10.1071/WF22048
International Journal of Wildland Fire

Supplementary Material

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]

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

*Correspondence to: Email: npavlovic@sonomatech.com

SUPPLEMENTAL INFORMATION FOR:

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

Crystal D. McClure[1], Nathan R. Pavlovic[1,*], ShihMing Huang[1], Melissa Chaveste[1], and Ningxin Wang[1]

[1] Sonoma Technology, Inc., 1450 N. McDowell Blvd., Suite 200, Petaluma, CA, 94954, USA

***** Correspondence: npavlovic@sonomatech.com; Tel.: 707-665-9900

1. Agency Data Sources

Data from each agency source were reviewed and quality controlled. Data records that were missing any relevant date information or fell outside the 2003-2020 for MODIS and 2012-2020 for VIIRS period were removed from the dataset. The incident type was reviewed and filtered to only include incidents for wildfires and prescribed burns (excluding agricultural or pile burn fires). Each dataset was subset to the state of California for this analysis and checked for duplicates on key attributes. Duplicate records were removed to obtain a single record for each fire per dataset. Records with a negative duration (the end date before the start date) were removed. For a conservative assessment on fire duration, fires that lasted longer than the following guidelines were removed from the dataset: (1) less than 41 hectares (100 acres) with a duration of more than 30 days, (2) under 405 hectares (1,000 acres) with a duration of more than 60 days, and (3) over 405 hectares (1,000 acres) with a duration greater than 180 days. Each data source was spot-checked when the quality check was complete. To merge multiple agency data records, we identified key input variables for use in data-processing scripts. The key variables for each dataset are summarized in Table S1.

1.1 USFS FPA FOD

The FPA FOD dataset contains spatial wildfire occurrence data from 2003 to 2017. The data was obtained from Karen Short at the U.S. Forest Service in May 2020. Final data was provided for 2003 through 2015, and data for 2016 and 2017 were provided as a draft version. The official documentation for the dataset is available from https://www.fs.usda.gov/rds/archive/catalog/RDS-2013-0009.4.

1.2 CAL FIRE FRAP

The FRAP Ignitions data contain fire ignition data from the Calstats database (formerly known as CAIRS). FRAP data for 2018 and 2019 was acquired from CAL FIRE in June 2020. The official documentation for the dataset is available from https://www.usfa.fema.gov/data/nfirs/support/documentation.html.

1.3 CAL FIRE FRAP Online

Raw fire ignition data for 2020 was acquired from the CAL FIRE Calstats online dataset in April 2021. The 2020 FRAP Ignitions dataset was not finalized for 2020 at the time of analysis, therefore, we chose to move forward with the information available from CAL FIRE's online database.

1.4 USFS FIRESTAT

The FIRESTAT Fire Occurrence data contain ignition points from wildland fires started on National Forest System Lands. The data was obtained for 2018 and 2019 from data.gov in July

2020 and for 2020 fires in April 2021. The official documentation for the dataset is available at https://data.fs.usda.gov/geodata/edw/edw_resources/meta/S_USA.Fire_Occurrence_FIRESTAT_YRLY.xml.

1.5 GeoMAC

The GeoMAC data contain the locations and perimeters of fires, as well as the date of each fire. The data for 2018-2019 was obtained from the National Interagency Fire Center in June 2020. The official documentation for the dataset is available at this link: https://www.arcgis.com/sharing/rest/content/items/a829aefbe4e5471490d8f3d47ca5410d/info/metadata/metadata.xml?format=default&output=html.

1.6 NIFC

The NIFC fire perimeter datasets were used to replace the GeoMAC dataset that was retired in 2019. This dataset also includes location and date ranges for each fire. The data for 2020 was obtained from NIFC in April 2021.

1.7 ICS-209

The ICS-209 data contains a record of fire incidents in two tables, "SIT209_HISTORY_INCIDENTS" and "SIT209_HISTORY_INCIDENT_209_REPORTS." The data was acquired from the SIT-209 from FAMWEB in June 2020 for 2018 and 2019 and in April 2021 for 2020 fires. The official documentation for the dataset was obtained from the SME FAM-IT Helpdesk Support.

2. Extended Methodology

2.1 Agency Data Sources

Table S1 provides a list of key variables from each agency data source that is merged into a single dataset for use in the fire activity application and during the merge step in processing.

Table S1. List of the source columns from the individual datasets used for the key columns in the final agency dataset. Some of the columns were created during the processing of the data.

