org/uc/item/5w85z07g

Yao J, Brauer M, Raffuse S, Henderson SB (2018) Machine learning approach to estimate hourly exposure to fine particulate matter for urban, rural, and remote populations during wildfire seasons. *Environmental Science & Technology* **52**(22), 13239–13249. doi:10.1021/acs.est.8b01921

Ying L, Shen Z, Yang M, Piao S (2019) Wildfire detection probability of MODIS fire products under the constraint of environmental factors: a study based on confirmed ground wildfire records. *Remote Sensing* **11**, 3031. doi:10.3390/rs11243031

Zou Y, O'Neill SM, Larkin NK, Alvarado EC, Solomon R, Mass C, Liu Y, Odman MT, Shen H (2019) Machine learning-based integration of high-resolution wildfire smoke simulations and observations for regional health impact assessment. *International Journal of Environmental Research and Public Health* **16**, 2137. doi:10.3390/ijerph16122137

Data availability. The data that support this study will be shared upon reasonable request to the corresponding author.

Conflicts of interest. The authors declare no conflicts of interest.

Declaration of funding. This research was funded by the Pacific Gas and Electric Company.

Acknowledgements. The authors thank the colleagues who helped make this project happen, including Scott Strenfel, Richard Bagley, Evan Duffey, Sean Gilleran, and Kan Ito for contributing ideas and reviewing data products.

Author affiliation
[A]Sonoma Technology, Inc., 1450 N. McDowell Boulevard, Suite 200, Petaluma, CA 94954, USA.