

Statewide maps of tree mortality 2016-2018

Prepared on behalf of Pyregence Working Group 2
Metadata and model validation report

Christopher B. Anderson & David C. Marvin
Salo Sciences, Inc., San Francisco CA
October 21, 2022



Figure 1: The USFS estimated that over 120 million trees died across California due to pest and drought over the past decade, with severe loss occurring from 2016-2018. The spatial distribution of this process shifted between years but was concentrated in the central and southern Sierra Nevada ecoregion, which is mapped in detail with this new dataset.

Introduction

California's forests have experienced severe stress from drought, heat, fires, and pest outbreaks over the past decade. In this new era of megafire, tree mortality is expected to exacerbate wildfire risk and further increase the economic and human health costs of catastrophic fire. However, there are few datasets and wildfire modeling tools that can quantify and map these effects, making it difficult to understand the rapidly shifting landscapes of wildfire hazard and exposure, and limiting the state's ability to manage and mitigate mortality-driven wildfire risks.

The Pyrengence Tree Mortality Working Group (WG2) is working to provide quantitative and qualitative evaluations of the impacts of elevated tree mortality and surface fuel buildup on fire risk. This report describes a key data product developed by WG2: annual, statewide, high resolution maps of tree mortality density from 2016-2018.

DATASET QUALITIES

- Metric: dead tree cover (%)
- Range: 0.0 - 1.0 (continuous)
- Spatial extent: California
- Spatial resolution: 30m
- Temporal extent: 2016-2018
- Temporal resolution: annual

DATASET DESIGN

Dead tree cover refers to the horizontal area of a pixel occupied by dead trees, as viewed from above. It is a direct analog to Canopy Cover, a canopy fuel metric provided by LANDFIRE. Scaled from 0-1, it can be interpreted as a sub-pixel fractional cover estimate.

This dataset was designed to provide spatially and numerically continuous estimates of tree mortality density that could be used as an input to the next generation of wildfire behavior models. Several of the dataset qualities—the spatial and temporal resolutions—were selected to align with typical fire modeling and fuels mapping paradigms. The continuous fuel load mapping approach, however, is a slight departure from current fire modeling norms.

Spatially, many fire modeling groups quantify fuel distributions based on the LANDFIRE family of fuel maps, which are provided at 30m resolution. Temporally, the USFS quantifies shifts in the geography and intensity of tree mortality across California via aerial detection surveys, which are collected annually to track progressive mortality patterns.

We provide these datasets at 30m resolution to match the spatial scale of other well-known fuels data, updated annually to capture progressive mortality trends from 2016-2018, a period representing some of the most extensive dieback of the past decade.

The continuous data type is a departure from traditional fuel models. Many fuel models, including Scott & Burgan, are categorical datasets that link fuel classes to a look-up table of biophysical parameters. The dead tree cover metric instead directly quantifies continuous variation in observed tree mortality.

