Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Rapid Loss of the Sierra Nevada's Largest Trees Driven by Fire

Fabien H. Wagner

fwagner@ctrees.org

CTrees

Dan J. Dixon

University of California, Davis

Christopher W. Woodall

CTrees

Mayumi C. M. Hirye

University of California, Los Angeles

Felipe Saad

University of Maryland, Baltimore County

Griffin Carter

CTrees

Ricardo Dalagnol

CTrees

Lorena Alves

University of California, Los Angeles

Cynthia Creze

University of California, Los Angeles

Stephen C. Hagen

CTrees

Zhihua Liu

CTrees

Christopher Mihiar

CTrees

Adugna Mullissa

CTrees

Le Bienfaiteur Sagang Takougoum

University of California, Los Angeles

Bryan Shaddy

CTrees

Yan Yang

CTrees
Dafeng Zhang
University of California, Los Angeles
Sassan Saatchi
Jet Propulsion Laboratory

Article

Keywords: Giant sequoia, Individual tree delineation, Deep learning, Sentinel-2 MSI, Wildfire, Tree mortality

Posted Date: October 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11054037/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Rapid Loss of the Sierra Nevada's Largest Trees Driven by Fire

Fabien H. Wagner[1*], Dan J. Dixon[2], Christopher W. Woodall[1], Mayumi C. M. Hirye[1,3,4], Felipe Saad[5], Griffin Carter[1], Ricardo Dalagnol[1], Lorena Alves[1,3], Cynthia Creze[1,3], Stephen C. Hagen[1], Zhihua Liu[1], Christopher Mihiar[1], Adugna Mullissa[1,3], Le Bienfaiteur Sagang Takougoum[1,3], Bryan Shaddy[1,3], Yan Yang[1], Dafeng Zhang[1,3], Sassan Saatchi[1,3,6]

[1*]CTrees, Pasadena, CA, 91105, USA.
[2]Department of Land, Air and Water Resources, University of California, Davis, Davis, CA, 95616, USA.
[3]Institute of Environment and Sustainability, University of California, Los Angeles, Los Angeles, CA, USA.
[4]Quapá Lab, Faculty of Architecture and Urbanism, University of São Paulo, 05508080, São Paulo, SP, Brazil.
[5]Department of Geography and Environmental Systems, University of Maryland Baltimore County, USA.
[6]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Avenue, Pasadena, CA, 91109, USA.

*Corresponding author(s). E-mail(s): fwagner@ctrees.org ;

Abstract

Large trees disproportionately contribute to biomass storage, habitat structure, and ecosystem functioning. However, their distribution and health dynamics remain poorly quantified at a regional scale. Here, a deep learning model (U-Net-ID) and canopy height models derived from sub-meter aerial imagery from 2020 were used to delineate all individual trees with crown area $\geq$ 100 m^2 across the Sierra Nevada Floristic Province. The model was trained using more than 3.3 million synthetic tree crowns and achieved a median IoU of 0.602 when validated against an independent dataset of 20,273 crowns. A total of 6,515,705 large trees were mapped, occurring across approximately 78.7% of the Sierra Nevada Floristic Province. The spatial distribution of large trees showed associations

with elevation, temperature, and precipitation. Using Sentinel-2 time series from 2020 to 2025, tree health dynamics were characterized by extracting spectral trajectories for each crown and applying BFAST breakpoint detection algorithm combined with a disturbance classification framework to identify mortality, disturbance, and recovery trajectories of individual trees. Wildfires, estimated from CAL FIRE fire perimeters, were identified as the dominant driver of large-tree mortality, killing 10% of all large trees in the Sierra Nevada, with mortality strongly concentrated during the extreme 2020–2021 fire seasons.

Keywords: Giant sequoia; Individual tree delineation; Deep learning; Sentinel-2 MSI; Wildfire; Tree mortality

1 Introduction

California State has exceptional ecological importance, with over 6,500 native vascular plant species, >25% of which are endemic, and contains a major biotic region, the California Floristic Province, one of the 35 world's biodiversity hotspots [1–4]. The Sierra Nevada Mountains Floristic Province covers ∼16% of California and spans approximately 620 km from the southern Cascades to the Tehachapi Mountains, covering ∼65,582 km^2. Most of it lies within the California Floristic Province, particularly on the western slopes and foothills, which host a high diversity of native plant species, while eastern high-elevation zones transition into the Great Basin Floristic Province [5]. Elevations range from 150 m in the western foothills to 4,421 m at Mount Whitney (the highest peak in the contiguous United States), with steep eastern escarpments and rugged western slopes. Vegetation is mainly composed of mixed-conifer forests [6], and it is home to some of the largest and most iconic tree species on Earth: the giant sequoias (*Sequoiadendron giganteum*), which are endemic to the western slope of the Sierra Nevada [7, 8]. The giant sequoias have remarkable longevity, with individuals reaching 3,200 years or more, and remarkable size, e.g., the General Sherman Tree is known to be the current largest known tree on Earth, with 1,486.6 m^3 and a weight of nearly 1,400 Mg [8, 9].

In addition to *Sequoiadendron giganteum*, Sierra Nevada are also home to many other very large tree species that can commonly reach more than 30-40 m, such as ponderosa pine (*Pinus ponderosa*), Jeffrey pine (*Pinus jeffreyi*), sugar pine (*Pinus lambertiana*), Western White pine (*Pinus monticola*), incense cedar(*Calocedrus decurrens*), California red fir (*Abies magnifica*), Sierra white fir (*Abies lowiana*) or Common Douglas-fir (*Pseudotsuga menziesii*), which together contribute to the high aboveground biomass and carbon storage observed across Sierra Nevada forests [7, 8]. Regional biomass mean density can range from 92.4 to 199.2 $Mg.ha^{-1}$, reach close to 500 $Mg.ha^{-1}$ in the highest range of canopy cover and height [10], and the highest values are observed in giant sequoia stands, with reported AGB reaching up to 2,683 $Mg.ha^{-1}$ [11]. Large trees are a key element of natural forest ecosystems, disproportionately producing and storing biomass, shaping habitat structure, and supporting

biodiversity [12–14]. The endemism and abundance of large trees are among the key reasons why the Sierra Nevada is such an important region for conservation.

However, Sierra Nevada forests are under threat. It includes decline of the number of large trees in last decades [15] and decrease in carbon sequestration [16]. These forests experience increasing climatic stress such as drought and glacier disappearance [17]. They show increasing vulnerability to disturbances caused by drought stress, insect infestation such as bark beetle outbreaks, and wildfires [9, 18, 19]. For example in 2020 and 2021, high-severity fire events have occurred in the Sierra Nevada with disastrous consequences for the forests, even if fire is usually considered as a normal component of those ecosystems [20–25]. For example, while recent evidence has shown that prescribed burns can reduce giant sequoia mortality [26], it has been reported that around 17.6% of all large giant Sequoia trees died due to fire since 1984, and most of this mortality, ∼80%, occurred during the 2020 and 2021 fire events [9]. Furthermore, higher tree density surrounding giant sequoias has been associated with increased mortality following wildfire in field-plot studies [27]. Consequently, urgent efforts are needed to allow efficient forest conservation and management and adapt to those new threads at scale.

Field inventories remain the standard approach for monitoring individual trees and assessing their health status. However, field data acquisition is labor intensive, spatially limited, and difficult to maintain over large areas (such as the Sierra Nevada) and long periods of time. As a result, increasing efforts have recently been made to monitor individual trees using airborne and satellite remote sensing.

Recent studies have highlighted the capacity of remote sensing to monitor forest mortality and disturbances. Using NAIP imagery for 2020, Cheng et al. [28] estimated 91.4 million dead trees across 27.8 million hectares of vegetation in California (including the Sierra Nevada), with 60% of mortality occurring in small groups of three trees or fewer within 30 × 30 m grids, a pattern often missed by conventional monitoring. Similarly, multi-year canopy height differences (2019–2020 and 2020–2021) showed that wildfire accounted for more than 60% of large forest disturbances in California [29]. Using monthly ∼5 m forest cover maps between 2020 and 2022, Carter et al. [30] found that the Sierra Nevada experienced the highest loss rates, with 21.4% of forest area disturbed between 2020 and 2021. Using PlanetScope, Dixon et al. [24] demonstrated that time series and deep learning can detect heterogeneous patterns of tree survival and mortality at 3 m resolution following large California wildfires. More recently, Dixon et al. [26] mapped giant sequoia mortality using PlanetScope imagery across 19 groves in Sequoia and Kings Canyon National Parks following the 2020 Castle and 2021 KNP Complex wildfires. Current monitoring approaches typically fall into two categories: (i) very high-resolution, static or temporally sparse individual-tree mapping from VHR imagery such as NAIP [28, 31], and (ii) medium-resolution, high-frequency satellite products (e.g., Sentinel-2, PlanetScope, Landsat) that capture temporal dynamics but mainly at stand scale [24, 26, 29, 30]. This creates a scale mismatch between spatially explicit individual tree detection and temporally dense monitoring of forest change. Here, we aim to bridge this gap by combining large-scale individual tree mapping (crown ≥ 100 m^2) from NAIP-derived canopy height models [32] with their associated temporally dense Sentinel-2 observations at 10 m resolution.

To perform tree delineation in VHR canopy height models with high object densities, such as our Sierra Nevada CHM, the U-Net-ID model is particularly well suited [33]. It directly predicts binary inner-segment masks corresponding to tree crowns eroded by one pixel, enabling each crown to be represented as a unique inner region without any explicit instance-separation step. Unlike Mask R-CNN and related approaches [34–38], it does not rely on bounding boxes or anchor definitions, which can become limiting in dense forest scenes. In contrast to Deep Watershed Transform methods [28, 39], individual trees are directly encoded in the binary mask and can be extracted using simple polygonization post-processing, avoiding watershed reconstruction and manually tuned thresholds. This formulation makes the approach highly scalable for large remote-sensing mosaics with variable and high canopy densities, while also simplifying the learning problem into a standard semantic segmentation task that is already well resolved in deep learning with models like U-nets [40].

Once all large individual tree crowns are delineated, the temporal signal from Sentinel-2 can be extracted and used for classical time-series analyses to detect trends and abrupt changes, such as with `BFAST` [41, 42]. `BFAST` and its extensions have been used for large-scale forest disturbance monitoring with Sentinel-2, enabling the detection of changes in forest condition resulting from disturbances such as deforestation or bark beetle outbreaks [43, 44]. Sentinel-2, alone or in combination with other sensors, has also been successfully used to map tree species across large regions and ecological domains. For example, Sentinel-2 time series, in association or not with other sensors, have been used combined with deep learning to detect conspicuous tree species across their entire distribution range [45, 46]. These studies suggest that Sentinel-2 time series contain information to monitor the condition and disturbance history of individual large trees once their crowns have been delineated.

In this work, we present (i) a wall-to-wall map of large-tree density across the Sierra Nevada Floristic Province, derived using a U-Net-ID deep learning model adapted for individual tree segmentation from canopy height models. The model was trained using synthetic data and applied to an existing 2020 Sierra Nevada canopy height model. Segmentation performance was evaluated using independent polygon datasets comprising 20,273 giant sequoia trees. We also analyze the relationships between large-tree distribution and key environmental variables. We then (ii) analyze the health status of all large trees (i.e., crown area ≥ 100 m^2) across the Sierra Nevada over the period 2020–2025 using BFAST time-series analysis and a disturbance-based classification derived from Sentinel-2 time series. This remote-sensing-based classification is validated against field observations of live/dead status. Finally, (iii) we provide a regional map of tree health status and quantify the role of fire in driving disturbance and mortality patterns across the Sierra Nevada.

2 Results

2.1 Validation of the individual tree delineation model

Among the 20,273 field-delineated Sequoias polygons, only one did not intersect with a polygon from our CHM-based crown delineation (detection rate $> 99.99\%$). Comparison between the field-delineated crown areas and the intersecting predicted polygons

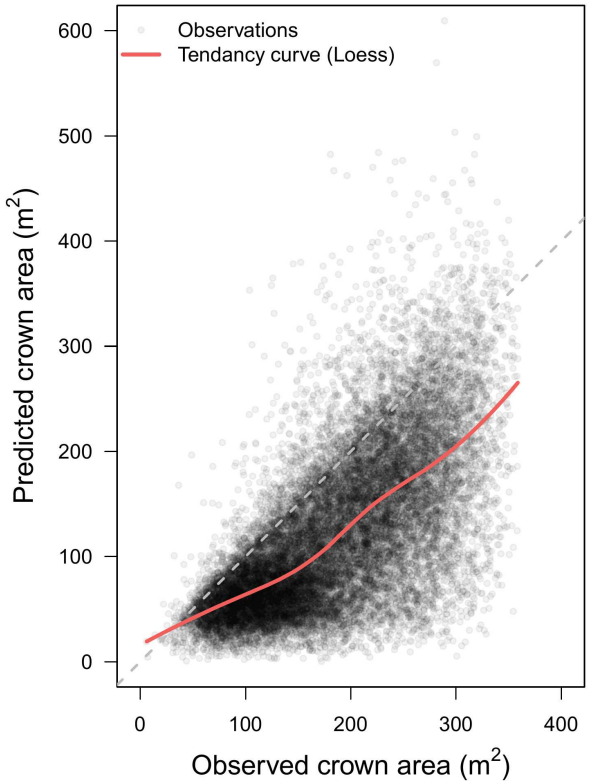

Fig. 1 Comparison of predicted versus observed crown area (m^2) for the 20,273 Sequoia trees with observed polygons delineated manually from CHM derived from LiDAR 2015/2016 data based on field GPS coordinates (1 polygon had no intersection). The 1:1 line is depicted in dashed gray. A tendency curve based on LOESS smoothing is shown to illustrate the relationship between observed and predicted crown area.

with the most similar size showed that our method tends to underestimate crown area overall, Fig. 1. The comparison resulted in an RMSE of 85.53 m^2, a Pearson correlation coefficient of 0.660, and a mean bias (predicted minus observed area) of -60.26 m^2.

Cases where the observed polygon area exceeded 300 m^2 while the predicted polygon area remained below 100 m^2 were generally associated with oversegmentation of crowns or smaller detected polygons. Conversely, for predicted polygons larger than 400 m^2, slight undersegmentation was observed, typically corresponding to the merging of two neighboring crowns, but rarely more than two.