Key Columns	FPA FOD	ICS-209	GeoMAC	FRAP	FRAP Online	FIRESTAT	NIFC
Start_Date	DISCOVERY_DATE_ONLY	DISCOVERY_DATE	min_date	ALARM_DATE	incident_date_created	IGNITION	min_date
Start_Time	DISCOVERY_TIME	NA	NA	ALARM_TIME	NA	NA	NA
End_Date	CONT_DATE_ONLY	EXPECTED_CONTAINMENT_DATE	max_date	CONTAIN_DATE	incident_date_extinguished	FIRE_OUT	max_date
End_Time	CONT_TIME	NA	NA	CONTAIN_TIME	NA	NA	NA
Ign_Lat	LATITUDE	POO_LATITUDE	latitude	LAT83	incident_latitude	POO_LATITUDE	latitude
Ign_Lon	LONGITUDE	POO_LONGITUDE	longitude	LON83	incident_longitude	POO_LONGITUDE	longitude
Final_Area	FIRE_SIZE	FINAL_AREA	gisacres	ACRES	incident_acres_burned	TOTAL_ACRES_BURNED	GISAcres
ID	FPA_ID	INCIDENT_IDENTIFIER	uniquefireidentifier	INCIDENTID	incident_id	FIRE_NUMBER	LocalIncidentID_Modified
Name	FIRE_NAME	INCIDENT_NAME	incidentname	FIRE_NAME	incident_name	FIRE_NAME	IncidentName
Fire_Type	FIRE_TYPE	FIRE_TYPE	FIRE_TYPE	FIRE_TYPE	fire_type	FIRE_TYPE	fire_type
Cause	STAT_CAUSE_DESCR	INCIDENT_DESCRIPTION	cause	CDF_DESC	Desc_Cause	STATISTICAL_CAUSE	Desc_Cause

2.2 Event

The Event module spatiotemporally aggregates fire perimeters into complete fire events records. Based on previous studies (Larkin et al., 2020), we use a 5-day (120-hour) temporal search and a 4-km spatial search to associate fire perimeters in space and time for MODIS and VIIRS data. To test for potential temporal splitting of fires due to lags of more than 120 hours between detections (caused by fire inactivity or detection issues due to clouds or smoke), we performed sensitivity tests by adjusting the temporal and spatial search value and found no substantial difference in the final output within the analyzed range of values for these parameters (96-144 hours and 2.5-4 km). Based on the sensitivity analysis and previous literature, the temporal and spatial search values used in this study are expected to minimize the temporal lag issue.

For each time step, we spatially and temporally aggregate fire perimeters within the spatial and temporal search windows. If a fire perimeter has no new growth after 120 hours, it is considered "completed" and removed from the active search. Each unique fire 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. On a day with multiple active fire detections, we use the centroid latitude and longitude from the first (temporally) fire perimeter to calculate the ignition location. If there were multiple distinct fire perimeters at the first timestep, the centroid is based on all distinct fire perimeters but is required to be within the geometry of one of the distinct fire perimeters. This could mean that the centroid is not at the actual ignition location if one perimeter started burning before another, but the satellite did not detect the fire until multiple perimeters were burning. While the ignition location centroid value may not be as useful for fires that have multiple distinct fire perimeters at the first timestep, the sub-daily fire growth information and perimeters contain the relevant information on actual spatial ignition location and timing.

Within the Event module, a special case occurs when a fire cluster perimeter can be associated with two or more separate fire perimeters. Under normal conditions, two possibilities can occur during association: (1) if a cluster does not match any previous fire event (nothing within the spatial or temporal search windows), that cluster becomes a fire event or (2) if a cluster intersects a previously identified fire event, the cluster will be added to the growth record and appended to the total perimeter of the existing fire event. However, if a fire perimeter intersects multiple existing fire events, we consider this a case where fires have merged, as can occur when multiple flame fronts are burning on a single fire. If this occurs, we merge all the growth records and perimeters from existing fire events and the new fire perimeter to create a new, unique fire event record. Figure S1 shows these cases visually.

