UNIVERSITY OF CALIFORNIA,
IRVINE

Mapping and Modeling Individual Tree Mortality in California's Sierra Nevada After the 2012-2016 Drought

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Earth System Science

by

Nicole M. Hemming-Schroeder

Dissertation Committee:
Professor James T. Randerson, Chair
Professor Steven D. Allison, Chair
Professor Padhraic Smyth

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

While at the University of California, Irvine, I had the privilege of working with several professors, researchers, graduate students, and undergraduate students. I would like to thank my advisors, Professors James Randerson and Steven Allison. I appreciated learning and engaging in discussions on topics ranging from global biogeochemical cycling to environmental justice throughout my five years of study and research in your research groups. I would also like to thank my committee member, Professor Padhraic Smyth. I enjoyed the computer science course with you and learned a lot from our discussions. Your collective insights and feedback on my research progress have been beneficial to me in becoming an Earth system scientist.

In addition to my doctoral committee, several people have collaborated with me on the work of this dissertation. Thank you very much to Aurora Gutierrez, Carl Norlen, Markelle Kelly, and Luciana Chavez Rodriguez for your contributions and advice on the work presented here. The specific details for each chapter are as follows:

Chapter 1 of this dissertation is adapted from an open-access manuscript published in the *Journal of Geophysical Sciences – Biogeosciences* with permission from John Wiley & Sons and coauthors Aurora Gutierrez, Steven Allison, and James Randerson. This work was published under a Creative Commons license (CC by 4.0, https://creativecommons.org/licenses/by/4.0/). Steven Allison and James Randerson directed and supervised research which forms the basis for the thesis/dissertation.

Chapter 2 of this dissertation is adapted from an unpublished manuscript with permission from coauthors Carl Norlen, Markelle Kelly, Padhraic Smyth, Steven Allison, and James Randerson.

Chapter 3 of this dissertation is adapted from an unpublished manuscript with permission from coauthors Luciana Chavez Rodriguez, Padhraic Smyth, Steven Allison, and James Randerson.

Part of the summary in Chapter 4 is adapted from an open-access manuscript published in the *Journal of Geophysical Sciences – Biogeosciences* with permission from John Wiley & Sons and coauthors Aurora Gutierrez, Steven Allison, and James Randerson. This work was published under a Creative Commons license (CC by 4.0, https://creativecommons.org/licenses/by/4.0/). Steven Allison and James Randerson directed and supervised research which forms the basis for the thesis/dissertation.

I am also grateful to my funding sources throughout my graduate career. In my first year, I received funding through the Graduate Division's Recruitment Fellowship and the Department of Earth System Science at the University of California, Irvine. My second year was funded by the National Science Foundation Research Traineeship Award 1633631. My final three years have been supported by the National Aeronautics and Space Agency

(NASA) Future Investigators in NASA Earth Space Science and Technology program under award 80NSSC20K1626, which is also the source of the funding which supported my work on Chapters 1-3 of my dissertation. In addition, I was funded as a teaching assistant for two quarters and had the privilege of working with undergraduate students learning about environmental data science.

I would also like to acknowledge the Research Cyberinfrastructure Center (RCIC) at the University of California, Irvine (UCI), which maintains the high-performance computing cluster and associated data storage and software tools used throughout this dissertation. These resources helped this work be possible in the timeline in which it was completed.

VITA

Nicole M. Hemming-Schroeder

Education

University of California, Irvine — Irvine, CA
Ph.D. Earth System Science — Sep. 2023
M.S. Earth System Science — Dec. 2020

California State University, Fullerton — Fullerton, CA
M.A. Applied Mathematics — Sep. 2016 – May 2018

University of Wisconsin, Milwaukee — Milwaukee, WI
Post-Baccalaureate Teacher Certification Program — Sep. 2010 – June 2012

University of Wisconsin, Madison — Madison, WI
B.S. Atmospheric and Oceanic Sciences — Sep. 2008 – Aug. 2010

Madison Area Technical College — Madison, WI
Liberal Arts Transfer Student — Aug. 2006 – May 2008