Examples of Sequoias segmentation are presented in Fig. 2 and show, despite the systematic bias in size and some slight displacement, a good agreement between manual and U-Net-ID delineation. Note that the drawn polygon by hand can be larger

than the real crown, Fig. 2a. As the trees are on slopes, there can be a slight misalignment between the observed field data and the predicted crown delineated from CHM image, Fig. 2b.

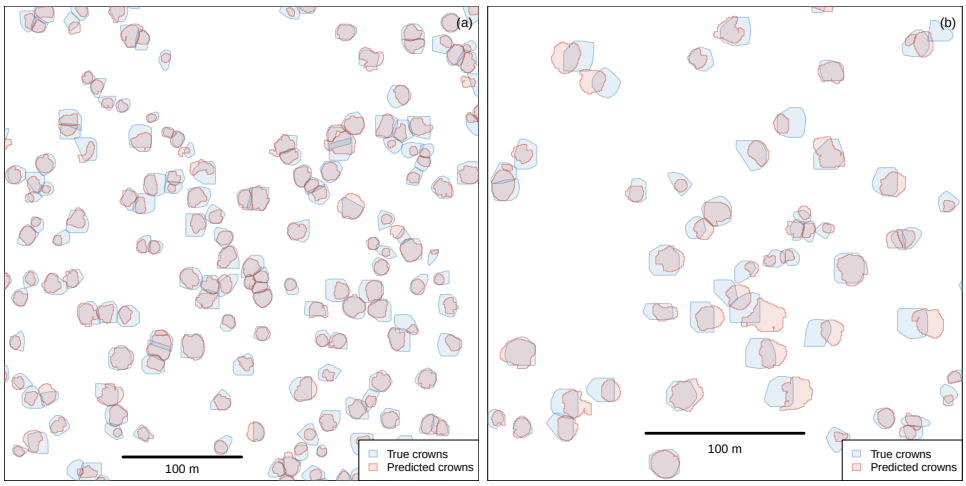

Fig. 2 Example of validation images for giant sequoia trees with manual delineation in blue and the U-Net-ID-based delineation in red.

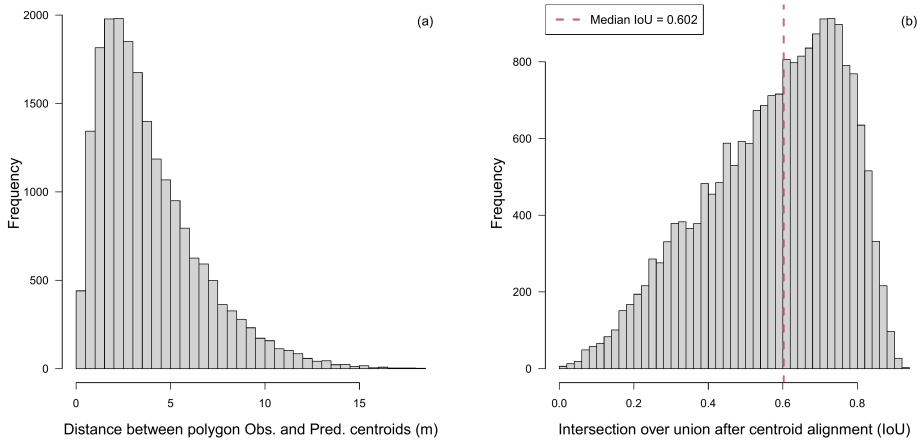

Fig. 3 Distance between the predicted and observed tree polygon centroids for the 20,272 Sequoia trees with a predicted polygon (a), and distribution of the Intersection over Union (IoU) between predicted and observed Sequoia polygons after centroid alignment (n=20,272) (b). Centroids were aligned to account for small geolocation errors between intersected predicted and observed polygons.

The median distance between observed and predicted Sequoias polygons centroids is 3.2 m (percentile 2.5–97.5: 0.54–10.64 m), Fig. 3a. This misalignment substantially

affects the Intersection over Union (IoU). Without centroid alignment, the mean and median IoU were 0.405 and 0.399, respectively, compared with 0.572 and 0.602 after alignment. Following centroid alignment, 50% of the predicted crowns had an IoU above 0.602 (Fig. 3b). After accounting for centroid alignment offset, the results indicate an improved agreement between predicted and observed crowns, supporting the robustness of the U-net-id delineation approach.

2.2 Distribution of large tree crowns across the Sierra Nevada

We identified a total of 6,515,705 large tree crowns ($\geq$ 100 m^2) distributed across approximately 78.7% of the Sierra Nevada Floristic Province (Fig. 4). Their spatial distribution is heterogeneous, with most areas ($\sim$98%) exhibiting low densities (0–5 trees ha^{-1}) and localized clusters reaching up to $\sim$20 trees ha^{-1}. Higher densities are primarily concentrated in mid-elevation forested bands along the western slope, while large portions of the southern and eastern regions show sparse distributions, with an almost complete absence of large crowns at the highest elevations (mainly along the eastern border) and lowest elevations (mainly in the west). Most of the Sequoias groves are located in the southern Sierra Nevada and show mean densities ranging from 0.08 to 10.60 trees ha^{-1}.

Some patterns appear in the association between large tree density and environmental and climate variables, Figure 5. Higher densities are observed for elevation ranging from 600 to 1700 m, and the ratio between large tree pixels and available pixels reaches a maximum around 1250 m, Figure 5a. Above 1700 m, large tree density gradually diminishes until reaching zero. Regarding slope, the distribution across classes is relatively similar to the general distribution of slope in the Sierra Nevada, Figure 5b. For orientation, there is a slight difference, with slightly more large tree pixels on north-facing slopes and slightly fewer on east-facing slopes, Figure 5c. For precipitation, there is a ratio above 1 for all classes above 800 mm, indicating that more large trees are observed where precipitation is higher, Figure 5d. For minimum temperature, there is a ratio above 1 for all classes between 2.2 and 9 °C, Figure 5e. For maximum temperature, Figure 5f, the pattern is very similar to elevation, indicating that they may carry similar information, with higher densities observed for maximum temperature ranging from 16 to 22 °C and a peak of densities between 19 and 20 °C.

2.3 Validation of the remotely sensed disturbance indices

Our dataset of large crown intersect with 1015 sequoias (amongst 3036) which have been monitored for health status after the 2020 Castle Fire and the 2021 Windy and KNP Complex fires [47]. These 1015 Ground-based observations of alive and dead status show a good overall agreement with our satellite-derived classifications of health status, Table 1. The alive and alive_with_breakpoint (bkpt) classes exhibit a high consistency with field observations, with 96.7% and 95.3% of cases, respectively corresponding to ground truth. For these two classes the ndvi remains the same during the complete time series.

The bkpt_ndvi_sup80 (3) class presents a more mixed signal, with 62.0% alive and 38.0% dead, reflecting intermediate canopy disturbance conditions, and that that a

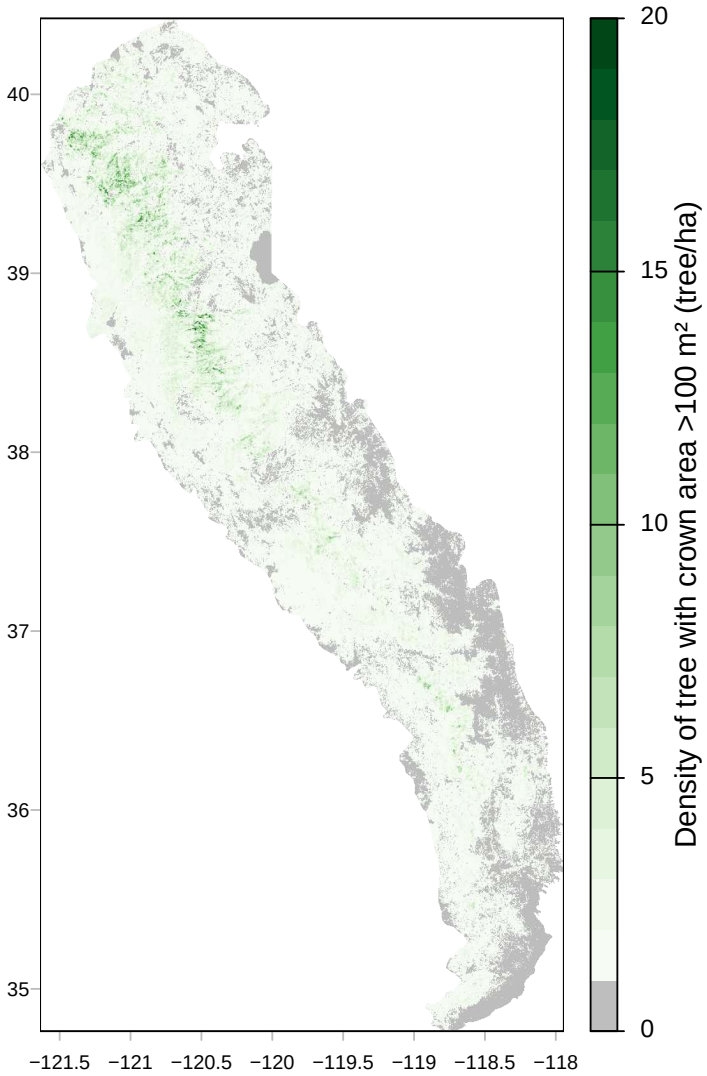

Fig. 4 Distribution of the 6,515,705 tree crowns $\geq$ 100 m^2 delineated with the U-Net-ID model in the Sierra Nevada.

quick recovery of the ndvi (more than 80% of the pre-fire ndvi) is positively associated with survival.

The bkpt_ndvi_inf80 (4) class shows a strong association with mortality, with 72.0% of cases classified as dead in ground-based observations, indicating that low NDVI breakpoint conditions are a strong indicator of mortality. The dead (6) class aligns with

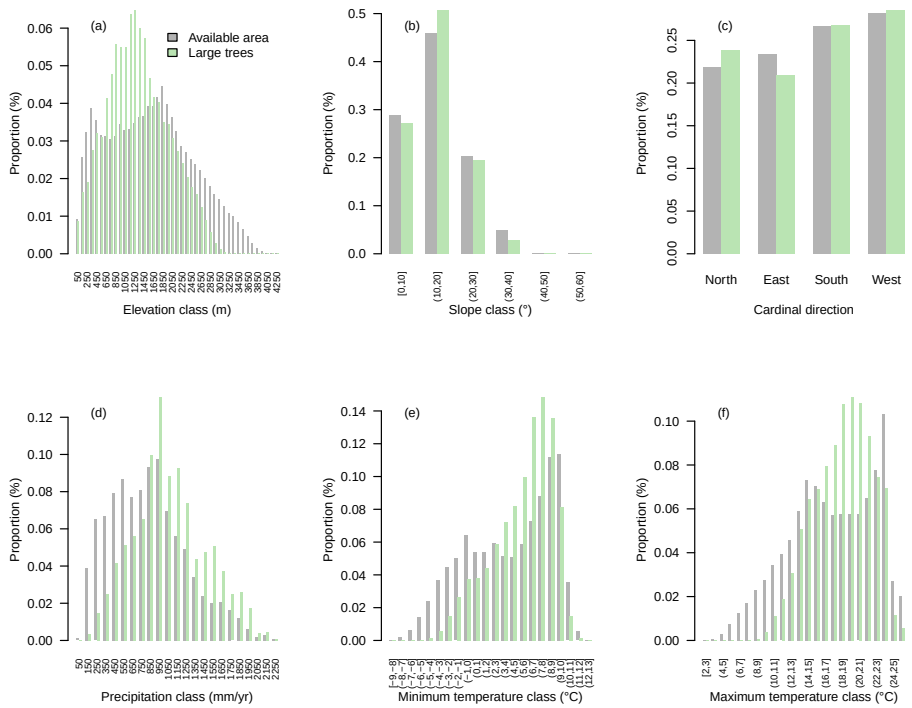

Fig. 5 Proportion of all Sierra Nevada pixels or pixels containing large trees across classes of environmental and climatic variables: (a) elevation (m), (b) slope (°), (c) slope orientation (North, East, South, West), (d) mean annual precipitation (mm yr^{-1}), (e) annual mean minimum temperature (°C), and (f) annual mean maximum temperature (°C). For elevation and precipitation, class values on the x-axis represent class means to ease visualization. Values of the ratio of observed (large trees) over expected (available area) equal to 1 indicate a uniform representation, values above 1 indicate over-representation of large trees, and values below 1 indicate under-representation.

field observations in 68.9% of cases, while 31.1% of pixels remain classified as alive, suggesting residual spectral or structural signal in standing dead trees or classification uncertainty.

The permanent_alteration (5) class is predominantly associated with alive observations (88.9%), indicating that structural change does not necessarily correspond to mortality. Finally, the already_dead_or_not_sequoia (0) class is mostly consistent with dead observations (75.0%), as expected given its definition.

Overall, results highlight strong separability between intact canopy conditions and high-confidence mortality signals, while intermediate breakpoints classes capture transitional or mixed structural states.

	Ground based observations	
	Alive	Dead
Satellite based observation		
alive (1)	202 (96.7%)	7 (3.3%)
alive_with_bkpt (2)	142 (95.3%)	7 (4.7%)
bkpt_ndvi_sup80 (3)	57 (62.0%)	35 (38.0%)
bkpt_ndvi_inf80 (4)	142 (28.0%)	365 (72.0%)
dead (6)	14 (31.1%)	31 (68.9%)
permanent_alteration (5)	8 (88.9%)	1 (11.1%)
already_dead_or_not_sequoia (0)	1 (25.0%)	3 (75.0%)
Total	566 (55.8%)	449 (44.2%)

Table 1 Ground-based observations of giant sequoia mortality compared to our remotely sensed classification of tree crown ($\geq$ 100 m^2) status classes for the 1,015 intersecting trees between both datasets.