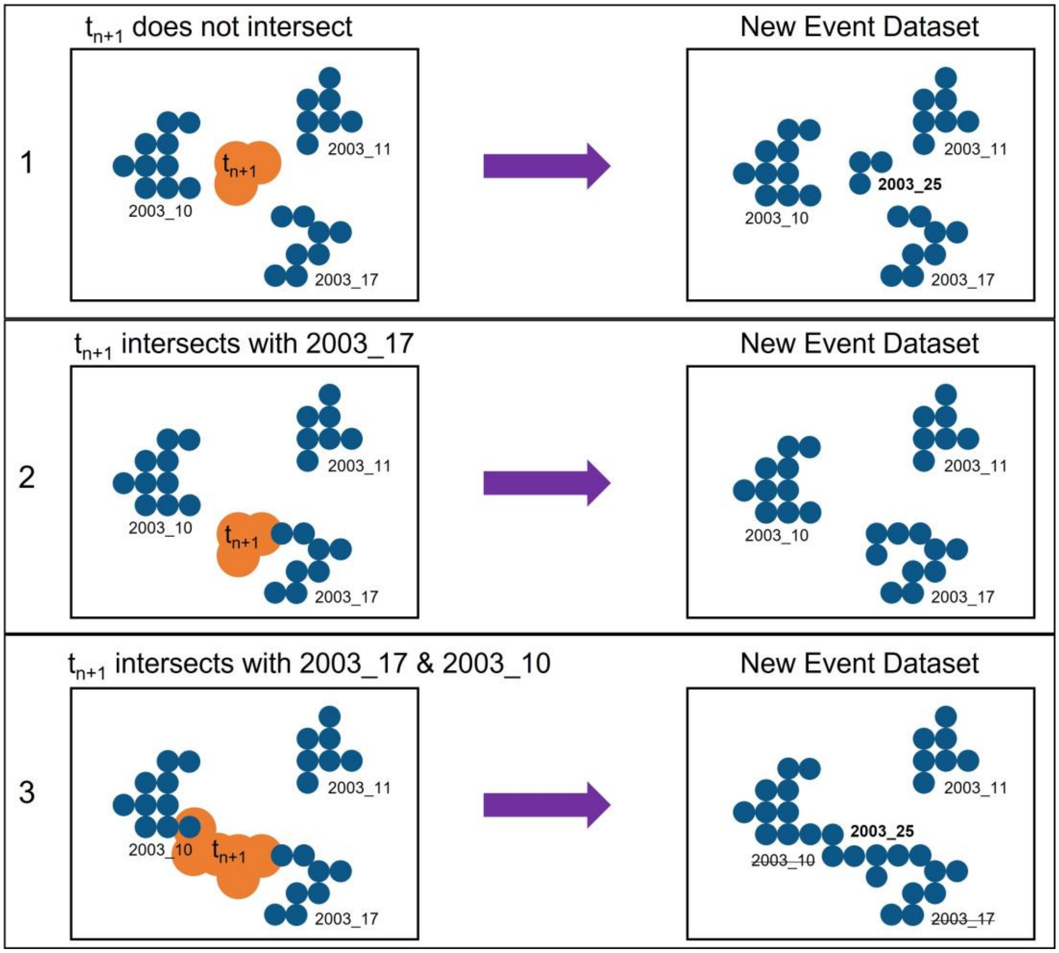

Figure S1. Event module cases. On the left, the new fire perimeter is represented by 't_{n+1}' and is shown as buffered, large orange dots. The fire events that are already identified are shown by smaller blue dots and unique fire identifier labels are added as "Year_FireNumber." On the right, we show how the Event module adds new fire perimeters into the fire event record if the new fire perimeter does not intersect (top), only intersects one fire event (middle), and intersects multiple fire events (bottom). The new fire perimeter is either given a new fire identifier (top), merged with an existing fire event (middle), or multiple fires and the new perimeters are merged into a new fire event (bottom).

2.3 Merge

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 spatial and temporal search windows to identify potential matches and selecting a match from all candidates using a priority-based system. All potential satellite fire perimeter matches for the agency record are identified by searching within a 48-hour temporal window and 4-km spatial window based on the agency information. A decision tree is used to select the best single satellite perimeter match and is described in Figure S2.

If multiple fire perimeter matches are identified for a single agency record, we step into the decision tree shown in Figure S2 to determine which fire perimeter is appropriate to merge with the agency record. In the decision tree, we determine if the unbuffered agency location intersects with any fire perimeters. If more than one or none intersect with the agency location, we use start date, location, and agency fire area to compare with the fire event records to determine the best match. In the decision tree, we choose to trust the agency start date of a fire more than agency fire area. This is because the areas reported in agency data can be very small (i.e., 0.01 acres) or just round-number estimates, which do not correlate well with a 300- or 500-m radius satellite-derived fire perimeter.