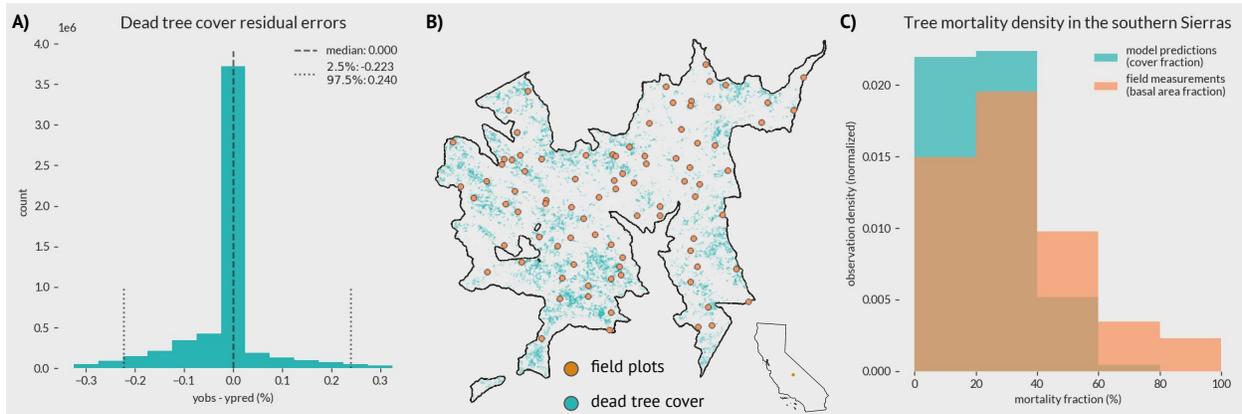


Figure 2: Model validation compared to machine learning training data (A) and to a group of field plots (B, C) independently collected by researchers at the USGS Western Ecological Research Center. We found that tree mortality predictions are generally conservative relative to each reference dataset. A) Compared to the high resolution training data, satellite-based model predictions show low prediction bias with median residual errors centered on zero. Mean absolute error scores were 4.2%, and root mean squared error scores were 9.0%. This difference between MAE and RMSE indicates higher error rates in areas of extreme mortality, where satellite model predictions were generally lower than NAIP observations. B) USGS researchers used a Generalized Random Tessellation Stratified Sampling approach to measure tree mortality in a series of field plots in Sequoia & Kings Canyon National Park. C) We compared histograms of the mortality fractions measured in the field (orange) with satellite predictions (blue), showing a lower frequency of extreme mortality predictions in the satellite maps. In addition to the epistemic model uncertainty, this conservative pattern is likely driven in part by differences in measurement type. Field mortality fractions were estimated by comparing the dead to live ratio of *basal area*; satellite predictions estimate the *cover fraction* of dead trees. Both are horizontal area measurements, but cover fraction is normalized to include patterns like canopy closure or bare soil fraction. Basal area ratios solely measure the relative woody stem area in a plot occupied by dead trees, with no consideration of foliage density, stem density, or total vegetation cover. Basal area mortality fractions will by definition always be higher than corresponding cover fractions, which we expect explains much of the disagreement.

MODELING METHODS

We trained a deep learning model to map dead tree cover using multi-temporal satellite data from multi-spectral and radar sensors. The underlying response variable was created from 1m resolution maps of red-phase conifer mortality, which was averaged to 10m to match the resolution of the satellite covariates. The 10m model was applied to statewide data across each year and masked using a 10m canopy cover map to reduce false positive predictions over bare ground. The masked model predictions were then averaged to 30m to match the desired output resolution.

High resolution training data were generated across a series of priority sites using a combination of machine learning and traditional earth observations/image interpretation techniques. We first applied k-means clustering to 1m resolution National Agriculture Imagery Program (NAIP) data from 2016 and 2018, which groups each pixel into a series of spectrally-similar clusters. Since leaf-on, red-phase dead trees are spectrally similar to bare ground and leaf litter across the range measured by NAIP, these land cover types were often clustered together. So we developed a deep learning model to map tree cover using NAIP imagery, trained on high resolution, LiDAR-derived rasters, and used this to separate dead trees

from bare ground and leaf litter. These high resolution, binary (0/1) maps of dead trees were created for four sites to capture mortality gradients in the ecoregion where tree loss was most severe: one in the northern Sierra Nevada ecoregion, two in Sierra National Forest, and one in Sequoia & Kings Canyon National Park.

The binary maps of tree cover were averaged to 10m, now continuously scaled from 0-1, representing the fraction of each pixel occupied by dead trees, and matching the resolution of the satellite covariates. These included 10-band multi-spectral data from Sentinel-2 and 2-band radar backscatter data from Sentinel-1. The full feature stack was 24 bands, as we included observations from spring and fall time steps.

Multi-spectral data can measure the greenness, dryness, and structure of vegetation, as optical data are directly sensitive to the fractional cover of photosynthetic vegetation, non-photosynthetic vegetation, and bare soil. Radar data are a good complement to multi-spectral data because backscatter signals are directly sensitive to canopy water content. Short-wavelength C-band radar measured by Sentinel-1 is most sensitive to water content in leaves and branches. These features are important because, in red-phase mortality, trees break down their photosynthetic pigments, transitioning leaves to non-photosynthetic tissue, which increases spectral discrimination from live

trees. Less photosynthesis also means less evapotranspiration, which leads to lower water content in leaves and branches.

Multi-temporal satellite data are critical for discriminating between red phase mortality and phenological senescence. Many of the underlying biophysical changes occur in seasonal leaf-off sequences, like the loss of chlorophyll and the drying and dropping of leaves, increasing the likelihood of false positive detection. Fortunately, the timing of these changes is different between mortality and senescence. Phenologically-driven leaf-off conditions are typically short, and balanced by long periods of leaf-on conditions. Post-mortality green-up is rare, however, and red-phase conditions tend to persist for at least a year. Multi-temporal satellite observations quantifying intra-annual variation is critical for distinguishing between these forms of canopy expression. So data from both seasons and both sensors were merged in a 24-band stack then rotated with principal components analysis and normalized to reduce feature covariance and standardize unit scales.