Research Experience

University of California, Irvine — Irvine, CA
Graduate Student Researcher — June 2018 – Present
Advisors: James Randerson and Steven Allison

- Used R and high-performance computing to map and model individual tree mortality from airborne lidar and multispectral data from the National Ecological Observatory Network in the Sierra Nevada and link mortality fraction to Landsat time series
- Prepared presentations and publications for scientific work
- Completed coursework on ecosystem ecology, biogeochemical cycling, remote sensing, physical climatology, and computer science

The Aerospace Corporation — El Segundo, CA
Student Project Lead, CSU Fullerton Capstone Project — Jan. 2018 – May 2018
Advisors: Tien Nguyen, Andy Guillen, and Charles Lee

- Designed an index for assessing the resilience of a satellite enterprise to radio frequency interference for the purpose of assessing optimal locations for additional ground stations

Research Experience, Continued

University of Wisconsin, Madison — Madison, WI
Undergraduate Researcher, Ecometeorology Lab — Sep. 2009 – June 2010
Advisor: Ankur Desai

- Investigated the relationship between water table height and carbon flux residuals of eddy covariance measurements and models

Space Science and Engineering Center — Madison, WI
Student Hourly, Antarctic Meteorological Research Center — Sep. 2008 – June 2010
Advisor: Matthew Lazzara

- Wrote bash and McBASI scripts for satellite data mapping and archival
- Participated in 2009-2010 Antarctic field season to assist with the maintenance of automatic weather stations

Teaching Experience

University of California, Irvine — Irvine, CA
Teaching Assistant, Department of Earth System Science — Spring 2021 and Winter 2022

- ESS 116 Introduction to Environmental Data Science (Winter 2022)
    - Created introductory lectures for the MATLAB-based lab component of the course, facilitated the labs, helped revise assessments, and graded student work
- ESS 118 Analysis, Modeling, and Visualization of Multidimensionnal Environmental Data (Spring 2021)
    - Created Python Jupyter notebooks and associated videos to lead the discussion section for this course and created rubric and intermediate deadlines to scaffold course project

Garden Grove Unified School District — Garden Grove, CA
Mathematics Teacher, Bolsa Grande High School — Aug. 2014 – June 2018

- Provided training in restorative practices and LGBTQ inclusion to build community in the school
- Supported student learning and academic literacy in Algebra I and Geometry through cooperative learning activities and Cornell notes

Sun Prairie Area School District — Sun Prairie, WI
Grade 9 Math Teacher, Cardinal Heights Upper Middle School — Aug. 2012 – June 2014

- Taught Algebra I and Geometry

- Participated in daily collaboration to design lessons and reflected weekly with a mentor to strengthen teaching practice

University of Wisconsin, Milwaukee Milwaukee, WI
Peer Tutor, Panther Academic Support Services Spring 2012
- Used example problems, scaffolding, and verbal questioning to tutor Survey of Calculus and Trigonometry

Madison Area Technical College Madison, WI
Peer Tutor, Learning Center Spring 2006 and
- Tutored mathematics and essay-writing for English 1 Spring 2007

Publications

Hemming-Schroeder, N. M., Gutierrez, A. A., Allison, S. D., & Randerson, J. T. (2023). Estimating Individual Tree Mortality in the Sierra Nevada Using Lidar and Multispectral Reflectance Data. *Journal of Geophysical Research: Biogeosciences*, e2022JG007234. https://doi.org/10.1029/2022JG007234

Hemming-Schroeder, N. M., Gutierrez, A. A., Allison, S. D., & Randerson, J. T. (2023). *Data and code from Estimating Individual Tree Mortality in the Sierra Nevada Using Lidar and Multispectral Reflectance Data*. https://doi.org/10.5281/zenodo.7812035

Griffiths, H. M., Eggleton, P., **Hemming-Schroeder, N. M.**, Swinfield, T., Woon, J. S., Allison, S. D., et al. (2021). Carbon flux and forest dynamics: Increased deadwood decomposition in tropical rainforest tree-fall canopy gaps. *Global Change Biology, 27*(8), 1601-1613. https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.15488