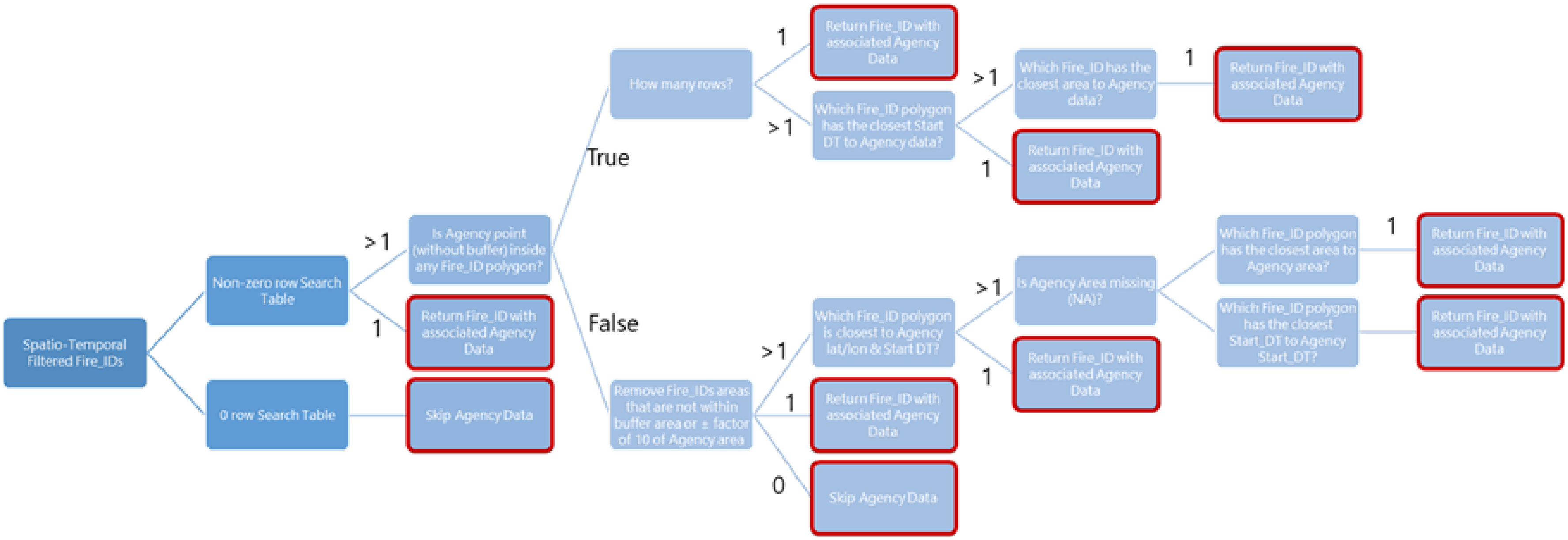

Figure S2. Merge Decision Tree. Spatiotemporally filtered fire data enters the decision tree at the left. The number indicates how many rows are left in the search table. Red outlined boxes indicate a termination of the decision tree and output of a single fire associated with an agency record. For the 'False' decision, the first step can be skipped if the agency area is sufficiently low (< 809 hectares/2,000 acres).

After merging the agency records with fire perimeters, more than one agency record can be associated with a single fire event. Therefore, we reconcile intra-agency (two or more records from the same agency data source) and inter-agency (two or more records from different agency sources) data records. The output of this operation consolidates duplicate fire names, areas, and locations while keeping any unique information from multiple agency sources.

We first perform an intra-agency reconciliation. Fire areas are checked for duplicates (if any values are within 0.5% of each other, they are considered duplicates and the smaller value is removed) and then summed for all intra-agency records. For agency fire name, type, cause, and ID number, we first look for the most frequent value (mode), but if there isn't a mode value, we use the value from the largest fire area record. For start date and time, we use the minimum agency record. For end date and time, we use the maximum agency record. For location, we use the agency latitude and longitude record from the minimum start date and time.

For inter-agency reconciliation, we use either a 'trust-first' method, a mode method, or a hierarchy method (in that order). For agency fire area, name, type, cause, and id, we use the 'trust-first' method. If the data sources are available, this method uses GeoMAC/NIFC for fire area and ICS-209 for all other variables. If neither of the 'trust-first' agencies are included in the record, we move on to looking for a mode value. If there is no mode value, we use the hierarchy shown in Table S2. If a higher-ranked agency (in the hierarchy) has NA for a variable, but a lower ranked agency has a non-NA value, the non-NA value is used. For start/end date and time, we start with the mode method. If no mode is found, we use the hierarchy. For agency location, we use the hierarchy table to determine which agency record to use for a fire event. By the end of this process, each matched fire event has a single, reconciled, associated agency record. The inter-agency reconciliation is only used for 2018-2020 data, during which more than one agency dataset was used.

Table S2. Hierarchy of Agency Data. The priority column lists the agency variable with the numbered columns indicating the rank of each agency for that particular variable.

Priority	1	2	3	4	5	6	7
Area	geomac	NIFC	ICS209	firestat	FRAP	FRAP_online	FPA_FOD
Name	ICS209	geomac	NIFC	firestat	FRAP	FRAP_online	FPA_FOD
Type	ICS209	geomac	NIFC	firestat	FRAP	FRAP_online	FPA_FOD
ID	ICS209	geomac	NIFC	firestat	FRAP	FRAP_online	FPA_FOD
Cause	ICS209	geomac	NIFC	firestat	FRAP	FRAP_online	FPA_FOD
Start_DT	ICS209	geomac	NIFC	firestat	FRAP	FRAP_online	FPA_FOD
End_DT	ICS209	FRAP	FRAP_online	geomac	NIFC	firestat	FPA_FOD
Location	ICS209	firestat	FRAP	FRAP_online	geomac	NIFC	FPA_FOD
Tons_Agency	ICS209	firestat	FRAP	FRAP_online	geomac	NIFC	FPA_FOD