We trained a U-net model, a 2-D Convolutional Neural Network model architecture that belongs to a class of modern pattern recognition algorithms. These models are particularly useful in ecological contexts. 2D convolutions applied at each block quantify the spatial orientations of nearby objects, adding critical contextual information about the surrounding neighborhood to the model, such as the density of nearby trees or the presence of a road. Since there are such strong density-dependent spatial patterns in tree mortality distributions—driven locally by bark beetle population density and tree species composition, and by drought and elevation at landscape scales—contextual information is a critical feature to include for mapping mortality with satellite data.

We sampled image tiles from a series of randomly-sampled point locations across the four sites. We sampled points at a rate of 5 points per km^2 , held out 15% of samples as our test data, then applied a subsequent 85/15 training/validation split for model training. Our final training data included 91,092 training samples, 16,075 validation samples, and 18,909 test samples. Each sample was of shape [64, 64, 24] (height, width, depth). To adjust for sample imbalance—the sparsity of dead relative to live/no trees—we computed pixel-wise sample weights based on an inverse sample frequency algorithm to minimize under-prediction.

The first layer of our network architecture was a 1x1 convolution layer, otherwise known as a network-in-network transformer, which applies dimensionality reduction to reduce the likelihood of overfitting. Then, we applied a traditional U-Net architecture with a 4-

block structure of [4, 3, 2, 2], where each value represents the number of sequential 2D convolution layers in each block. 3x3 convolutions were applied and zero-padded to maintain the 2D shape at each block. Downsampling and upsampling was applied at each block, with passthrough layers concatenating features from the corresponding descending and ascending U-Net blocks. We applied a random dropout rate of 10% at each layer, used ReLU for the internal activation function, and a linear transformation for the output layer activation function. Mean squared error was the loss function, and we used the Adam optimizer with a learning rate of $1e^{-4}$. Model results are discussed in detail in Figure 2.

Conclusions

These satellite-derived maps of tree mortality were designed to provide a quantitative, spatially-explicit evaluation of the impacts of elevated mortality on surface fuel buildup and fire risk. The mapped patterns are consistent with independent USFS measurements of mortality (Fig. 3) and clearly align with the extent of a recent high intensity, mortality-driven wildfire (Fig. 4). This dataset reveals the spatial and temporal patterns of progressive mortality in fine detail, and we expect it could provide great value to the state for understanding and mitigating the wildfire risks posed by extreme tree mortality events.

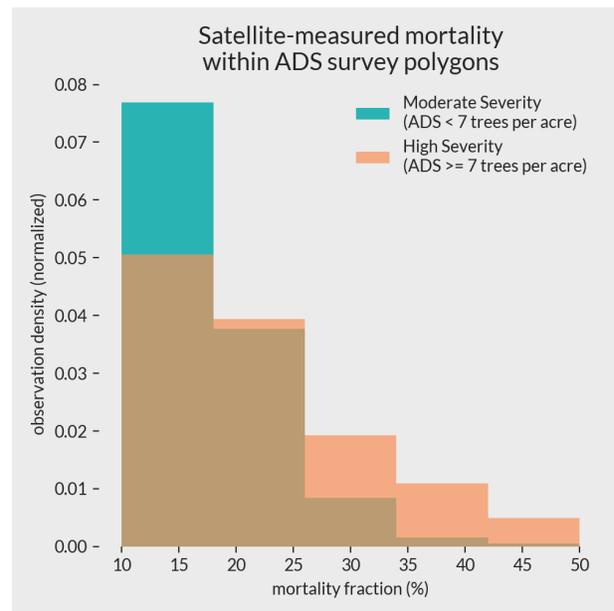


Figure 3: Satellite maps predict higher dead tree cover rates in regions of moderate and high severity mortality, as mapped by the USFS Aerial Detection Survey (ADS).