Sulman, B. N., Desai, A. R., **Schroeder, N. M.**, Ricciuto, D., Barr, A., Richardson, A. D., et al. (2012). Impact of hydrological variations on modeling of peatland CO2 fluxes: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research: Biogeosciences, 117*(G1), G01031. https://doi.org/10.1029/2011JG001862

Selected Presentations

Hemming-Schroeder, N. M., Kelly, M., Norlen, C. A., Allison, S. D., Randerson, J. T., (2023). Modeling individual tree mortality in the Sierra Nevada in response to the 2012-2016 California drought. NASA Carbon Cycle and Ecosystems Joint Science Workshop, Baltimore, MD, *Poster Presentation*

Hemming-Schroeder, N. M., Gutierrez, A. A., Allison, S. D., Randerson, J. T., (2021). Modeling tree mortality in the Sierra Nevada under drought conditions. AGU Fall Meeting 2021, New Orleans, LA, *Oral Presentation*

Hemming-Schroeder, N. M., Sulman, B. N., Ashton, L., Yatsko, A. R., (2021) Roles of insects and other animals in woody plant turnover: Consequences for the Earth system. AGU Fall Meeting 2021, New Orleans, LA, *Poster Session*

Downs, S., Wang, J., **Hemming-Schroeder, N. M.**, Randerson, J. T. (2020). Utilizing GEDI data to classify changes in forest structure due to disturbance in California. AGU Fall Meeting 2020, Virtual Conference, *Poster Presentation*

Hemming-Schroeder, N. M., Allison, S. D., Randerson, J. T. (2019). Toward an improved estimate of the carbon pool in coarse woody debris and its representation in Earth System Models. AGU Fall Meeting 2019, San Francisco, CA, *Poster Presentation*

Peer Review

- *Remote Sensing of the Environment*, 2023
- *Journal of Advances in Modeling Earth Systems*, 2022

Outreach

- *Graduate Student Mentor for Noyce Transfer-to-Teaching Scholarship Program*, UC Irvine, Winter 2022
  - Led a two-day workshop for CalTeach mathematics education transfer students on creating math lessons from scientific research
- *Center for Ecosystem Climate Solutions*, UC Irvine, 2020
  - Served as co-mentor for the summer internship program
  - Led a workshop on using R for geospatial analysis
- *Climate Literacy, Education, and Inquiry (CLEAN)*, UC Irvine, 2019-2021
  - Co-led activities with other graduate students on carbon cycling for 6th-grade students at MacArthur Intermediate Fundamental School in Santa Ana, CA
  - Co-created a lesson on tree mortality and wildfire in California
- *Grandparents University*, University of Wisconsin, Madison, 2010
  - Led an interactive presentation for children and their grandparents on what it is like to do field work in Antarctica

Grants, Awards, and Fellowships

- Outstanding Graduate Student Research Award, Department of Earth System Science, UCI, 2021
- Future Investigators in NASA Earth and Space Science and Technology (FINESST) award recipient, 2020-2023
- UCI MAPS (Machine Learning and Statistics in the Physical Sciences) Fellowship (NSF-NRT), 2019
- Community Earth System Model Tutorial Participant at National Center for Atmospheric Research, 2019
- UC Irvine Graduate Dean's Recruitment Fellowship, 2018
- Guy DeRosa Safety in Schools Grant, California Teachers Association, 2016
- Antarctic Service Medal, 2010

ABSTRACT OF THE DISSERTATION

Mapping and Modeling Individual Tree Mortality in California's Sierra Nevada After the 2012-2016 Drought

by

Nicole M. Hemming-Schroeder

Doctor of Philosophy in Earth System Science

University of California, Irvine, 2023

Professor James T. Randerson, Chair

Professor Steven D. Allison, Chair

Widespread tree mortality events occur in temperate forests during periods of severe drought. Severe droughts like the 2012-2016 California drought have become more frequent over the last several decades, threatening forests no longer aligned with their current climate conditions. To better predict what will happen to these forests in the future, we need an improved understanding of the response of forests to severe drought and the capability to predict tree mortality risk at regional scales. The increasing frequency of flight campaigns by organizations like the National Ecological Observatory Network (NEON) provides an opportunity to generate detailed maps of individual tree mortality to support risk analyses at the scale of the individual tree and support satellite-based estimates of tree mortality over a broader domain.