2.4 Export

The overall fire events export includes whole fire statistics such as: a unique fire identifier, ignition latitude and longitude, start and end date, final fire perimeter size, final pixel number, and total FRE. The daily growth files include 24-, 48-, and 72-hour growth values for every day the fire was detected by a satellite. The daily growth export also includes daily FRP and FRE values to evaluate the intensity of the fire on any given day. The sub-daily fire activity and growth export includes the date and time of the satellite overpass, the centroid location of the fire perimeter during the overpass, the satellite that detected the fire, FRP, number of fire pixels, the current area of the fire during the overpass, and the amount of fire growth that occurred between the current overpass and all previous detections of that fire. This dataset is particularly valuable for sub-daily evaluation of fire growth and movement. In addition to the database exports, an overall event perimeter, daily growth perimeters, sub-daily fire perimeters, and sub-daily fire growth perimeters are exported as shapefiles to visualize and map all fire events at each level.

The 97.5th and 2.5th quantile ranges of the percent difference between merged agency reported fire area vs. satellite-derived fire area were calculated. Based on review of the data, fire events beyond these thresholds were identified as suspect and included an "S" flag in the database exports to identify when agency and satellite data were erroneously merged or have incorrect/incomplete data.

To finalize fire perimeters, we fill any holes inside fire perimeters that are ≤ 0.5 km^2 for MODIS and ≤ 0.1 km^2 for VIIRS. This fills in the small gaps in the fire total perimeter, like those in Figure 1 in the main text. For this daily growth export, we chose to calculate 24-, 48-, and 72-hour growth for each day of the fire. Growth values are calculated by identifying the fire area from midnight to midnight each day. This means that even if a fire started at 13:30 on day 1, the 24-hour growth for the first day would only be from 13:30 to 23:59 on day 1. The 48-hour growth on day 1 includes the growth from the first 24 hours and any additional growth from 24-48 hours (midnight to midnight on the second day). Similarly, the 72-hour growth includes growth from the 24- and 48-hour windows and any additional fire growth from 48-72 hours (midnight to midnight on the third day). Therefore, growth values are additive. On day 2, we exclude any growth from the first day (0-24 hours on day 1), so the 24-hour growth only includes midnight to midnight on the second day. Similarly, for the 48- and 72-hour growth on day 2, we add any growth from the third and fourth day, respectively. Figure S3 provides an example of how growth is calculated for the 24-, 48-, and 72-hour periods on days 1 and 2.