2.4 Geographical distribution of large-tree disturbances in the Sierra Nevada and their relation to recent fires

The distribution of health status classes for tree crowns larger than 100 m^2 in the Sierra Nevada has seen by remote sensing show clear spatial patterns, Fig 6a. Large trees classified as alive, form a nearly continuous distribution along the mountain range. Areas with large trees affected by disturbances were concentrated in localized clusters where breakpoint detections and mortality classes were more frequent. Trees classified as dead represented the third visible class proportion of the total population and were primarily associated with regions that have experienced strong spectral changes.

Comparison with the fire occurrence map for the 2020–2025 period reveals a strong spatial correspondence between disturbance classes and areas affected by recent wildfires, Fig 6b. Regions with high densities of breakpoint detections, reduced NDVI recovery, and tree mortality overlapped with fire perimeters, in the northern Sierra Nevada and in several large fire complexes in the central and southern portions of the range. Substantial numbers of large trees remained classified as alive within burned areas, indicating heterogeneous fire impacts and varying levels of post-fire canopy persistence or overestimation of the fire perimeters.

Over the period 2020–2025, total fire perimeter covered 9977.758 km^2, with most of the fire concentrated in 2020 and 2021, accounting for 39.91% and 46.86%, respectively, for a total of 86.75%. The eight largest fires occurred in 2020 and 2021, Fig 6b, and 1,355,019 tree crown observations in burned areas were observed during the 2020–2021 period (87.93%).

Outside fire perimeters, large trees were mostly classified as alive, suggesting that wildfire is the dominant driver of recent large-tree mortality and canopy alteration detected by our analysis. The overall pattern highlights the concentration of recent disturbances within major fire while confirming the persistence of extensive populations of large trees across much of the Sierra Nevada.

Of the 6.52 million large tree crowns identified across the Sierra Nevada, 1.54 million (23.6%) were located within CAL FIRE fire polygons from 2020-2025 (Table 2 and Fig 6).

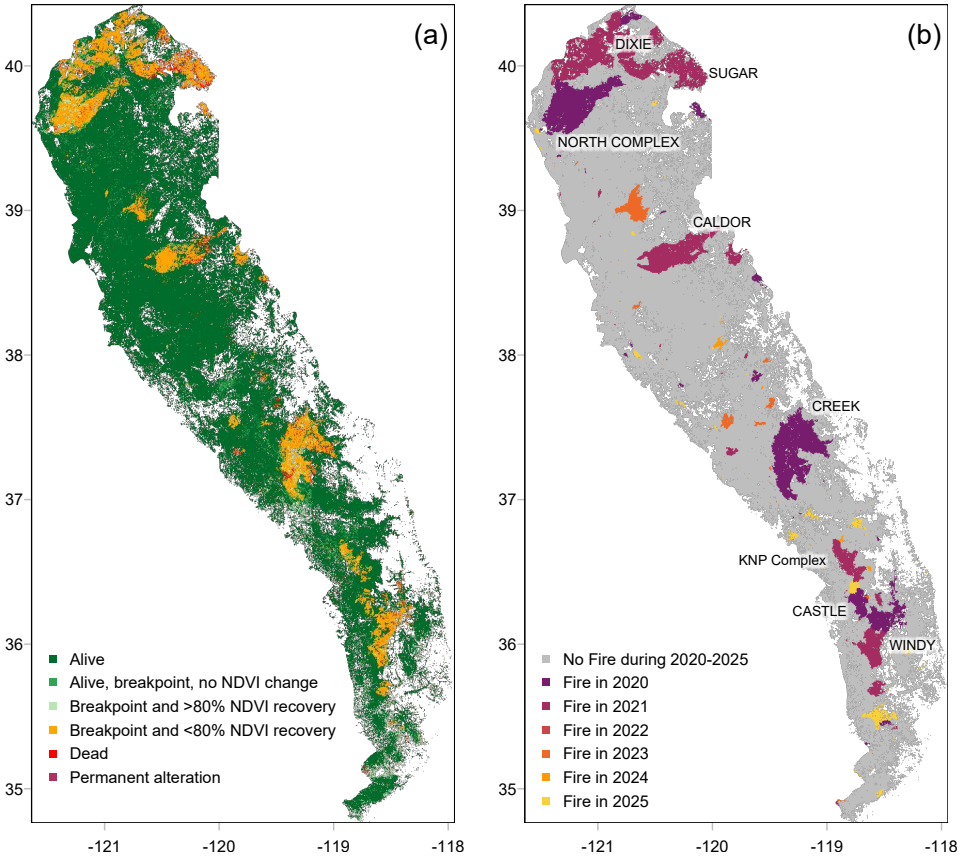

Fig. 6 Alive, disturbed and dead status classes of tree crowns $\geq$ 100 m^2 in the Sierra Nevada with similar spectral characteristics as the sequoias (a). To ease visualization, pixel values represent the most represented class at 500 m spatial resolution. (b) Fire occurrence during the 2020-2025 period in the Sierra Nevada at 500 m spatial resolution for the pixel that contains at least one tree crowns $\geq$ 100 m^2. The names of the eight largest fires, that all occurred in 2020 and 2021, are displayed.

Most crowns were classified as alive or alive with a breakpoint but no detectable NDVI change, representing 3.76 million (57.7%) and 0.81 million (12.5%) crowns, respectively. The vast majority of these crowns occurred outside burned areas (92.8% and 81.3%, respectively). Nevertheless, trees belonging to these two remotely sensed alive classes remained common within fire perimeters, accounting for 271,755 and 152,017 crowns, respectively, representing 27.5% of all large tree crowns located within burned areas.

Classes of disturbance were more concentrated within burned areas. Crowns classified as breakpoint with less than 80% NDVI recovery (750,148 crowns) showed the strongest association with wildfire, with 92.4% occurring inside fire polygons. Similarly, 79.6% of crowns classified as dead were located within burned areas. The breakpoint with greater than 80% NDVI recovery class exhibited an intermediate pattern,

Table 2 Large tree crown health status estimated from remote sensing and their distribution within CAL FIRE fire perimeters.

Class	Class ID	Burned (N, %row, %col)	Unburned (N, %row, %col)	Total (N)
alive	1	271755 (7.23%, 17.63%)	3486934 (92.77%, 70.09%)	3758689
alive_with_bkp	2	152017 (18.71%, 9.86%)	660432 (81.29%, 13.28%)	812449
bkp_sup80	3	152543 (59.61%, 9.9%)	103346 (40.39%, 2.08%)	255889
bkp_inf80	4	692992 (92.38%, 44.97%)	57156 (7.62%, 1.15%)	750148
dead	6	146775 (79.61%, 9.52%)	37591 (20.39%, 0.76%)	184366
permanent_alteration	5	20576 (37%, 1.34%)	35029 (63%, 0.7%)	55605
previous_dead_or_diff	0	103376 (14.96%, 6.71%)	587426 (85.04%, 11.81%)	690802
not_enough_data	255	990 (12.76%, 0.06%)	6767 (87.24%, 0.14%)	7757
TOTAL		1541024	4974681	6515705

with 59.6% of crowns occurring within fire perimeters, suggesting substantial post-fire recovery in some locations.

The composition of burned areas further highlights the dominant effect of wildfire on large-tree health conditions seen by remote sensing. Within fire polygons, the breakpoint with less than 80% NDVI recovery class represented the largest category (45.0% of all crowns in burn perimeter), followed by live trees (17.6%), breakpoint with greater than 80% NDVI recovery (9.9%), and dead trees (9.5%). In contrast, unburned areas were dominated by live tree classes (alive and alive with breakpoints), which accounted for 70.1% and 13.3% respectively of all crowns outside fire perimeters.

These results indicate that recent wildfire is the primary driver of large-tree canopy alteration and mortality across the Sierra Nevada. The classes most strongly associated with severe decline, namely dead and breakpoint with less than 80% NDVI recovery, were mainly concentrated within burned areas, where they represented 146,775 and 692,992 crowns, respectively.

Assuming that respectively 31.1% and 28% of crowns assigned to the dead and breakpoint with less than 80% NDVI recovery classes are in fact alive on the ground as observed in the field validation dataset, and that 38% of crowns assigned to breakpoint with more than 80% NDVI recovery class are actually dead (Table 1), this would translate to approximately 660,000 large tree crowns within burned areas that are very likely dead. This is $\sim$ 10.1% of the Sierra Nevada large trees. This number greatly exceeds the corresponding estimates outside burned areas ($\sim$ 1.6%).

The presence of a small number of dead and severely declining trees outside mapped fire perimeters suggests that additional factors such as drought, insects, disease, logging, localized disturbances or background mortality may also contribute to large-tree mortality at a much lower magnitude than wildfire.

A substantial fraction of large trees persisted within burned landscapes—however, it is important to note that CAL FIRE fire perimeters may also include unburned patches, so this signal may partly reflect spatial heterogeneity within fire perimeters rather than survival.

Table 3 Large tree crown health status estimated from remote sensing within the sequoia groves and their distribution in CAL FIRE fire polygons.

Class	Class ID	Burned (N, %row, %col)	Unburned (N, %row, %col)	Total (N)
alive	1.00	18346 (50.92%, 41.57%)	17680 (49.08%, 76.47%)	36026
alive_with_bkp	2.00	7096 (65.67%, 16.08%)	3710 (34.33%, 16.05%)	10806
bkp_sup80	3.00	2682 (87.05%, 6.08%)	399 (12.95%, 1.73%)	3081
bkp_inf80	4.00	12223 (98.38%, 27.69%)	201 (1.62%, 0.87%)	12424
dead	6.00	2330 (96.24%, 5.28%)	91 (3.76%, 0.39%)	2421
permanent_alteration	5.00	831 (86.47%, 1.88%)	130 (13.53%, 0.56%)	961
previous_dead_or_diff	0.00	626 (40.83%, 1.42%)	907 (59.17%, 3.92%)	1533
not_enough_data	255.00	3 (75%, 0.01%)	1 (25%, 0%)	4
TOTAL		44137	23119	67256

2.5 Large-tree disturbances in the sequoia groves

Within the 97 Sequoia groves, a total of 67,256 large tree crowns were identified, of which 44,137 (65.6%) were located within CAL FIRE fire perimeters from 2020–2025 (Table 3). The distribution of health status classes shows a strong influence of wildfire even within these typically fire-adapted ecosystems. The most common class was alive (36,026 crowns), with nearly equal representation inside and outside burned areas (50.9% vs. 49.1%), although the majority of unburned groves were classified as alive (76.5% of all unburned crowns).

Disturbance-related classes were strongly concentrated within burned areas. The breakpoint with less than 80% NDVI recovery class was almost entirely associated with fire perimeters (98.4%), representing the dominant disturbed condition (27.7% of all burned crowns). Similarly, dead crowns were also overwhelmingly located within burned areas (96.2%), while remaining rare outside fire perimeters (3.8%). The breakpoint with greater than 80% NDVI recovery class also showed a strong fire association, with 87.1% of crowns occurring within burned areas.

Assuming that respectively 31.1% and 28% of crowns assigned to the dead and breakpoint with less than 80% NDVI recovery classes are in fact alive on the ground as observed in the field validation dataset, and that 38% of crowns assigned to breakpoint with more than 80% NDVI recovery class are actually dead (Table 1), this would still translate to approximately 11400 large trees within burned areas of the groves that are very likely dead, which would represent $\sim$ 17.0% of the large trees in the groves. In unburned patches, the estimate of dead trees is lower corresponding to around $\sim$ 0.54% of the large trees in the groves.

Overall, burned areas accounted for the majority of all non-alive conditions, including breakpoint, permanent alteration, and dead classes, indicating substantial post-fire canopy modification within Sequoia groves. The persistence of a substantial fraction of live crowns within fire perimeters suggests heterogeneous fire effects and partial survival of these large trees. These patterns confirm that wildfire is also the dominant recent driver of canopy disturbance within Sequoia groves.

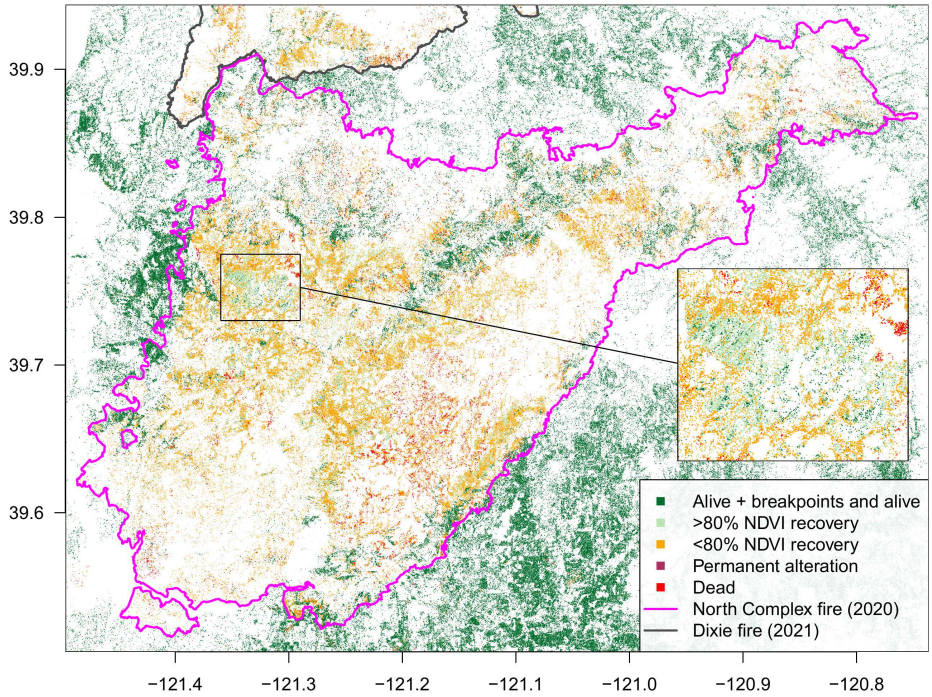

Fig. 7 Spatial pattern of large tree disturbance in the North Complex Fire. Only the main live/disturbed health classes are shown to ease visualisation.