Appendix A. Progressive mortality preceding the Creek Fire

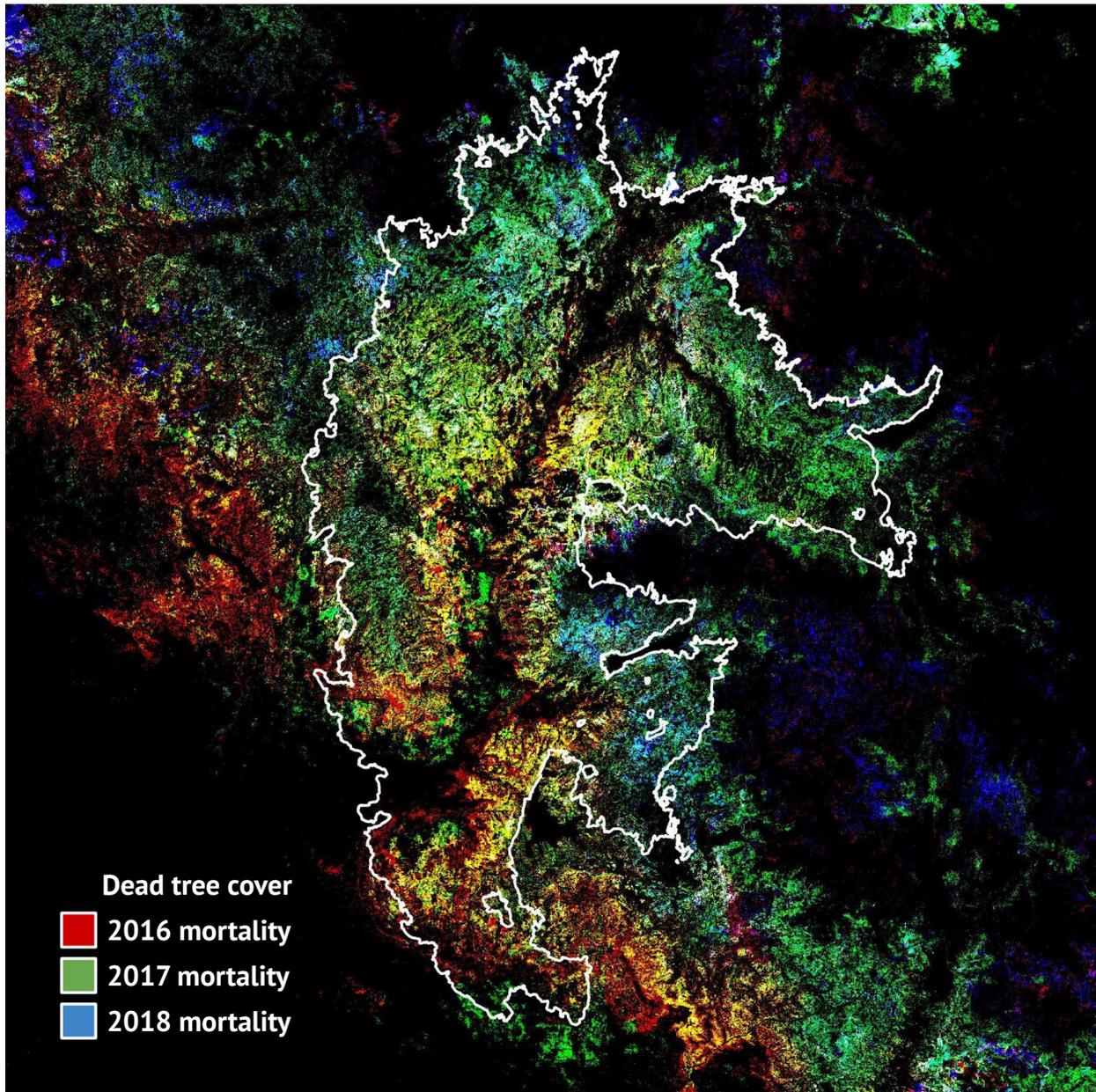


Figure 4: The Creek Fire, outlined in white, burned at extremely high intensity in August 2020 following increases in mortality from 2016-2018. This was one of the first instances of an uncontrolled, high intensity fire burning a region that experienced severe, progressive tree mortality. As the burn perimeter approximates the spatial patterns of satellite-mapped tree mortality, we expect these satellite-derived maps to improve predictions of future mortality-driven wildfires.