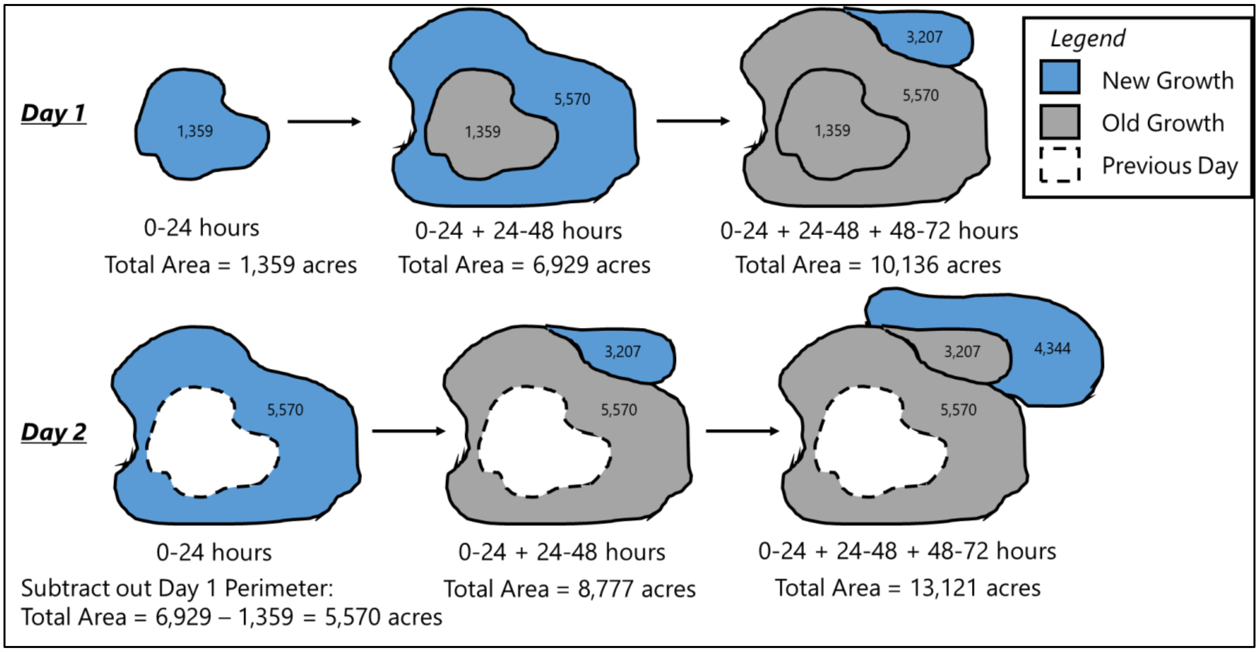

Figure S3. Daily Growth Calculation. Fire perimeters for Day 1 and 2 are shown with 24-, 48-, and 72-hour intervals. New growth in an interval is shown in blue. Old growth during an interval is shown in gray. The previous day's growth is shown unfilled with a dashed line. Fire growth is given below each interval for Day 1 and 2.

Special cases can cause zero or NA growth to be reported for daily growth. NA growth indicates that no fire detections were made during the growth window. A reported growth of zero indicates that satellite fire detections were present during the growth window, but that they did not extend beyond the perimeter from the previous time period. Three cases that can result in these values are shown in Figure S4. Zero growth (case #1) between a 24- and 48-hour growth window occurs when fire detections do not extend beyond the previous fire perimeter. A zero-growth case (#2) in the 24-hour growth period means that the fire has not extended beyond the previous day's fire perimeter (i.e., no growth). The NA growth example occurs when there are no satellite detections during a growth window. For example, if there were no satellite detections between 48 and 72 hours, we mark this growth window as NA. This is because we want to distinguish between when there is fire detection but no new growth (i.e., 0 growth), and when there are no fire detections and we have no information about growth (i.e., NA).