2.6 North Complex Fire: fire intensity and disturbances

The distributions of large-tree health classes derived from remote sensing inside the North Complex Fire show clear and consistent spatial patterns of fire effects (Figure 7 and Table 4). Large occurrences of disturbed large-tree crowns are found inside the fire perimeter, dominated by breakpoints with less than 80% NDVI recovery (48%), breakpoints with more than 80% NDVI recovery (17.6%), and dead crowns (5.5%). Tree mortality (class dead) forms localized hotspots of severe impact, with the largest one in the southeastern portion of the fire. These areas were frequently embedded within broader zones dominated by the class breakpoint with less than 80% NDVI recovery. Inside impacted areas, some pockets of alive trees are found surrounded by the class breakpoint with more than 80% NDVI recovery (Fig. 7). The disturbance pattern observed from remote sensing appears to reflect a gradient of post-fire responses to fire intensity, rather than abrupt transitions between healthy and dead crowns. Living crowns (classes alive and alive with breakpoint and no change in NDVI) remained common throughout the fire perimeter (∼27%), likely indicating very low-intensity fire or unburned areas within the perimeter defined by CAL FIRE.

Table 4 Large tree crown health status estimated from remote sensing within the largest extent of the North Complex Fire in December 2020.

Class	Class ID	Burned (N, %col)
alive	1	51429 (13.09%)
alive_with_bkp	2	54832 (13.95%)
bkp_sup80	3	69085 (17.58%)
bkp_inf80	4	188490 (47.97%)
dead	6	21758 (5.54%)
permanent_alteration	5	4211 (1.07%)
previous_dead_or_diff	0	3118 (0.79%)
not_enough_data	255	18 (0%)
TOTAL		392941

3 Discussion

3.1 Mapping all large trees of a floristic province

Here, we demonstrate that large individual trees can be delineated directly from very high spatial resolution canopy height models at a regional scale. We identified 6,515,705 trees with crown areas ≥ 100 m^2 across the Sierra Nevada Floristic Province in 2020, highlighting the importance of these forests as a reservoir of large trees, including giant sequoias and other large conifers. The U-Net-Id network delineated individual crowns directly from CHM imagery with a median IoU of 0.602 on an independent validation dataset comprising 20,723 large sequoia trees. For comparison, a recent dead-tree crown delineation study conducted across California using the same 2020 imagery dataset reported an IoU of 0.53 [28]. Furthermore, the relatively high IoU obtained by our model indicates that training with synthetic data can successfully transfer to real-world conditions, demonstrating that the transfer of knowledge from synthetic to real data is feasible for large-tree crown delineation.

Our tree delineation approach differs from previous methods in three key aspects. First, while most studies perform delineation directly from very high-resolution optical imagery, our method uses canopy height models (CHMs) [32]. Training and prediction are performed on CHMs representing trees in 3D space from a nadir view, similar to LiDAR, which is particularly advantageous in mountainous regions. In NAIP imagery, varying acquisition geometries and off-nadir viewing angles can distort and spatially displace tree crowns relative to their true trunk positions [32], complicating multi-sensor or multi-date comparisons. Working directly with CHMs reduces these artefacts by relying on a consistent 3D-derived representation. In addition, in CHMs generated by deep learning models [32], individual tree segmentation is simplified, as crowns appear as smooth, blob-like structures due to CNN-based reconstruction. It remains unknown whether this behaviour generalizes to true LiDAR-derived CHMs.

The second advantage of the approach is the exclusive use of synthetic training data. By generating training labels directly from simulated crown geometries, the method removes dependence on manually annotated datasets, which are time-consuming, costly, and often subject to interpretation differences among annotators. Previous studies relied on manually delineated trees (e.g., 27,000 in [28], 14,137 in

[48]) or hybrid datasets combining manual and synthetic labels, including 434,551 self-generated trees and 2,848 hand-annotated samples in [38], and 1,506 manual crowns with 19,656 synthetic images in [36]. In contrast, we generated 3,356,877 simulated tree crowns with diverse overlap and structural configurations in a few hours, a scale not achievable through manual annotation. This framework enables the generation of diverse crown shapes, sizes, forest compositions, and interaction patterns, while producing fully reproducible labels with consistent crown boundary definitions. Unlike human annotation, identical inputs always yield identical labels, improving consistency for model training. A limitation is that simulated crowns remain idealized "blob-like" structures, well suited to conical crowns, but it is unclear how well the approach generalizes to highly irregular or fragmented crown architectures. Nevertheless, the approach performs well for large-tree mapping in the Sierra Nevada.

Finally, the third advantage of our approach is that the deep learning component is used only for segmentation, with individual tree delineation achieved through simple post-processing, while most existing methods perform instance segmentation, which is more complex, memory intensive, and relies more heavily on human-defined decisions [28, 36, 38, 48]. In our approach, the model is trained and applied solely as a segmentation algorithm, and polygon creation is performed only during post-processing. This removes decisions related to merging overlapping polygons that can arise in instance segmentation approaches such as Mask R-CNN [36]. In addition, GPU memory requirements are independent of the number of trees present in an image, unlike Mask R-CNN for example where the number of expected objects (anchors) must be specified in advance [36, 38, 49]. This is particularly important in forests, where a single image may contain a very large number of trees. As a result, prediction can be efficiently applied to entire NAIP-based CHM images containing millions of trees. The final polygonization of the binary masks is a simple GIS operation that can handle millions of objects without requiring GPU resources.

3.2 Large tree distribution in the Sierra Nevada

To our knowledge, this is the first map of all large individual trees produced for the entire Sierra Nevada Floristic Province. In contrast to [50], who estimated more than 2.7 million very large trees (DBH $\geq$ 101.6 cm) from airborne LiDAR-derived individual tree measurements and allometric equations across three Sierra Nevada landscapes covering approximately 396,000 ha, our analysis provides wall-to-wall coverage of the Sierra Nevada Floristic Province and identifies 6,515,705 large tree crowns ($\geq$100 m^2) distributed across 78.7% of the region.

The estimated number of large trees indicates that they remain a major structural component of these forests despite pre- and post-2020 disturbances, including drought, bark beetle outbreaks, and high-severity wildfires [25, 28, 50]. Large trees are ecologically important because they store a disproportionate fraction of forest biomass and carbon, provide critical wildlife habitat, and strongly influence ecosystem functioning [12–14, 51, 52]. This regional quantification provides a baseline for monitoring future changes in Sierra Nevada large trees structure and health.

While the distribution of giant sequoias and other large tree species is already known qualitatively or from localized field studies [7], for example, with giant sequoias

primarily occurring on the western slope within an elevation range of approximately 900–2700 m [8, 9], and large trees generally associated with stands of high canopy cover and density [50], our map enables a quantitative and statistical analysis of all large tree distribution across the Sierra Nevada Floristic Province. Although it remains difficult to fully attribute these patterns to natural distributions due to anthropogenic influences in parts of the region or unmeasured characteristics at this scale, such as soils and underground hydrological processes, we nevertheless observe consistent associations with environmental and climatic variables.

Large tree distribution is strongly associated with elevation, with most occurrences concentrated between 600 and 1,700 m. Above approximately 1,700 m, large trees distribution decreases gradually to reach zero near 3,000 m, suggesting a potential physiological or climatic limitation to their persistence at high elevation. Maximum annual temperature shows a very similar pattern to elevation, with large tree occurrence concentrated between 16 and 22 °C, suggesting a strong co-variation between these variables. Precipitation also shows a clear relationship, with large tree density peaking around 1,200–1,600 mm yr^{-1}, indicating a preference of the largest trees for relatively moist environments.

Our results show only a marginal over-representation of large trees on the western side compared to the eastern side, and relatively small differences between north- and south-facing slopes. This suggests that, at the scale of all large trees, topographic orientation have less importance than elevation- and climate-related gradients which appear to be the dominant controls.

3.3 Large tree health status characterization using Sentinel-2 time series

Most large trees fall within the range of giant sequoia-derived RGB values, Fig. 11. Although giant sequoia-derived spectral characteristics are not sufficient to identify the species, as previously observed for conspicuous species with distinctive yellow flowering phenology [46], they fortunately allow time series filtering and analysis for the vast majority of large crowns. Only 10.6% of large trees exhibit spectral values outside this range, indicating either already dead trees or trees with distinct spectral characteristics, most commonly associated with lower NDVI values (Table 2).

Previous studies have demonstrated that spectral indices, such as the delta Normalized Burn Ratio (dNBR), the relativized delta Normalized Burn Ratio (RdNBR), and the relativized burn ratio (RBR), are related to fire intensity and post-fire vegetation responses in California forests, the Sierra Nevada, and across the United States [9, 53–55]. However, these approaches typically rely on simple regressions between index values and fire severity, comparisons of pre- and post-fire conditions that require fire perimeter polygons, or relationships between spectral indices and live/dead tree status. Furthermore, for Sentinel-2, the computation of these indices requires the use of short-wave infrared (SWIR) bands (e.g., B12) available at 20 m spatial resolution, which is substantially larger than the crowns of most individual trees. In contrast, our approach relies exclusively on the temporal behavior of NDVI and RGB bands at 10 m resolution and does not require fire perimeter information. By identifying breakpoints

and analyzing post-breakpoint trajectories, we are able to characterize disturbance intensity and recovery dynamics at the scale of individual large trees.

Our framework provides additional insight into disturbance dynamics by capturing not only pre- and post-fire conditions, but also the temporal evolution of crown condition following fire. Overall, the approach appears reliable when compared with field estimates of disturbance severity (Table 1). However, some disturbance processes require several years before their effects can be fully observed from remote sensing data.

For example, a proportion of trees classified as dead, characterized by a substantial decline in NDVI and no apparent recovery following the breakpoint, were still alive according to field observations. In these cases, the crown was often heavily scorched, displaying brown or orange foliage, and several years may be required before recovery becomes visible in satellite imagery. This situation represents approximately 30% of the trees classified as dead or severely disturbed from remote sensing observations.

Conversely, around 38% of trees exhibiting a breakpoint followed by strong NDVI recovery were classified as dead in the field. Although these trees were dead, their NDVI values increased over time, indicating a greening signal. This recovery may result from lateral crown expansion of neighboring trees, the exposure of previously obscured vegetation following canopy loss, or the establishment of new vegetation within canopy gaps. In such cases, remote sensing alone cannot fully explain the underlying processes responsible for the observed spectral response.

Our results are consistent with delayed mortality and delayed crown recovery observed in large conifers following severe fire injury. Trees subjected to high crown scorch or partial cambial damage may remain spectrally green for several seasons before experiencing mortality [56, 57], or alternatively show delayed flushing and partial recovery after an initial decline in canopy condition [58]. These time lags between physiological damage and observable spectral response help explain discrepancies between field observations and satellite-derived classifications in post-fire environments.

In contrast, the classification of living trees without disturbance appears highly reliable, with fewer than 5% of trees classified as alive being identified as dead during field assessments. These results suggest that our framework is particularly effective at distinguishing stable living crowns from disturbed crowns.

Our approach to define disturbance classes is closely linked to visually interpretable spectral responses, relying on variations in RGB and NDVI values (Fig. 11 and 12; and see Section 4.4.2). We believe that this framework can be more effectively interpreted by humans than approaches based solely on a single spectral index. Rather than reducing crown condition to one metric, our classification incorporates multiple spectral responses, including changes that are readily visible to the human eye. This facilitates the interpretation of the resulting classes and allows expert knowledge of crown condition and tree health to be directly incorporated into the class definition process.

3.4 Fire and large tree deaths

Fire was identified as the main driver of the large tree death and our conservative estimate is that around 660,000 large trees or $\sim 10\%$ of the large trees of Sierra Nevada Floristic Province dies as a consequence of fire between 2020 and 2025 (Fig. 6 and Table 2 and section 2.4). This is drastically above the rate observed outside fire perimeter which is around 1.6%. While fire was the dominant driver during the 2020–2025 period, other disturbance agents such as drought and bark beetle outbreaks also contributed to mortality and, in some cases, acted as competing or co-dominant drivers of large-tree loss in 2020, prior to the major 2020 fire season (Supplementary Fig. 6 in [28]). Together with the strong spatial correspondence between disturbance classes and fire occurrence shown in Fig. 6, our findings support the conclusion that wildfire has been the dominant recent disturbance affecting large trees across the Sierra Nevada.

Intense fire seasons took place in 2020 and 2021 in the Sierra Nevada and were associated with severe drought, elevated temperatures, and exceptionally dry fuels, which promoted the occurrence and spread of large (>10,000 ha), high-severity wildfires [20–23, 25]. Fire in 2020 and 2021 accounted for 86.75% of the total fire perimeter over 2020–2025, which covered 9977.758 km^2. The largest fires were also in 2020 and 2021, Fig 6b. Mortality was concentrated during the extreme 2020–2021 fire seasons, as around 88% of all observations in burned areas occurred during those years.

Patterns of disturbance reveal strong spatial heterogeneity in fire severity (Fig. 7). For example, within the North Complex Fire, some patches contain nearly all trees classified as dead, whereas other pockets of surviving trees embedded within a gradient of increasing severity (Fig. 7). Similar patterns have been described as fire refugia, where unburned or lightly burned patches persist within fire perimeters and provide important refuges for surviving trees and post-fire ecosystem recovery [59]. We also observe extensive areas where all trees remain alive, which may correspond either to unburned areas within the CAL FIRE perimeter or to locations affected by very low-intensity fire with little to no detectable impact on trees (i.e., trees classified alive and breakpoints with change in NDVI), Fig. 7. The observed mosaic of dead and surviving trees is consistent with recent studies showing that wildfire effects in Sierra Nevada forests are inherently heterogeneous [24, 59, 60]. However, the consequences of this spatial heterogeneity for the survival, recovery, and long-term dynamics of large trees remain poorly understood and the regional-scale dataset produced here could provide a baseline for future investigations.

Inside the sequoia groves affected by the 2020 and 2021 fires, our remote sensing estimates indicate that around 17.0% of large trees died due to fire. This compares to the estimate of 17.6% obtained from field assessments of large giant sequoias killed by wildfire since 1984, of which $\sim$13.9% occurred during the 2020 and 2021 fire seasons (Fig. 5 in [9]). The slightly higher mortality estimated here suggests that large giant sequoias may have been less affected than the other common large tree species found within the groves. These lower mortality rates could be related to species-specific traits, particularly the thick, fire-resistant bark and elevated crowns of giant sequoias [61]. In the unburned areas of the groves, mortality was also lower than in the Sierra Nevada as a whole, which may reflect the dominance of long-lived and fire-resistant giant sequoias.This pattern is also consistent with recent evidence showing that prescribed

burns substantially reduced giant sequoia mortality during the 2020 Castle and 2021 KNP Complex fires, with burns conducted within the previous 10 years reducing mortality odds by 77% and preventing an estimated 1,888 giant sequoia deaths [26].