In the first chapter of this dissertation, my coauthors and I used lidar and multispectral surface reflectance from the NEON airborne observation platform to map

individual tree mortality over a 160 km^2 area during and after the 2012-2016 drought for two sites in California's Sierra National Forest. We derived tree locations and crown perimeters from the lidar point clouds and used surface reflectance and changes in crown perimeters between 2013 and 2017 to map 2017 tree mortality for more than one million trees. We found that cumulative tree mortality after the drought could be as high as 50%. In addition, we found that the subsequent effects of wildfire after the drought can be severe, with the Blue Fire of 2021 killing almost the trees within its perimeter. While tree mortality at low elevations appeared to saturate near 50%, cumulative tree mortality at higher elevations can be considerably lower (25%), with elevated rates of mortality continuing in the years after the drought subsides.

In Chapter 2, we used the data set from Chapter 1 to retrospectively investigate biophysical drivers of tree mortality risk from a severe drought at the two sites in the Sierra Nevada using a machine learning method called extreme gradient boosting. Our classification models of tree mortality performed better on the lower elevation site (74% accuracy on the validation and held-out test data sets), which experienced high mortality (50%) during the drought. The most significant driving variables we explored at this site were tree height, distance to rivers, and canopy cover fraction. We found that our models trained on data from one study area did not perform well at the other, highlighting the importance of developing tree mortality benchmarking data sets, which encompass a broad domain for training predictive models of tree mortality.

In Chapter 3, we aimed to create a model to estimate tree mortality fraction over a broader domain. We linked the individual tree mortality data set from Chapter 1 to the Landsat time series using one-dimensional convolutional neural networks. The R^2 values

for the relationship between the mortality fraction observations and convolutional neural network predictions was 0.44 for the entire data set, including pixels with no trees, and 0.57 when we filtered for pixels with at least four trees. Our model enables the expansion of tree mortality estimates to broader spatial domains, which may help uncover fundamental interactions among biophysical drivers of tree mortality needed to generalize process-based tree mortality models at regional scales. This is important because these models are used to predict the biosphere's response to current and future climate to help predict future concentrations of atmospheric carbon.

Our approaches and datasets provide a means to estimate tree mortality and predict tree mortality risk at the scale of individual trees across our study domain in the Southern Sierra Nevada to broader scales across California. These analyses may improve our understanding of forest dynamics after severe drought and subsequent wildfire to improve projections of forest structure and carbon cycling in the Anthropocene.

INTRODUCTION

Increases in greenhouse gas concentrations since 1750 have led to significant warming, particularly since the Industrial Revolution in the 1950s (IPCC, 2014; Steffen et al., 2015). There has also been an increase in the frequency and magnitude of El Niño and La Niña events since 1960 (Cai et al., 2023). These modes of the El Niño Southern Oscillation (ENSO) and corresponding interactions with other large-scale climate modes like the Atlantic Multidecadal Oscillation drive shifts in global precipitation (Kitzberger et al., 2007). In particular, multiyear La Niña events are associated with drought and wildfire in the Southwestern United States (Okumura et al., 2017). Analysis of Earth System Model simulations from phase 6 of the Coupled Model Intercomparison Project (CMIP6) indicate that consecutive La Niña events may be more frequent in the future (Geng et al., 2023).

While typical, slow-onset droughts have increased over the same period, rapid-onset, flash droughts have also increased in frequency (Yuan et al., 2023). In the late spring of 2012, the United States experienced a severe flash drought (Hoerling et al., 2014) following two La Nina events in 2010-2011 and 2011-2012 (Rippey, 2015). In California, precipitation remained low over the subsequent four years between 2012 and 2015, leading to a 1500 mm overdraft from evapotranspiration (Goulden & Bales, 2019). This overdraft dried the soil to depths of 5 to 15 m, leading to significant tree mortality, particularly in the Southern Sierra Nevada (Goulden & Bales, 2019). The United States Forest Service (USFS) estimated from hand-drawn annual Aerial Detection Survey polygons that 129 million trees were dead in California by the end of the drought (USDA, 2017). Key mechanisms of tree mortality under these conditions include hydraulic failure

(Anderegg et al., 2015a) and the combined effects of warm drought and beetle kill (Raffa et al., 2008; Robbins et al., 2021).