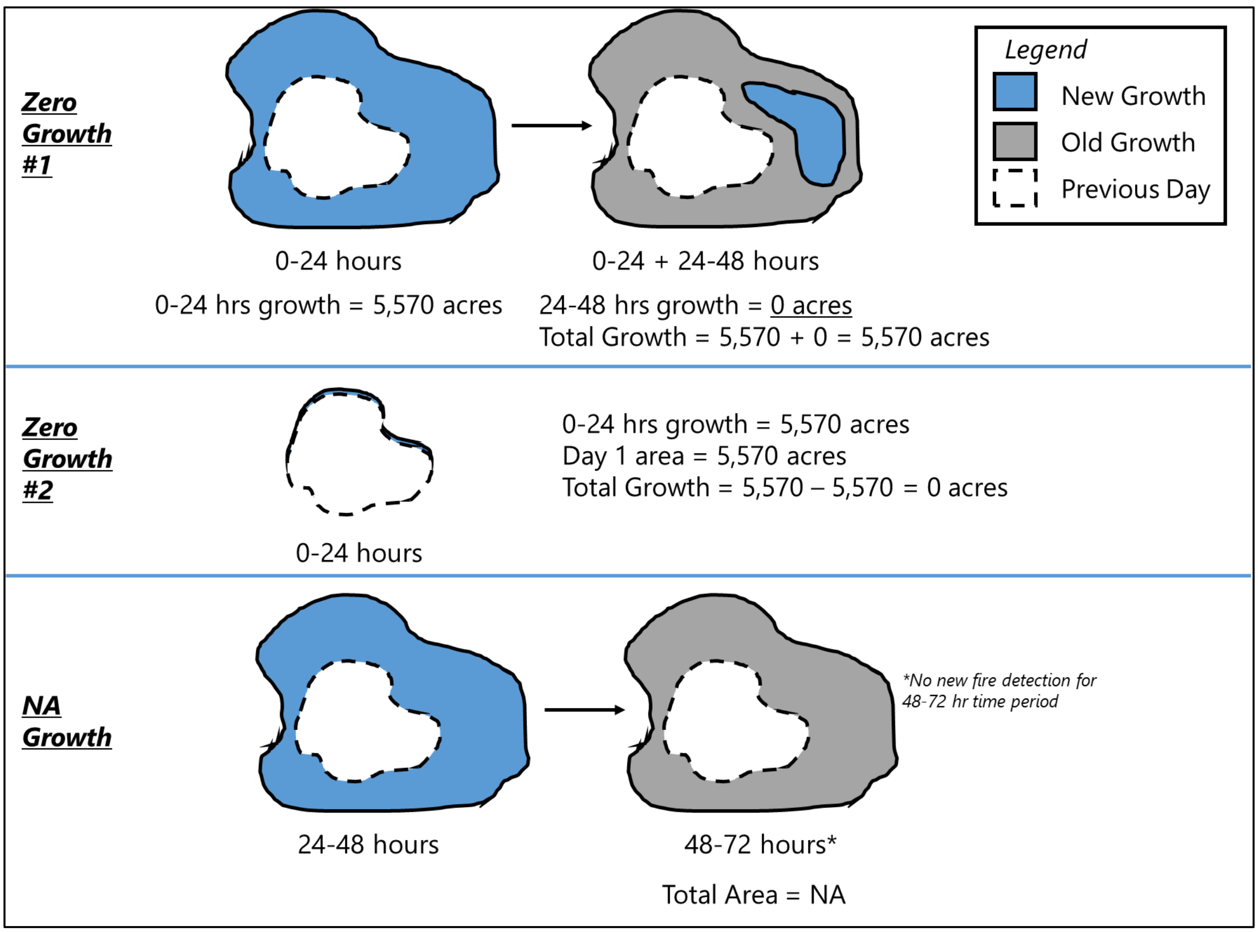

Figure S4. Fire Growth Special Cases. Zero and NA growth cases are illustrated here. New growth in an interval is shown in blue. Old growth during an interval is shown in gray.

We also calculate pixel growth for the same time windows (24, 48, and 72 hours). However, the pixel area is not derived directly from the number of pixels at each growth step. Because we remove holes from the daily fire perimeters, we need to back-calculate the number of pixels from the daily perimeters. We do this by taking the growth area per interval and dividing it by $\pi \cdot r^2$ (where r is the buffer radius used in the Cluster module [i.e., 500 m for MODIS and 300 m for VIIRS]). This gives us a non-integer number of pixels for each growth window. Buffer overlaps for individual detections can result in fractional pixels, and the non-integer is always rounded up to a whole number of pixels per growth window. We then multiply the number of pixels by the HMS-derived average fire area per MODIS pixel (https://www3.epa.gov/ttn/chief/conference/ei20/session2/sraffuse.pdf) (100 acres/pixel) or VIIRS pixel (14 acres/pixel) to estimate the actual area burned based on the number of fire detection pixels. Both fire perimeter and pixel area growth values are reported in the daily growth export, however, only the fire perimeter area is evaluated in this study. The fire pixel area is discussed here for completeness.

3. Extended Analysis

For the sub-daily analysis, we investigate whether the difference between flight time and satellite overpass causes an issue with increasing time difference. The top-left plot in Figure S5 shows the F-score (a 0-1 test of accuracy calculated from the precision and recall values for each satellite and NIROPS comparison) for MODIS and VIIRS-derived fire perimeters versus time difference. For VIIRS, we see that the F-score remains steady between 0.6 and 0.7 until greater than 12 hours' time difference. For MODIS, the median F-score is more variable, between 0.4 and 0.6 until greater than 12 hours' time when the F-score drops to 0.3. Neither shows a clear pattern of reduced F-score with time, except at greater than 12 hours' time, which could be an artifact of the very few samples at that time range (n-values at the bottom of the plot). Overall, we can compare the satellite-derived fire perimeter area vs. the NIROPS area in the top-right plot of Figure S5. Fire areas are overestimated in both satellite-derived fire perimeter datasets compared to the NIROPS perimeters, but this is likely due to the large mismatch in resolution. The overestimation is relatively consistent (r^2 = 0.51-0.61), however, with some spread likely due to difference in scan area from NIROPS vs. satellite swath, cloud cover, etc. Finally, the bottom row of Figure S5 shows the precision (left) and recall (right) values for MODIS/VIIRS versus NIROPS for each case study fire. Precision values show a median of around 0.5 for both satellite products (with exceptionally high precision for the Whaleback fire), likely due to differences in scan area and larger satellite pixels detecting heat outside of the high resolution NIROPS 'Intense Heat' perimeters. VIIRS shows very high recall values for all fires (median ranging between 0.75 and 0.90), with MODIS showing lower recall (medians between 0.40 and 0.85). This is likely due to MODIS being less likely to detect small heat signatures (Fusco et al., 2019; Hawbaker et al., 2008).