In this work, we developed a framework to map the health status of individual large trees across the Sierra Nevada using Sentinel-2 time series. The framework consists of two components: individual tree delineation from canopy height models using a U-Net-ID model, and disturbance analysis based on BFAST applied to Sentinel-2 time series extracted for each large tree crown over the period 2020-2025. The combination of breakpoint detection and post-breakpoint spectral trajectories allows disturbance intensity to be characterized with a good accuracy at the scale of individual crowns while remaining applicable over large regions. The resulting maps provide a detailed view of tree mortality, disturbance, and recovery following recent wildfire events, and identify fire as the dominant driver of mortality, responsible for the death of approximately 10% of the Sierra Nevada's large trees during the period 2020-2025. The next steps are to (i) extend individual tree mapping to all trees across the United States, (ii) expand the framework to monitor changes in endangered redwood forests and other large coniferous forests of California, (iii) investigate the integration of additional structural and spectral information to improve disturbance characterization, and (iv) assess whether similar time-series approaches can be used to monitor tree condition, mortality, and recovery globally using Sentinel-2 observations.

4 Materials and Methods

4.1 Study site

The study covered the Sierra Nevada Floristic Province (Fig. 8), located in eastern California between 34.7^o and 40.5^o N latitude [62]. The region covers a total area of 65,582 km^2, with elevations ranging from 30 to 4,421 m [63, 64]. Its highest point is Mount Whitney, the tallest mountain peak in the contiguous United States. Precipitation varies substantially across the region, ranging from $\sim$ 170 mm to 2,390 mm annually [65, 66]. Temperature also exhibits considerable spatial variability, with annual mean minimum and maximum temperatures ranging from -8.8 to 12.4^{o}C and from 2.6 to 25.8^{o}C, respectively [65, 66].

4.2 Individual tree delineation

To map all large trees in the Sierra Nevada with crown areas exceeding 100 m^2, we developed an individual tree delineation algorithm. In previous work, we produced a canopy height model (CHM) for the entire state of California, including the Sierra Nevada [32]. Individual tree crowns are visible and identifiable within this CHM (see Fig. 7 and Supplementary Figures in [32]). To delineate trees from the CHM, we developed a novel approach in which crown delineation is treated as a blob segmentation problem. First, we generated realistic synthetic CHMs from simulated tree crown polygons. Using this large synthetic dataset, we trained a U-Net-ID deep learning model to identify individual polygon instances directly from the CHM. The model, trained entirely on synthetic data, was then applied to real CHM data in a transfer learning

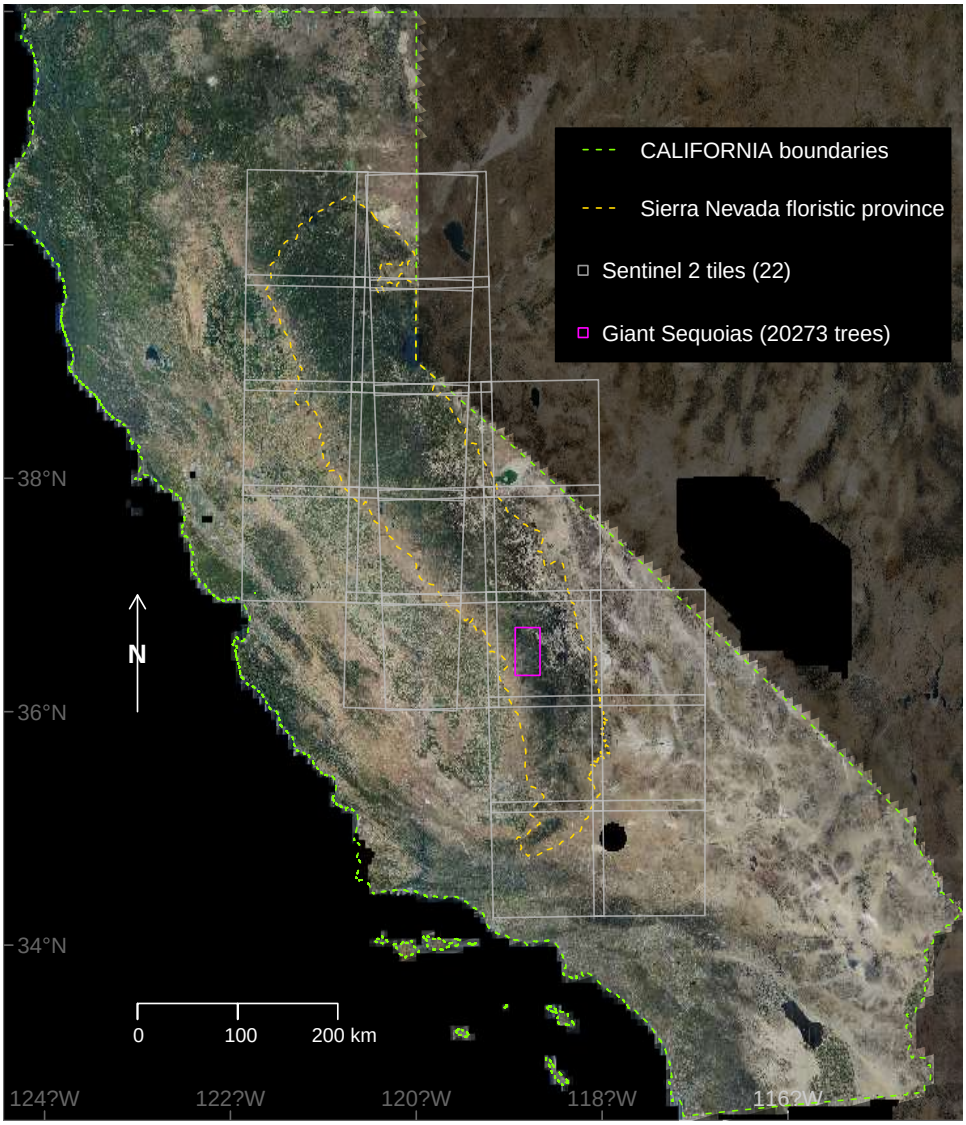

Fig. 8 Mosaic of NAIP aerial imagery acquired over California (USA) during the 2020 campaign, with the Sierra Nevada Floristic Province outlined in dashed yellow (∼65,582 km^2). The map also shows the extent of the 22 Sentinel-2 MSI tiles used for time-series analysis (grey) and the extent of the 20,273 Sequoia polygons used for individual tree crown model validation (magenta).

framework. The main advantage of this approach is that it enables simulation of a wide range of tree sizes, heights, densities, and spatial configurations, while also allowing extensive data augmentation. As a result, the model can learn to detect and delineate tree crowns without requiring any manually labeled training data or additional LiDAR processing.

4.2.1 Generation of synthetic canopy height models

We generate a fully synthetic canopy height model (CHM) by simulating 3D individual tree crowns and merging them into LiDAR–like canopy height model. The goal is to obtain realistic CHM with fully controlled samples, with known tree location, size, and geometry. Simulations are done in R using `terra`, `sf`, and `tensorflow`. Ground is flat (0 m), so CHM directly gives tree height and the resolution is around 0.6 m resolution to be close to the real resolution of our CHM of California [32]. As our CHM for California, simulated height is scaled by $\times 2.5$ to be stored in integer 8 bits while keeping some sub-meter definition.

Each simulated forest consist in a squared grid of 192×192 pixels with a central 160× 160 pixels tree placement window. Final CHM and labels are cropped to 128 × 128 pixels. Trees are randomly placed in the 160 pixels window.

To avoid dense clustering of stems and excessive crown overlap in the CHM, we defined a minimum–distance constraint between trees. A candidate tree i with crown radius R_i is accepted only if the Euclidean distance to every previously accepted tree j with radius R_j satisfies

$$\|X_i - X_j\| > (R_i + R_j) f_{\min} \tag{1}$$

where $X_i = (x_i, y_i)$, $X_j = (x_j, y_j)$ and $f_{\min} = 1/3$ is a constant distance factor. Candidate trees that do not respect this condition are rejected and resampled until a valid position is found or the defined maximum number of attempts is reached. Distance computations are carried out using tensors for improved efficiency.

The number of trees and their functional types is defined by scenarios of forest with different functional compositions and densities, Table 5. For example, large-tree forests mix conifers and deciduous with some giant Sequoia trees. Sequoia scenarios are dominated by giant Sequoia trees (20–50%). The number of trees varies by scenario (e.g., 15 to 40 for Sequoia scenario) to avoid saturation while keeping a dense canopy. Scenarios generated using an automated agent were used to simulate the different forests observed in the Sierra Nevada, producing mixed or homogeneous stands from shrubs to giant Sequoias forest.

Table 5 Example of giant sequoia forest scenarios. For each scenario, the number of trees per simulation N is drawn uniformly from the integer range in the second column. The remaining columns are sampling weights for functional types (row sums to 1).

Scenario (short name)	N range	con.	t.d.	l.c.d.	m.d.	s.c.	s.d.	sh.r.	sh.s.	g.d.	g.c.	g.r.
`dense_redwood`	20–30	.05	.05	.05	.05	.05	.05	.05	.05	.10	.10	.40
`redwood_conif_mix`	25–35	.20	.05	.05	.05	.10	.05	.05	.05	.05	.15	.20
`old_growth_redwood`	20–30	.10	.10	.10	.10	.05	.05	.05	.05	.10	.10	.20
`redwood_giant_mix`	25–35	.10	.10	.10	.10	.05	.05	.05	.05	.15	.15	.15
`pure_redwood`	15–25	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.50
`redwood_tall`	25–35	.10	.20	.10	.10	.05	.05	.05	.05	.10	.05	.15
`redwood_conif_dom`	25–35	.25	.05	.05	.05	.10	.05	.05	.05	.05	.15	.20
`mixed_redwood`	30–40	.12	.12	.12	.12	.08	.08	.08	.08	.10	.10	.10

Trees are defined by functional types with ranges of height and crown radius, Table 6. This includes conifers, deciduous, shrubs, and giant conifers and giant Sequoia types. For example, for Sequoia tree, ranges of height (H) and radius (R) are given by :

$$H \sim \mathcal{U}(80, 250), \quad R \sim \mathcal{U}(13, 27) \tag{2}$$

For each simulated tree we sample H and R from its type range.

Table 6 Tree functional types and their height and crown parameters used in the CHM simulation. Height values in the simulator are stored in scaled units; physical height (m) is obtained by dividing by 2.5. Crown radius is given in pixels in the simulator; crown radius (m), where one pixel corresponds to 0.6 m on the ground. The column H/R corresponds to the `height_radius_ratio` field in the same list.

Type (code name)	Height (scaled)	Height (m)	Crown radius (px)	Crown radius (m)	H/R
coniferous	50–100	20.0–40.0	4–6	2.4–3.6	12
tall_deciduous	60–100	24.0–40.0	3–5	1.8–3.0	15
large_crown_deciduous	35–80	14.0–32.0	6–8	3.6–4.8	6
medium_deciduous	40–70	16.0–28.0	4–6	2.4–3.6	9
sapling_coniferous	10–25	4.0–10.0	1–2	0.6–1.2	10
sapling_deciduous	8–20	3.2–8.0	1–3	0.6–1.8	6
shrub_rounded	7–15	2.8–6.0	2–4	1.2–2.4	3
shrub_spreading	7–12	2.8–4.8	3–5	1.8–3.0	2
giant_deciduous	80–120	32.0–48.0	10–15	6.0–9.0	7
giant_coniferous	90–130	36.0–52.0	8–12	4.8–7.2	10
gigantic_redwood	150–250	60.0–100.0	8–20	4.8–12.0	9

Then tree attributes are converted into rasterized canopy surfaces using shape templates.

For conifers and giant trees, narrow forms, conical crowns are used. For a tree with total height H and crown radius R, the idealized crown surface as a function of horizontal distance r from the stem is defined as

$$z(r) = H\left(1 - \frac{r}{R}\right)^s,$$

where s is a cone–shape parameter. Larger values of s produce more pointed crowns, whereas values close to one approximate straight–sided cones.

For broadleaf trees and shrubs, we use rounded crowns. In this case, the canopy surface is defined as

$$z(r) = H \cdot f\left(\frac{r}{R}; q\right),$$

where f is a smooth decreasing function, such as a cosine–based profile, that is equal to one at $r = 0$ and zero at $r = R$, and q is a roundness parameter that controls whether the crowns are relatively flat or dome–shaped.

To mimic a LiDAR CHM, only the highest point per pixel is kept and the id of the tree that has this highest point updated accordingly in a separate tree ID raster. Final outputs are cropped to 128 × 128 pixels. We further randomly add a blur to

the simulated canopy height as a data augmentation technique, rescale the maximum height to the maximum initial height and remove all height below 1.60 m.

Finally we saved the simulated canopy height model without data augmentation, with random blur augmentation, a tree-ID map in raster which associates each height in the canopy height model to its original tree and the individual polygons of the trees.

Our generated dataset consists of 59,134 synthetic CHM images of 128 × 128 pixels, containing a total of 3,356,877 simulated trees.