Appendix B. Statewide progressive tree mortality composite

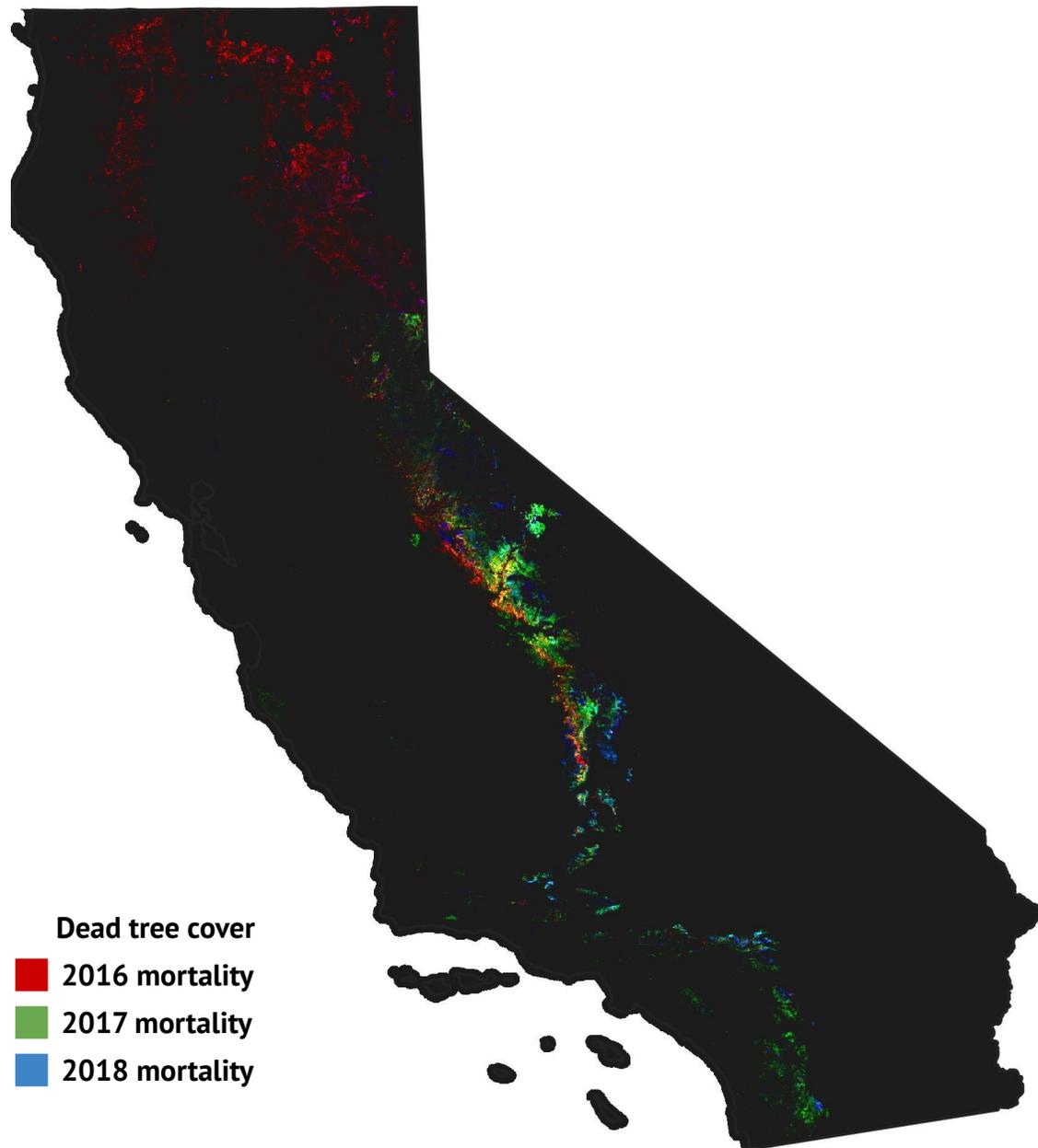


Figure 5: Progressive shifts in tree mortality distributions are mapped here with an RGB composite, where red tracks tree mortality in 2016, green tracks mortality in 2017, and blue tracks mortality in 2018. This map shows how dead tree distributions shifted up-slope in the central and southern Sierra Nevada over this period. Mortality was pronounced in the Klamath and Cascades ecoregions of northern California in 2016 (red), and increased in intensity in the Southern California Mountains ecoregion from 2017-2018 (green, blue).