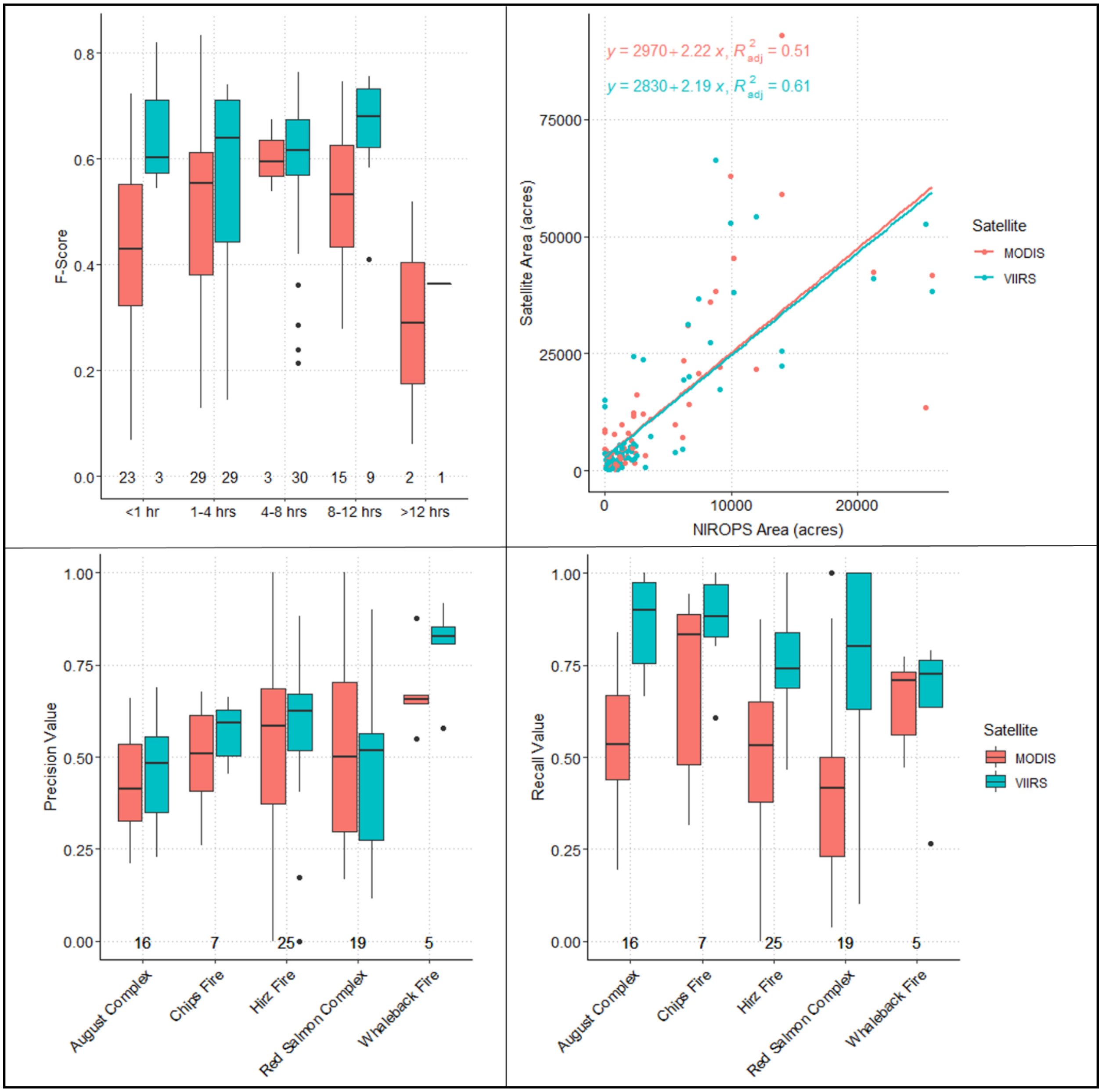

Figure S5. Validation of sub-daily fire perimeter products versus NIROPS data for the five case study fires. MODIS data is shown in red throughout all sub-plots, and VIIRS data is shown in blue. The top-left figure shows a boxplot of the F-score of each satellite-derived fire perimeter compared with NIROPS versus the time difference between the satellite overpass and the NIROPS flight. The top-right figure shows the MODIS and VIIRS fire perimeters versus the NIROPS-reported fire perimeters. The bottom plots show a box plot of the precision (left) and recall (right) for each case study fire separated by satellite. For the box plots, the count in each box can be found at the bottom of each plot. For the scatter plot, the regression equation and r^2 value can be found in the top left, color-coded by satellite.

References

Fusco E.J., Finn J.T., Abatzoglou J.T., Balch J.K., Dadashi S., and Bradley B.A. (2019) Detection rates and biases of fire observations from MODIS and agency reports in the conterminous United States. *Remote Sensing of Environment*, 220, 30-40, 2019/01/01/. Available at https://www.sciencedirect.com/science/article/pii/S0034425718304826.

Hawbaker T.J., Radeloff V.C., Syphard A.D., Zhu Z., and Stewart S.I. (2008) Detection rates of the MODIS active fire product in the United States. *Remote Sensing of Environment*, 112(5), 2656-2664, 2008/05/15/. Available at https://www.sciencedirect.com/science/article/pii/S0034425708000023.

Larkin N.K., Raffuse S.M., Huang S., Pavlovic N., Lahm P., and 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. *J. Air Waste Manage.*, 70(11), 1165-1185, doi: 10.1080/10962247.2020.1802365, 2020/11/01. Available at https://doi.org/10.1080/10962247.2020.1802365.