Process-based land carbon models may underpredict the impacts of drought on the carbon cycle (Kolus et al., 2019). A review of process-based models for drought-induced tree mortality showed that models that perform well over small study areas may not generalize well to regional scales (Trugman, 2021). One reason for poor performance at regional scales is that models do not always include key processes like bark beetle kill, which can affect relatively large and healthy trees when bark beetle populations reach epidemic levels (Stephenson et al., 2019; Trugman, 2021). A recent study modeling tree hydraulic stress and western pine beetle dynamics during the 2012-2016 drought found that the model could explain 91% of the annual variation in Ponderosa pine mortality fraction observed in each Forest Inventory and Analysis subregion on the validation data set (Robbins et al., 2021).

To uncover important drivers of regional tree mortality risk at finer spatial resolution, we need detailed tree mortality data over spatially continuous regions that encompass a wide domain of biophysical drivers of tree mortality. A recent study by Stovall et al. (2019) mapped the locations of 1.8 million individual trees from 2013 NEON lidar data and classified each tree as dead or alive using 1-meter resolution imagery from the National Agriculture Imagery Program (NAIP). A key finding in their study is that large trees were twice as likely to die as smaller ones (Stovall et al., 2019), which is consistent with beetle preference for large-diameter pine trees (Stephenson & Das, 2020). Because the NAIP imagery is not corrected for atmospheric effects, my coauthors and I argue in Chapter 1 that this study may have overestimated tree mortality in some regions

(Hemming-Schroeder et al., 2023). In a related study, Weinstein et al. (2021) mapped bounding boxes for 100 million trees for 2019 using convolutional neural networks on high-resolution RGB camera imagery from NEON. However, earlier NEON lidar data and hyperspectral data corrected for atmospheric effects captured forest structure and condition before and after the large tree mortality event near the end of the drought.

In Chapter 1, we used high-resolution lidar and hyperspectral data from NEON during and after the 2012-2016 drought to characterize individual tree mortality after the drought. We generated shapefiles of individual tree crown perimeters with live and dead labels for over 1 million trees in the Sierra Nevada forests of California for every year of data availability (2013, 2017, 2018, 2019, and 2021). While the hyperspectral data set in 2013 is missing some tiles, the 2013 lidar data is complete. This allows us to characterize changes in forest structure from treefall and logging between 2013 and 2017 to aid in our tree mortality characterization efforts. Because two wildfires burned through our study sites between 2019 and 2021, we also characterized tree mortality from fire.

Previous work investigating patterns in tree mortality in Sequoia National Park derived from 2-meter resolution hyperspectral data and canopy height models in 2015 showed that topographic features, including elevation, distance to rivers, and granite cover, impact tree mortality risk (Paz-Kagan et al., 2017). We aimed to use our data set of more than 1 million individual trees with live and dead classifications at the nearby Sierra National Park to describe biophysical drivers of tree mortality risk at the scale of the individual tree. We selected a classification model based on insights by Chollet and Allaire (2018), who note that recent classification competitions in machine learning have been won using extreme gradient boosting, a tree-based method developed by Chen and

Guestrin (2015). Unlike random forests, which use an average of the result of many decision trees to make a classification decision, extreme gradient boosting uses batches of decision trees to iteratively generate new trees that focus on the limitations of earlier trees. In Chapter 2, we use extreme gradient boosting on the individual tree mortality data set from Chapter 1 to model individual tree mortality risk from a combination of prospective biophysical drivers, including topography, initial Landsat-based vegetation indices prior to the drought, and climate variables.