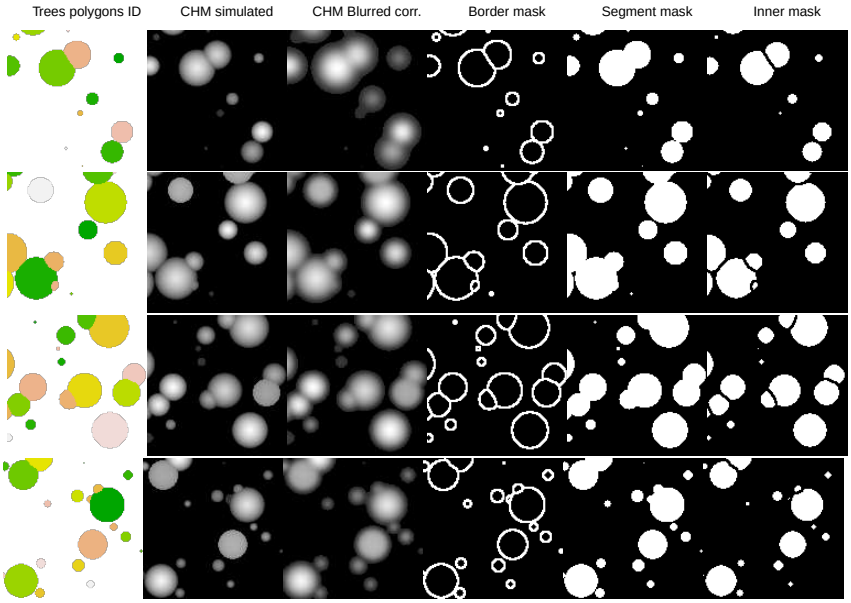

Fig. 9 Example of simulated individual tree polygons and the associated simulated CHM, simulated blurred CHM and masks in raster format.

4.3 Neural Network Architecture

The individual tree segmentation from the Sierra Nevada NAIP-derived CHM images in California was performed using the U-Net-Id model, Fig. 10 [32, 33]. The U-Net-Id model is designed to convert an instance segmentation problem into a semantic segmentation problem of binary masks, which is well resolved using U-Net-like architectures [40].

The inputs of the U-Net-ID model are synthetic canopy height model images of 128 × 128 pixels at a spatial resolution of 0.6 m, Fig. 10. The image is processed through three independent paths, each beginning with a separate U-Net model [33]. The three U-Nets predict the binary masks corresponding to the object border, the

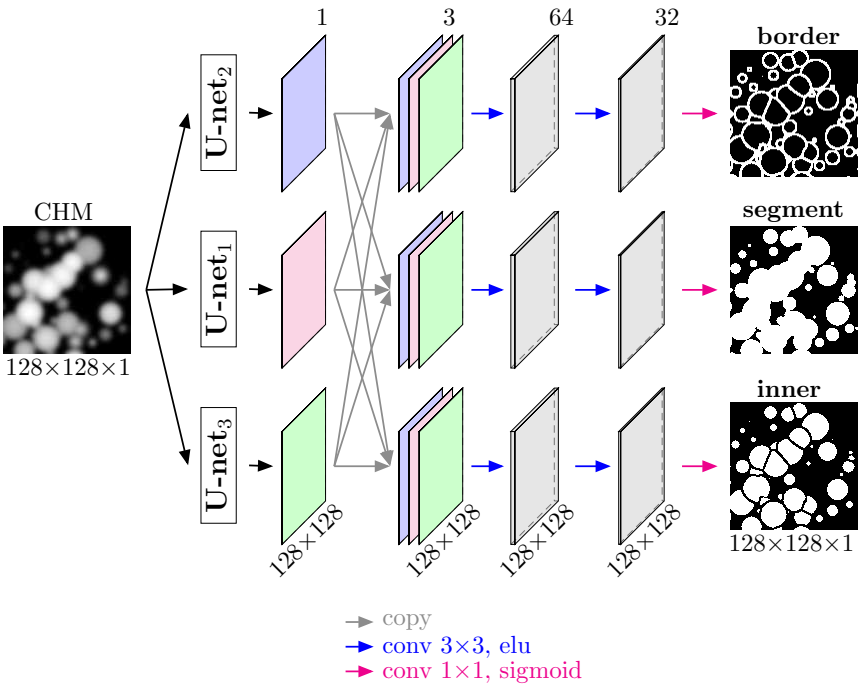

Fig. 10 U-Net-Id model architecture used for individual tree segmentation from Canopy height model, adapted from Wagner et al. [33]. The three U-net units used in our model have the same layers as in Wagner et al. [33] only that the input and output of the U-net units here are 128 × 128. The outputs of the U-net units are represented with three different colors: magenta (U-net$_1$), blue (U-net$_2$), and green (U-net$_3$). The number of channels is indicated above the cuboids, and the vertical numbers indicate the row and column size in pixels. The operations (skip connections and convolutions) performed in each layer and their sizes are indicated by the colored arrows.

object segment, and the object inner segment. These training masks are generated directly from the simulated individual tree polygons that were used to produce the simulated CHM image (see section 4.2.1). The object segment corresponds to the original polygon, the border is defined as the area between a buffer of 0.6 m and -0.6 m around the polygon boundary, and the inner segment is obtained by applying a negative buffer of 0.6 m to the original polygon. The resulting masks therefore overlap along the object boundary, creating transition zones equivalent to $\sim$ 1 to 3 pixels in width. The inner segment is individual to each tree. The three 128 × 128 activation maps produced by the U-Nets are then concatenated. This is followed by two convolutional layers with 64 and 32 filters, respectively. Finally, separate predictions for the segment, border, and inner segment are produced using a final convolutional layer with a sigmoid activation function.

The instance individualization is performed in post-processing by extracting the inner segments, which are unique and do not touch each other, and expanding them with a 1-pixel (0.6 m) buffer. The resulting objects are then combined with predicted segments that do not intersect any inner segment. These isolated segments correspond

to the smallest trees, with crown sizes of only 1–4 pixels, which do not have an inner segment. Finally, a unique ID is added.

4.3.1 Training and testing

The training sample comprised a total of 59,134 128 × 128 pixels one-band CHM image patch their associated 3 bands mask patchs. 90% (53,372) of the samples were used for training and 10% (5,762) for internal model testing. Before being inputted into the U-Net model, image patches underwent random vertical and horizontal flips as data augmentation.

The network was trained for 2,048 epochs with a batch size of 128 images. The loss function was designed as a sum of two terms: binary cross-entropy and Dice coefficient-related loss of the three predicted masks [67–69]. During network training, we used standard stochastic gradient descent (SGD) optimization with the RMSprop optimizer (learning rate of 0.0001). The SGD procedure minimizes the loss function using random mini-batches of data, enabling faster convergence and improved generalization [70]. The model with the best validation loss (i.e., Dice coefficient-related loss of 0.02324226 and a Dice coefficient of 0.97362) was kept for prediction. The training of the model took around 12 hours using an Nvidia RTX4090 Graphics Processing Unit (GPU) with 24 GB of memory.

4.3.2 Prediction

For prediction, the 1,885 NAIP-based CHM tiles covering the Sierra Nevada Floristic Province were resized by adding columns and rows, resulting in images sized at 11,392 × 13,440 pixels with an aspect ratio of 1024 and a 128-pixel border. This adjustment was made to meet the input size requirements for prediction. The standardized-size NAIP tiles were subsequently divided into sub-images of 1,152 × 1,152 pixels with a 64-pixel overlap between them. Predictions were made on these 1,152 × 1,152-pixel images, and the 64-pixel border on each sub-image prediction was removed to mitigate border artifacts [40]. The resulting 1024 × 1024 images were merged and cropped to the original NAIP extent, and, finally, a 60-pixel border (36 m) was removed to mitigate any remaining border artifacts. This last step did not affect the final height map since NAIP images have an overlap of $\sim$ 400 meters between them. The computing time for predicting California tree height using the RTX4090 GPU was 4 days.

4.3.3 Instance polygonization and tree crown statistics

Polygonisation of individual tree crowns was performed in post-processing using the `terra` and `sf` R packages [71–73]. Individual crowns were defined as either (i) inner segments buffered by 0.6 m to reconstruct full crowns size, or (ii) isolated predicted segments that do not intersect any inner segment, i.e. very small trees.

Due to border overlap between NAIP tiles, polygon conflicts along image borders were resolved by retaining for each tile only polygons intersecting the left and bottom edges of the NAIP original image tiling grid, ensuring that each tree appears in only one tile and avoiding double counting.

For each individual crown, a unique identifier was assigned. Centroids were extracted for each polygon, crown area computed, and crown radius was approximated assuming circular crowns ($radius = \sqrt{area/\pi}$). Height statistics were then computed from the extracted CHM raster values for each polygon, including mean, median, minimum, and maximum height values.

4.3.4 Instance validation

We used a giant sequoia reference dataset comprising 20,273 trees, generated from field-validated stem maps and crown polygons delineated from airborne LiDAR data collected between 2015 and 2017. Crown delineation was performed using an unsupervised watershed segmentation algorithm, and only trees with a one-to-one correspondence between stem locations and delineated crown polygons were retained (see Dixon et al. [26] for a full description of the dataset and methodology). Because the predicted dataset contained substantially more tree instances, including both giant sequoias and other tree species, we matched each observed crown polygon to the predicted intersecting polygon with the closest crown area, while ensuring that each predicted polygon could be assigned only once. The agreement between observed and predicted crown polygons was evaluated using the root mean square error (RMSE), Pearson correlation coefficient (r), and bias, calculated as predicted minus observed crown area. The spatial displacement between observed and predicted crown centroids was also quantified. Finally, we assessed the spatial overlap between matched crown polygons using the intersection over union (IoU), reporting both the mean and median IoU across matched tree instances, with and without accounting for centroid displacement.

4.4 Sentinel-2 data cubes preparation and analysis

We prepared a Sentinel–2 surface reflectance data cube for the Sierra Nevada region using the `sits` package [74]. As a first step, we delineated the region of interest using the California Floristic Province shapefile and selected only the polygons corresponding to the Sierra Nevada subsection. This polygon mask was reprojected to WGS84 and used to intersect the global Sentinel–2 tiling grid, which yielded 22 Sentinel–2 tiles overlapping the Sierra Nevada. The tile identifiers were stored in a list of tile names, including the tile that contains our sequoia field plots (11SLA). For each tile in this list, we used `sits_cube` to query the Sentinel–2 L2A collection hosted on Amazon Web Services. We requested the blue (B02), green (B03), red (B04), near–infrared (B08), shortwave infrared (B11 and B12) and cloud mask (CLOUD) bands, and restricted the temporal coverage to the period from December 2017 to February 2025. To reduce the influence of clouds at the scene level, we exploited the `file_info` metadata returned by `sits_cube` and retained only images whose reported cloud cover was less than or equal to five percent over the tile, discarding more contaminated acquisitions.

The selected images for each tile were then copied from the cloud provider to local storage using `sits_cube_copy`, with multi–core processing activated to accelerate data transfer. We used a simple consistency check to ensure that all requested bands were present for every acquisition date, repeatedly calling `sits_cube_copy` until the local

directory contained the same number of images for each band and for the cloud mask. This procedure produced, for each Sentinel–2 tile intersecting the Sierra Nevada, a locally stored, band–complete archive of Sentinel–2 L2A scenes with low overall cloud cover, which was subsequently used as input to the regularization and time–series analysis steps described in the following sections.

4.4.1 Creation of the RGB reflectance envelope for alive and dead crowns

We extracted Sentinel–2 surface reflectance time series for individual sequoia crowns from field data by first intersecting field–validated sequoia polygons (`true_in.shp`) with the predicted sequoia crowns (`pred_in.shp`) and retaining only those predicted objects that overlapped a field crown and had a similar area. The centroids of the validated sequoia polygons were then used as point samples for extracting the multi–band time series from the Sentinel–2 cube (bands B02, B03, B04, B08 and B11). This procedure ensured that all subsequent reflectance values correspond to locations known to contain real sequoia trees.

The resulting time series still contained spurious reflectance values caused by clouds, shadows and atmospheric artefacts, which appear in the Sentinel–2 surface reflectance product as exact zeros or very low values.

To obtain a clean representation of the sequoia spectral signal, we applied a two–stage filtering to ensure that the Sentinel–2 reflectance time series correspond to realistic sequoia values. First, we cleaned the temporal trajectories of each spectral band (B02, B03, B04, B08 and B11) by removing obvious artefacts. Exact zero values in the surface reflectance were interpreted as invalid observations and replaced by the last previously observed non–zero value in the same band, i.e., a last–observation–carried–forward scheme restricted to zeros. We then treated very low reflectance values as unreliable (approximately corresponding to digital numbers of 1–10 in the Sentinel–2 product) and similarly replaced any value below a small threshold by the most recent valid reflectance in the same time series. This two–step temporal filtering was applied independently to all bands and to all sequoia crowns.

After the temporal filtering described above, we further restricted the time series to reflectance combinations that are consistent with the observed spectral signature of real field sequoia crowns. To do this, we first constructed a three–dimensional color mesh in RGB space using only manually inspected sequoia observations that were visually confirmed to be free of clouds, shadows and atmospheric artefacts, Fig. 11a. For each such reference observation we converted the filtered Sentinel–2 red, green and blue reflectances (bands B04, B03 and B02) to pseudo–8-bit digital numbers and used these triplets as vertices in a three–dimensional point cloud,

$$\mathbf{c}_k = \big(R_k, G_k, B_k\big),$$

where R_k, G_k and B_k denote the scaled reflectances of the red, green and blue bands for reference sample k. Note that some trees burned during the study period, so crown colors can range from green for living trees to brown for dead crowns. The imagery

also includes variations caused by illumination and atmospheric conditions, as well as occasional snow appearing in clear blue tones.

From this point cloud, we derived a closed 3D mesh that approximates the region of RGB space occupied by uncontaminated sequoia crowns Fig. 11b. In practice, this mesh can be interpreted as a three–dimensional envelope (for example, a convex hull or a triangulated surface) that encompasses all reference points $\mathbf{c}_k$ and defines the admissible color domain for sequoia reflectance. For each time–series observation t and each crown, we computed its RGB triplet $\mathbf{c}(t) = \big(R(t), G(t), B(t)\big)$ from the filtered bands and tested whether this point lies inside the sequoia mesh. Observations for which $\mathbf{c}(t)$ fell outside the mesh were discarded, whereas only those time steps whose RGB values lay inside the 3D sequoia envelope were retained for subsequent analysis. This final filtering step ensures that the retained reflectance values are not only temporally consistent, but also spectrally consistent with the empirical distribution of real crown reflectance in three–dimensional color space.