Appendix C. Comparison to USFS Aerial Detection Survey data

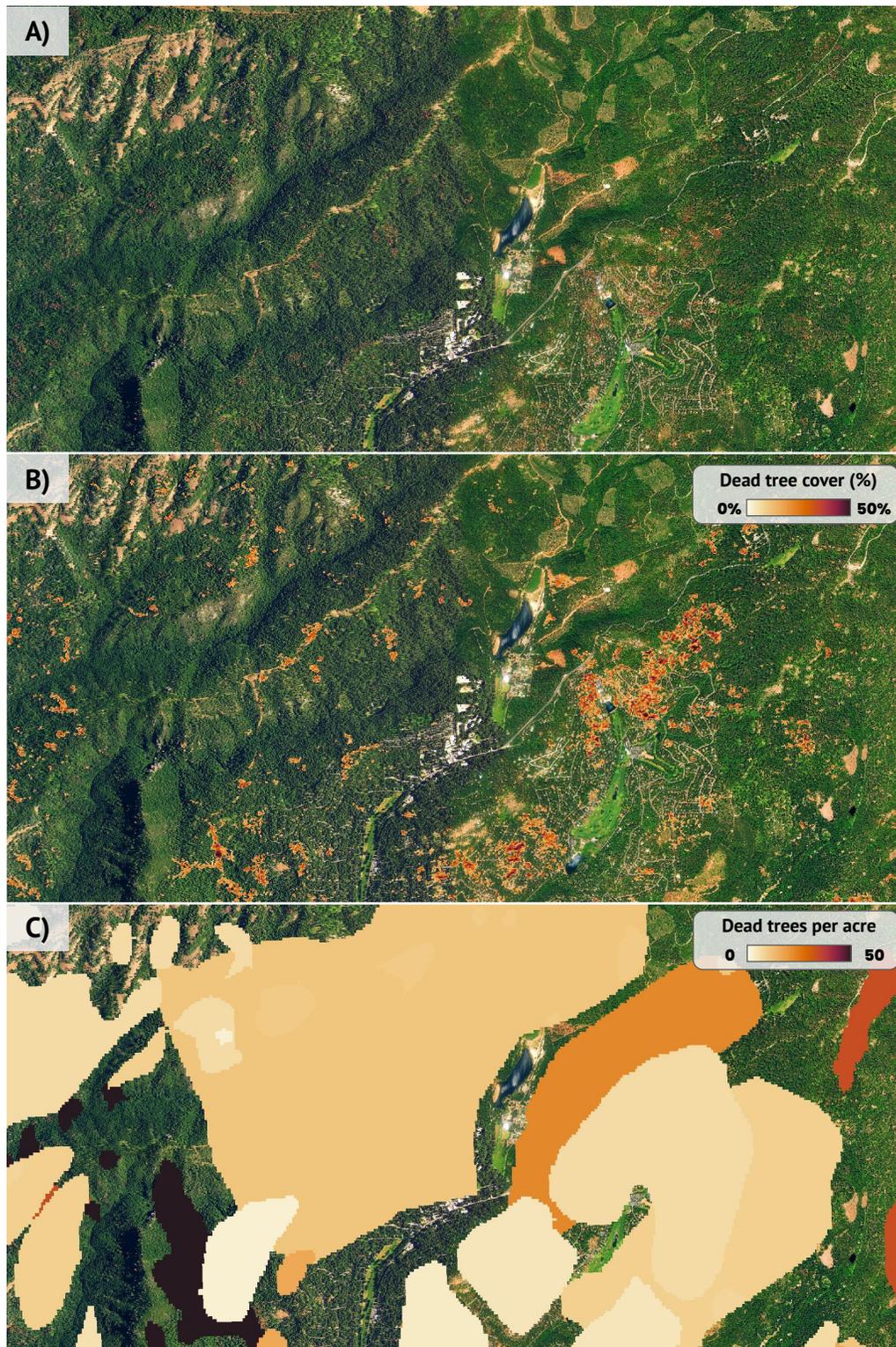


Figure 6: Comparisons between satellite tree mortality maps and USFS Aerial Detection Surveys show agreement in broad spatial mortality patterns, even in mixed-use wildland-urban interface areas. A) 1m NAIP imagery from 2016 for a region in the central Sierra Nevada ecoregion. B) Satellite dead tree cover predictions covering the same extent. C) ADS survey polygons, colored by dead trees per acre. Satellite predictions track the spatial patterns of mortality with high precision.