Finally, in Chapter 3, we aim to scale our individual tree mortality results, which cover an area of approximately 160 square kilometers, to a broader domain in the Sierra Nevada by linking tree mortality fraction within each Landsat pixel to the preceding Landsat time series. There are many previous methods for mapping land cover and detecting land cover change from Landsat time series. For example, LandTrendr uses a piecewise linear function to segment trends in the Landsat time series (Kennedy et al., 2010), and the Continuous Change Detection and Classification (CCDC) fits a linear harmonic model to the time series and looks for significant changes in the Landsat signal (Zhu & Woodcock, 2014). While the CCDC method performs very well when an entire pixel is affected by the change (e.g., clear-cutting for agriculture), changes that only affect a portion of a pixel, like insect-driven tree mortality, can be difficult for these models to predict (Ye et al., 2021). However, a recent study by Schiefer et al. (2023) showed that unmanned aerial vehicle observations of tree mortality at the centimeter-scale could be upscaled to Sentinel data using long short-term memory recurrent neural networks, a type of deep learning method where earlier time steps influence predictions for future steps.

For Chapter 3, we used another deep learning strategy, convolutional neural networks (CNNs), to analyze the Landsat time series. Convolutional neural networks identify key features in a data array and are often used for image classification. For example, when classifying an image as showing either a dog or a cat, a CNN may rely on features like the shape of the ears to make a classification. Because the CNN convolves a kernel, commonly a 3 x 3 matrix, through the 2D image, it can pick up on features like an ear in many places in the image. With respect to a 1D time series of one or more Landsat-derived spectral indices, this means that the algorithm may be able to pick up on key features in the time series, such as seasonal signals consistent with tree mortality.

Through these investigations, we aim to improve our characterization of tree mortality during and after a severe drought at the scale of the individual tree across a broad spatial domain and describe key biophysical drivers of tree mortality at a regional scale. These new data sets and insights may help inform models of tree mortality risk and forest dynamics as droughts become more frequent and severe in the future.

CHAPTER 1

Estimating Individual Tree Mortality in the Sierra Nevada Using Lidar and Multispectral Reflectance Data

Adapted from the following manuscript with permission from coauthors:

Hemming-Schroeder, N., Gutierrez, A., Allison, S., & Randerson, J. (2023). Estimating Individual Tree Mortality in the Sierra Nevada Using Lidar and Multispectral Reflectance Data. *Journal of Geophysical Research: Biogeosciences*, e2022JG007234.

1.1 Introduction

Temperate forests are estimated to store over 50 petagrams of carbon in live vegetation (Xu et al., 2021). Changing climate patterns may threaten this important carbon pool, leading to a release of stored carbon to the atmosphere. For example, warmer and drier summer conditions have already increased background tree mortality rates among subalpine conifer forests in Colorado since the 1980s (Andrus et al., 2021). Moreover, dynamic vegetation models predict that at least 50% of needleleaf evergreen trees in the Northern Hemisphere will reach mortality thresholds for temperature and pre-dawn moisture by 2100 (McDowell et al., 2016). One key mechanism of tree mortality under drought is thought to be hydraulic failure (Adams et al., 2017; Anderegg et al., 2015a). However, additional factors are needed to explain the magnitude and spatial patterns of large tree mortality events (Trugman et al., 2021). Previous work has shown that the combined effects of drought stress and growing populations of endemic bark beetles contribute to large mortality events in California (Anderegg et al., 2020). Endemic bark beetles were responsible for tree mortality percentages in Ponderosa and Sugar pines as high as 89.6% and 48.1%, respectively, in forests hit the hardest by the 2012-2016 California drought (Fettig et al., 2019).

During the 2012-2016 California drought, evapotranspiration in the Sierra Nevada exceeded precipitation in mid-elevation forests by about 1500 mm in the four years from

2012 to 2015 (Goulden & Bales, 2019). Canopy water content losses were observed throughout California's forests (Asner et al., 2016). Previous studies have found that topographic features related to hydrology play a key role in mortality risk during drought conditions in the Sierra Nevada. For example, lower elevation trees on shallower slopes are at higher risk of dying during drought (Paz-Kagan et al., 2017; Stovall et al., 2019). In addition, trees surrounded by granite outcrops, farther from rivers, and on southwestern slopes have been shown to have a higher risk of tree mortality (Paz-Kagan et al., 2017).