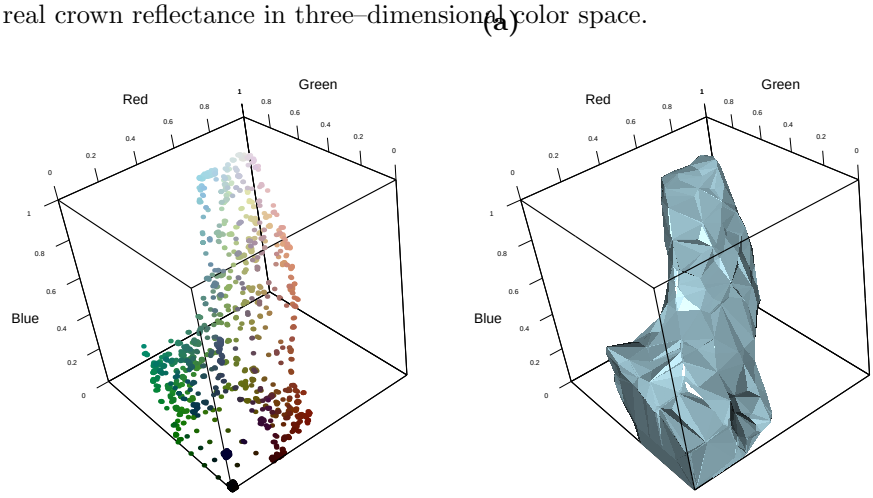

Fig. 11 Sample of points on the surface defining the RGB color domain from sequoia field crowns extracted from Sentinel-2 imagery (a), and a 3D mesh representing the color-domain surface of sequoia crowns used to filter the Sentinel-2 time series, i.e., all colors outside the mesh are discarded (b).

4.4.2 Time series analysis of Sierra Nevada's large trees

Spectral time series were extracted for each crowns polygon using the blue (B02), green (B03), red (B04), and near-infrared (B08) bands from multispectral satellite imagery. The Normalized Difference Vegetation Index (NDVI) was computed for each observation as:

$$NDVI = \frac{B08 + B04}{B08 - B04} \tag{3}$$

To reduce the influence of anomalous spectral observations, RGB values were first scaled and constrained to a valid reflectance range. The normalized RGB values were

then projected into a predefined spectral envelope representing physically plausible color combinations. Observations falling outside this envelope were excluded from further analysis.

The NDVI time series was constructed with a regular temporal sampling corresponding to the image acquisition frequency (approximately 16 days). Structural changes in the NDVI trajectory were detected using the breakpoint analysis implemented in `bfast` [41, 42]. The method decomposes the time series into seasonal and trend components using a harmonic seasonal model and identifies the most likely breakpoint in the trend component.

When a breakpoint was detected, the time series was divided into pre- and post-break segments. Mean spectral values (R, G, B) and NDVI were computed for each segment. Differences in spectral distributions between the two periods were evaluated using a permutational multivariate analysis of variance (PERMANOVA) based on Euclidean distances with 999 permutations. Homogeneity of dispersion between groups was also tested to ensure that significant differences reflected shifts in the spectral centroid rather than differences in variance.

Vegetation recovery after the disturbance was characterized by fitting a linear regression to the NDVI observations following the breakpoint. Time was expressed in days to estimate the rate of NDVI change per day. The slope of this regression represents the post-disturbance recovery rate, and its statistical significance was assessed using a t-test on the regression coefficient. Additional recovery metrics were derived, including the percentage of NDVI recovery relative to the pre-disturbance level and the estimated time required to reach full recovery based on the observed post-break trend.

Breakpoint analysis was used to classify tree crowns as alive, dead, recovering, or partially recovering. Only crowns with a statistically significant breakpoint (permutation $p<0.05$) occurring after 1 January 2020 were considered, thereby excluding earlier fluctuations unrelated to the recent fires.

To remove spurious breakpoints caused by residual noise, we evaluated post–break visible reflectance. The distribution of red and green reflectance values after the breakpoint was examined and a linear threshold in this colour space was used to separate healthy green crowns from browned or charred canopies. Crowns whose post–break colour remained in the green domain were reclassified as undisturbed, whereas crowns falling in the brown domain were retained as disturbed.

For the complete dataset, we identified seven possible trajectories of large crown color time series, illustrated in Fig. 12.

The first class corresponds to healthy living trees (sequoia or in the spectral range of sequoias) with no detected breakpoints (Fig. 12a), characterized by a mean NDVI close to 0.8 and annual seasonality. The second class corresponds to trees with a detected breakpoint but no substantial change in mean NDVI (Fig. 12b), indicating that the crown remains alive and healthy despite a short disturbance signal. The third class includes crowns with a breakpoint followed by $\geq 80\%$ NDVI recovery (Fig. 12c). These crowns experienced a disturbance but rapidly recovered to values close to their original NDVI and were therefore classified as alive and recovered. The fourth class includes crowns that were already dead or that did not exhibit a typical sequoia-like crown color

(Fig. 12d). These crowns are characterized by low NDVI values close to 0.4. This class may include already dead trees or large tree species with naturally different crown reflectance, such as large oaks located on the lower slopes of the Sierra Nevada. The fifth class includes crowns with a breakpoint followed by $\leq$ 80% NDVI recovery (Fig. 12e). These crowns experienced a disturbance but showed limited recovery, suggesting more permanent damage. The sixth class corresponds to trees that have died (Fig. 12f), characterized by a breakpoint followed by persistent brown coloration and no NDVI recovery in subsequent years. The seventh class represents crowns with gradual and permanent alteration without recovery (Fig. 12g). In these cases, the signal slowly transitions from green to brown, indicating a progressive loss of photosynthetic activity for unknown reasons, potentially related to decline, stress, or disease. At last, the class not represented in Figure 12 is a time series without enough data to fit the BFAST algorithm. All trees with crowns > 100 m^2 were classified into these 8 classes.

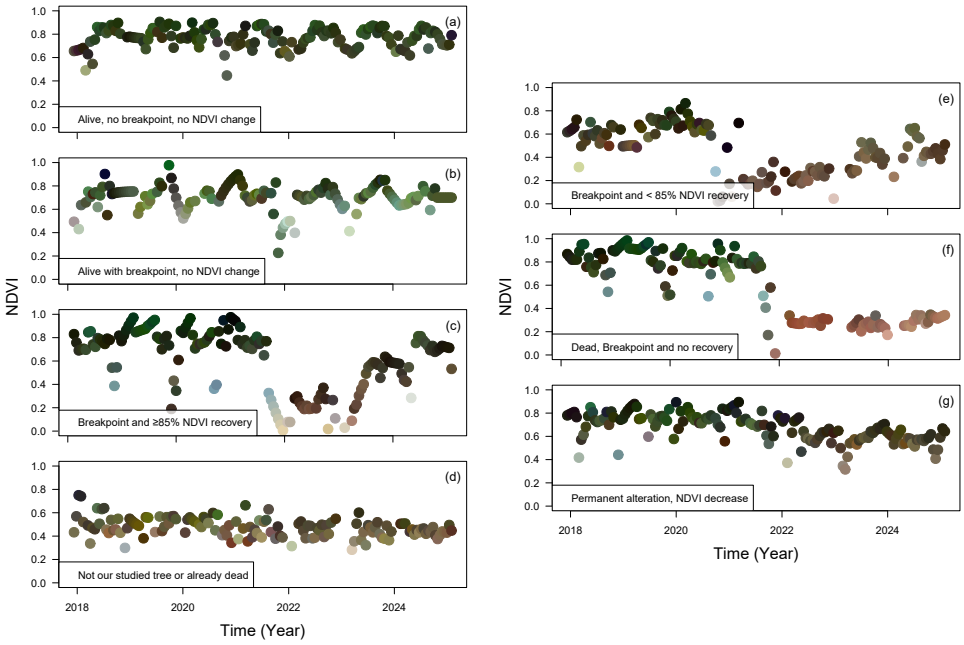

Fig. 12 Examples of the seven time series trajectories identified in the dataset and used to classify the large crowns into alive or disturbed classes.

To assess how our class interpretation corresponds to field reality of crown health condition, we compared our classification with field-based observations of giant sequoia (*Sequoiadendron giganteum*) condition, including survival status and crown condition of individual trees recorded within the footprints of the 2020 Castle Fire and the 2021 Windy and KNP Complex fires [47], https://www.sciencebase.gov/catalog/item/63d40479d34e06fef150e7b4. For each polygon in the satellite-based dataset, we performed a spatial intersection with ground-based observations of 'Last_Known_Status'.

When a polygon intersected a single observation, that status was directly assigned; when multiple observations with identical status were present, the common status was retained. In cases of conflicting observations, polygons were classified as Dead if at least one intersecting observation indicated mortality, otherwise the remaining status was retained. We kept dead in this case as this is the status that has the most influence on the change of the signal seen by the satellite.

After the validation step, the classification framework was applied to the complete large-tree dataset across the Sierra Nevada and we quantified the distribution of status classes across the region. The distribution of classes was compared with the spatial distribution of wildfire perimeters during the period. Historical wildfire perimeters were obtained from the CAL FIRE Fire and Resource Assessment Program (FRAP) Fire Perimeters database (Firep24_1), a statewide compilation of wildfire perimeters developed jointly by CAL FIRE, the U.S. Forest Service, the Bureau of Land Management, the National Park Service, and the U.S. Fish and Wildlife Service [75]. This analysis was conducted both at the scale of the entire Sierra Nevada and within the boundaries of known giant sequoia groves to assess patterns of mortality and disturbance in sequoia-dominated forests.

4.5 Environmental data

To test the association of large crown tree density trees with elevation, slope and orientation, elevation data from the Shuttle Radar Topography Mission (SRTM) were used, Fig. 13a-c [63]. Specifically, we used the 1 arc-seconds (∼30 m) spatial resolution digital elevation database Distributed by OpenTopography [64]. From this dataset, we used the variables elevation (m) and computed slope (^{o}) and orientation considering the 8 neighbor pixels (Fig. 13a-c).

The association of large tree densities with local climate was tested using the annual means of precipitation and air temperatures (Fig. 13d-f). Mean annual precipitation and air temperature over the region were computed for the period 1981–2025 using the PRISM climate dataset at 800 m spatial resolution. PRISM provides high-resolution gridded climate data derived from station observations and spatial interpolation across the United States [65, 66].

All environmental variables were resampled to a raster of 500 × 500 m spatial resolution using an average interpolation to match the resolution of the tree density dataset.

To describe the association of large trees with environmental variables, we first reclassified each environmental variable into classes and compared the proportion of large trees occurring in each class with the proportion of pixels belonging to that class across the Sierra Nevada Floristic Province [45]. Ratios greater than 1 indicate that large trees are overrepresented in a class, whereas ratios below 1 indicate underrepresentation. Ratios close to 1 indicate that large trees occur in similar proportion to the overall distribution of that class.

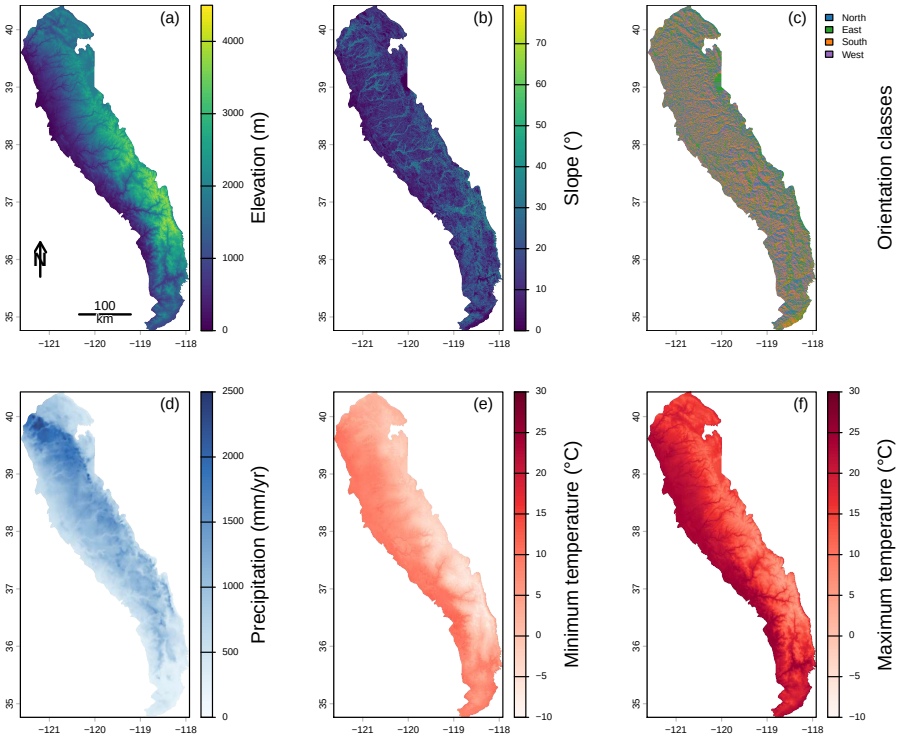

Fig. 13 Environmental and climatic variables used in the study to analyse large trees density distribution (a) elevation (m), (b) slope (^{o}), (c) Slope orientation in classes (North, East, South and West), (d) mean annual precipitation (mm.yr^{-1}), (e) annual mean of minimum temperatures (^{o}C), and (f) annual mean of maximal temperatures (^{o}C).

Acknowledgements

The authors wish to thank the Grantham Foundation and High Tide Foundation for their generous gift to UCLA and support to CTrees.org. Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA).

Declarations

- Conflict of interest/Competing interestsThe authors declare no competing financial interests.
- Data availability The Sentinel-2 MSI and NAIP data that support the findings of this study are publicly available from their sources. Processed data, products and codes produced in this study are available from the corresponding author upon request.
- Author contribution Conceptualization, F.H.W., M.C.M.H., R.D., G.C., S.S. and S.C.H.; methodology, F.H.W., D.J.D., C.W.W., M.C.M.H., F.S., R.D., G.C., C.M.

and S.S.; software, F.H.W., R.D., G.C., F.S. and M.C.M.H.; validation, F.H.W. and D.J.D.; formal analysis, F.H.W.; investigation, F.H.W., R.D., G.C., F.S., M.C.M.H. and S.S.; resources, S.S.; data curation, F.H.W., M.C.M.H., D.J.D. and F.S.; writing—original draft preparation, F.H.W.; writing—review and editing, F.H.W., R.D., G.C., M.C.M.H., L.B.S.T., F.S., L.A., C.C., S.C.H., Z.L., C.M. A.M., Y.Y., D.J.D., C.W.W., M.S.T., B.S., D.Z. and S.S.; visualization, F.H.W. and M.C.M.H.; supervision, S.S.; project administration, C.W.W., S.C.H. and S.S.; funding acquisition, C.W.W., S.C.H. and S.S. All authors reviewed and approved the final manuscript.

==========

References

[1] Raven, P. H. & Axelrod, D. I. *Origin and relationships of the California flora* Vol. 72 (Univ of California Press, 1978).

[2] Baldwin, B. G., Goldman, D. H. & Vorobik, L. A. *The Jepson manual: vascular plants of California* (Univ of California Press, 2012).

[3] Baldwin, B. G. Origins of plant diversity in the california floristic province. *Annual Review of Ecology, Evolution, and Systematics* **45**, 347–369 (2014).

[4] Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853 (2000).

[5] Van Wagtendonk, J. W., Fites-Kaufman, J. A., Safford, H. D., North, M. P. & Collins, B. Sierra nevada bioregion. *Fire in California's ecosystems* 249–279 (2018).

[6] Barbour, M. G. & Minnich, R. A. California upland forests. *North American terrestrial vegetation* 132–164 (1988).

[7] Griffin, J. R. & Critchfield, W. B. The distribution of forest trees in california. *USDA FOREST SERVICE RESEARCH PAPER PSW- 82 /1972* (1976).

[8] Stephenson, N. L. *Ecology and Management of Giant Sequoia Groves* (Centers for Water and Wildland Resources, University of California, Davis, CA, 1996).

[9] Shive, K. L. *et al.* The state of the giant sequoias: losses, risks, and opportunities. *Fire Ecology* **22**, 30 (2026).

[10] Winsemius, S. *et al.* Improved aboveground biomass estimation and regional assessment with aerial lidar in california's subalpine forests. *Carbon Balance and Management* **19**, 41 (2024).

[11] Sillett, S. C., Van Pelt, R., Carroll, A. L., Campbell-Spickler, J. & Antoine, M. E. Structure and dynamics of forests dominated by sequoiadendron giganteum. *Forest Ecology and Management* **448**, 218–239 (2019).

[12] Enquist, B. J., Abraham, A. J., Harfoot, M. B., Malhi, Y. & Doughty, C. E. The megabiota are disproportionately important for biosphere functioning. *Nature Communications* **11**, 699 (2020).

[13] Lutz, J. A. *et al.* Global importance of large-diameter trees. *Global Ecology and Biogeography* **27**, 849–864 (2018).

[14] Stephenson, N. L. *et al.* Rate of tree carbon accumulation increases continuously with tree size. *Nature* **507**, 90–93 (2014).

[15] McIntyre, P. J. *et al.* Twentieth-century shifts in forest structure in california: Denser forests, smaller trees, and increased dominance of oaks. *Proceedings of the National Academy of Sciences* **112**, 1458–1463 (2015).

[16] Domke, G. M., Oswalt, S. N., Walters, B. F. & Morin, R. S. Tree planting has the potential to increase carbon sequestration capacity of forests in the united states. *Proceedings of the national academy of sciences* **117**, 24649–24651 (2020).

[17] Jones, A. G. *et al.* Glaciers in california's sierra nevada are likely disappearing for the first time in the holocene. *Science Advances* **11**, eadx9442 (2025).

[18] Wang, J. A., Randerson, J. T., Goulden, M. L., Knight, C. A. & Battles, J. J. Losses of tree cover in california driven by increasing fire disturbance and climate stress. *AGU Advances* **3**, e2021AV000654 (2022).

[19] Williams, C. A., Gu, H., MacLean, R., Masek, J. G. & Collatz, G. J. Disturbance and the carbon balance of us forests: A quantitative review of impacts from harvests, fires, insects, and droughts. *Global and Planetary Change* **143**, 66–80 (2016).

[20] Safford, H. D., Paulson, A. K., Steel, Z. L., Young, D. J. & Wayman, R. B. The 2020 california fire season: A year like no other, a return to the past or a harbinger of the future? *Global Ecology and Biogeography* **31**, 2005–2025 (2022).

[21] Keeley, J. E. & Syphard, A. D. Large california wildfires: 2020 fires in historical context. *Fire Ecology* **17**, 22 (2021).

[22] Kennedy, M. C., Bart, R. R., Tague, C. L. & Choate, J. S. Does hot and dry equal more wildfire? contrasting short-and long-term climate effects on fire in the sierra nevada, ca. *Ecosphere* **12**, e03657 (2021).

[23] Williams, J., Safford, H., Enstice, N., Steel, Z. & Paulson, A. High-severity burned area and proportion exceed historic conditions in sierra nevada, california, and adjacent ranges. *Ecosphere* **14**, e4397 (2023).

[24] Dixon, D. J., Zhu, Y., Brown, C. F. & Jin, Y. Satellite detection of canopy-scale tree mortality and survival from california wildfires with spatio-temporal deep

learning. *Remote Sensing of Environment* **298**, 113842 (2023).

[25] Hung, M. J. & Williams, A. P. High-severity fire now dominant in california forests. *Proceedings of the National Academy of Sciences* **123**, e2532829123 (2026).

[26] Dixon, D. J. *et al.* Previous prescribed burns saved thousands of ancient sequoias during historically unprecedented wildfires. *Nature Communications* (2026).

[27] Hardlund, L. J. *et al.* How big is big enough? exploring drivers of fire-induced giant sequoia mortality from the individual to the stand level. *Fire Ecology* (2026).

[28] Cheng, Y. *et al.* Scattered tree death contributes to substantial forest loss in california. *Nature communications* **15**, 641 (2024).

[29] Favrichon, S. *et al.* Monitoring changes of forest height in california. *Frontiers in Remote Sensing* **Volume 5 - 2024** (2025). URL https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2024.1459524.

[30] Carter, G. *et al.* Detection of forest disturbance across california using deep-learning on planetscope imagery. *Frontiers in Remote Sensing* **5**, 1409400 (2024).

[31] USDA. National Agricultural Imagery Program (NAIP). Tech. Rep., United States Department of Agriculture (2020).

[32] Wagner, F. H. *et al.* Sub-meter tree height mapping of california using aerial images and lidar-informed u-net model. *Remote Sensing of Environment* **305**, 114099 (2024).

[33] Wagner, F. H. *et al.* U-net-id, an instance segmentation model for building extraction from satellite images—case study in the joanópolis city, brazil. *Remote Sensing* **12** (2020). URL https://www.mdpi.com/2072-4292/12/10/1544.

[34] He, K. S. *et al.* Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation* **1**, 4–18 (2015). URL http://dx.doi.org/10.1002/rse2.7.

[35] Lin, T.-Y., Goyal, P., Girshick, R., He, K. & Dollár, P. Focal loss for dense object detection (2017).

[36] G. Braga, J. R. *et al.* Tree crown delineation algorithm based on a convolutional neural network. *Remote Sensing* **12** (2020). URL https://www.mdpi.com/2072-4292/12/8/1288.

[37] Weinstein, B. G. *et al.* Deepforest: A python package for rgb deep learning tree crown delineation. *Methods in Ecology and Evolution* **11**, 1743–1751 (2020).

[38] Weinstein, B. G., Marconi, S., Bohlman, S., Zare, A. & White, E. Individual tree-crown detection in rgb imagery using semi-supervised deep learning neural networks. *Remote Sensing* **11**, 1309 (2019).

[39] Bai, T., Li, D., Sun, K., Chen, Y. & Li, W. Cloud Detection for High-Resolution Satellite Imagery Using Machine Learning and Multi-Feature Fusion. *Remote Sensing* **8**, 715 (2016).

[40] Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. *CoRR* **abs/1505.04597** (2015). URL http://arxiv.org/abs/1505.04597.

[41] Verbesselt, J., Hyndman, R., Newnham, G. & Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment* **114**, 106–115 (2010).

[42] Verbesselt, J., Hyndman, R., Zeileis, A. & Culvenor, D. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sensing of Environment* **114**, 2970–2980 (2010).

[43] Chen, N., Tsendbazar, N.-E., Hamunyela, E., Verbesselt, J. & Herold, M. Sub-annual tropical forest disturbance monitoring using harmonized landsat and sentinel-2 data. *International Journal of Applied Earth Observation and Geoinformation* **102**, 102386 (2021).

[44] Schiller, C. & Fassnacht, F. E. Comparative study of near real-time monitoring algorithms for early detection of bark beetle infestations in germany with sentinel-2. *European Journal of Remote Sensing* **59**, 2662660 (2026). URL https://doi.org/10.1080/22797254.2026.2662660.

[45] Wagner, F. H. The flowering of atlantic forest pleroma trees. *Scientific reports* **11**, 1–20 (2021).

[46] Saad, F. *et al.* Integrating optical and sar data enables crown-level maps of an emergent tree species, dipteryx panamensis. *Remote Sensing Applications: Society and Environment* 102076 (2026).

[47] Soderberg, D. N. & Das, A. J. Assessment of giant sequoia mortality and regeneration within burned groves in sequoia and kings canyon national parks (ver. 4.0, february 2026). *US Geological Survey (USGS) Data Release* 805 (2023).

[48] Ritz, A. L. *et al.* Applying a convolutional neural network (cnn) to virginia's forests: how forest type and age can impact individual tree segmentation. *International Journal of Remote Sensing* **47**, 218–245 (2026). URL https://doi.org/10.1080/01431161.2025.2598075.

[49] He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask r-cnn (2017).

[50] Kane, V. R. *et al.* Where are the large trees? a census of sierra nevada large trees to determine their frequency and spatial distribution across three large landscapes. *Forest Ecology and Management* **546**, 121351 (2023).

[51] Ali, A. & Wang, L.-Q. Big-sized trees and forest functioning: Current knowledge and future perspectives. *Ecological Indicators* **127**, 107760 (2021).

[52] Lutz, J. A., Struckman, S., Germain, S. J. & Furniss, T. J. The importance of large-diameter trees to the creation of snag and deadwood biomass. *Ecological Processes* **10**, 28 (2021).

[53] Miller, J. D. *et al.* Calibration and validation of the relative differenced normalized burn ratio (rdnbr) to three measures of fire severity in the sierra nevada and klamath mountains, california, usa. *Remote sensing of environment* **113**, 645–656 (2009).

[54] Miller, J. D. & Thode, A. E. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta normalized burn ratio (dnbr). *Remote sensing of Environment* **109**, 66–80 (2007).

[55] Parks, S. A., Holsinger, L. M., Voss, M. A., Loehman, R. A. & Robinson, N. P. Mean composite fire severity metrics computed with google earth engine offer improved accuracy and expanded mapping potential. *Remote Sensing* **10** (2018). URL https://www.mdpi.com/2072-4292/10/6/879.

[56] Jeronimo, S. M., Lutz, J. A., R. Kane, V., Larson, A. J. & Franklin, J. F. Burn weather and three-dimensional fuel structure determine post-fire tree mortality. *Landscape Ecology* **35**, 859–878 (2020).

[57] Reilly, M. J., Zuspan, A. & Yang, Z. Characterizing post-fire delayed tree mortality with remote sensing: sizing up the elephant in the room. *Fire Ecology* **19**, 64 (2023).

[58] Hanson, C. T. & North, M. P. Post-fire survival and flushing in three sierra nevada conifers with high initial crown scorch. *International Journal of Wildland Fire* **18**, 857–864 (2009).

[59] Blomdahl, E. M., Kolden, C. A., Meddens, A. J. & Lutz, J. A. The importance of small fire refugia in the central sierra nevada, california, usa. *Forest Ecology and Management* **432**, 1041–1052 (2019).

[60] Paudel, A. & Markwith, S. H. From the severity patch to the landscape: Wildfire and spatial heterogeneity in northern sierra nevada conifer forests. *Journal of Vegetation Science* **34**, e13207 (2023).

[61] Innes, R. J. Sequoiadendron giganteum, giant sequoia. *Fire Effects Information System,[Online]. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Missoula Fire Sciences Laboratory (Producer). Available: https://research. fs. usda. gov/feis/species-reviews/seqgig* (2026).

[62] Burge, D. O. *et al.* Data from: Plant diversity and endemism in the california floristic province. *V1* (2021). URL https://borealisdata.ca/file.xhtml?fileId=153128&version=1.0.

[63] Farr, T. G. *et al.* The shuttle radar topography mission. *Reviews of geophysics* **45** (2007).

[64] NASA. Shuttle radar topography mission (srtm)(2013). shuttle radar topography mission (srtm) global. distributed by opentopography. https://doi.org/10.5069/g9445jdf. accessed 2026-05-20. Tech. Rep., NASA (2026).

[65] Daly, C. *et al.* Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *International Journal of Climatology: a Journal of the Royal Meteorological Society* **28**, 2031–2064 (2008).

[66] PRISM Climate Group. Prism climate data. https://prism.oregonstate.edu. Accessed: 2025-12-16.

[67] Dice, L. R. Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302 (1945).

[68] Allaire, J. & Chollet, F. *keras: R Interface to 'Keras'* (2019). URL https://keras.rstudio.com. R package version 2.2.4.1.9001.

[69] Chollet, F. *et al.* Keras. https://keras.io (2015).

[70] Chollet, F. & Allaire, J. *Deep Learning with R* (Manning Publications Co., 2018).

[71] Hijmans, R. J. *terra: Spatial Data Analysis* (2026). URL https://CRAN.R-project.org/package=terra. R package version 1.8-93.

[72] Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* **10**, 439–446 (2018). URL https://doi.org/10.32614/RJ-2018-009.

[73] Pebesma, E. & Bivand, R. *Spatial Data Science: With applications in R* (Chapman and Hall/CRC, 2023). URL https://r-spatial.org/book/.

[74] Simoes, R. *et al.* Satellite image time series analysis for big earth observation data. *Remote Sensing* **13**, 2428 (2021).

[75] California Department of Forestry and Fire Protection (CAL FIRE). Fire perimeters (firep24_1 geodatabase) (2025). URL https://www.fire.ca.gov/what-we-do/fire-resource-assessment-program/fire-perimeters. Version Firep24_1; statewide wildfire perimeter database.