Stovall et al. (2019) used 2013 lidar from the National Ecological Observatory Network (NEON) and imagery between 2009 and 2016 from the National Agricultural Imagery Program (NAIP) to map and classify 1.8 million individual trees over two NEON sites in the Southern Sierra Nevada as dead or alive (Stovall et al., 2019). Large trees (greater than 30 meters in height) were two times as likely to die as smaller trees (trees between 5 meters and 15 meters in height) (Stovall et al., 2019). In contrast, a nearby study in Sequoia National Park over the same drought period showed that trees belonging to the genus *Pinus* were the only group to show an increasing trend in mortality with height, whereas most species showed an increase in resilience with height (Stephenson & Das, 2020). Since the western pine beetle prefers large *Pinus* trees, the authors reasoned that the pine beetles were responsible for the trend in increasing mortality with height for *Pinus* trees (Stephenson & Das, 2020).

To test hypotheses about drivers of tree mortality during drought conditions, we need accurate tree mortality maps that rely on robust and reproducible classification algorithms. In this study, we use an updated set of lidar and hyperspectral reflectance observations available from NEON for the years 2013, 2017, 2018, 2019, and 2021 to

create maps of tree mortality during and after the 2012-2016 California drought. Our approach makes use of these new data resources to build on the work of Stovall et al. (2019) within an open-source framework. We use all available lidar to identify the location of treetops and then identify the perimeters of individual tree crowns based on the lidar data for each year. Tree mortality is based on co-registered hyperspectral reflectance data from the same flights, which are more closely aligned with the lidar data than the NAIP imagery used in past work. NEON data providers also correct their hyperspectral reflectance data using a surface bidirectional reflectance distribution function. As a result, vegetation indices such as relative greenness (green surface reflectance divided by the sum of red, green, and blue surface reflectance) and normalized difference vegetation index (NDVI) derived from their hyperspectral datasets are potentially more robust with respect to variations in solar zenith angle, collection angle, and surface topography. In contrast, for NAIP imagery, little information about the time of day and angle of data collection is publicly available to correct for these variations.

In this study, we compare our revised method for identifying individual tree mortality following the 2012-2016 California drought to the method developed by Stovall et al. (2019) over a similar region and time interval, quantify tree mortality in subsequent years (2017-2021) in response to drought legacy effects and wildfire and evaluate environmental geospatial controls on tree mortality. The method we develop may be useful for tracking tree mortality in other forests where aircraft lidar and spectral data are available. Insights from the regional maps of tree mortality over time may help inform forest stewardship, models simulating forest ecosystem processes, and environmental policy.

1.2 Methods

1.2.1 Study site

We selected NEON's Soaproot Saddle and Lower Teakettle sites in the Sierra National Forest in the Sierra Nevada, which experienced severe drought between 2012 and 2016 (Figure A.1). The spatial extent of these sites is shown in Figure 1.1. The elevations range from 1000 to 1400 meters at Soaproot Saddle, which has a mean precipitation of about 900 mm per year (Krauss, 2018). Because the site is located along the rain-snow transition, winter snow melts quickly (Krauss, 2018). Dominant tree species at Soaproot Saddle include Ponderosa pine (*Pinus ponderosa*), incense cedar (*Calocedrus decurrens*), canyon live oak (*Quercus chrysolepis*), and California black oak (*Quercus kelloggii*) with high mortality observed among Ponderosa pine from bark beetle infestations during the drought (Krauss, 2018). Compared with Soaproot Saddle, Lower Teakettle has a higher elevation and wetter climate with respect to both precipitation and snow accumulation, which provides more moisture during the dry season. The elevation for Lower Teakettle ranges from 2000 to 2800 meters with a mean precipitation of 1220 mm per year (Krauss, 2018). In addition, snow accumulates at Lower Teakettle with a 30-year average maximum depth of 1140 mm (Krauss, 2018). In contrast to Soaproot Saddle, the forest of Lower Teakettle is characterized by red fir (*Abies magnifica*), white fir (*Abies concolor*), Jeffrey pine (*Pinus jeffreyi*), and lodgepole pine (*Pinus contorta*) (Krauss, 2018).

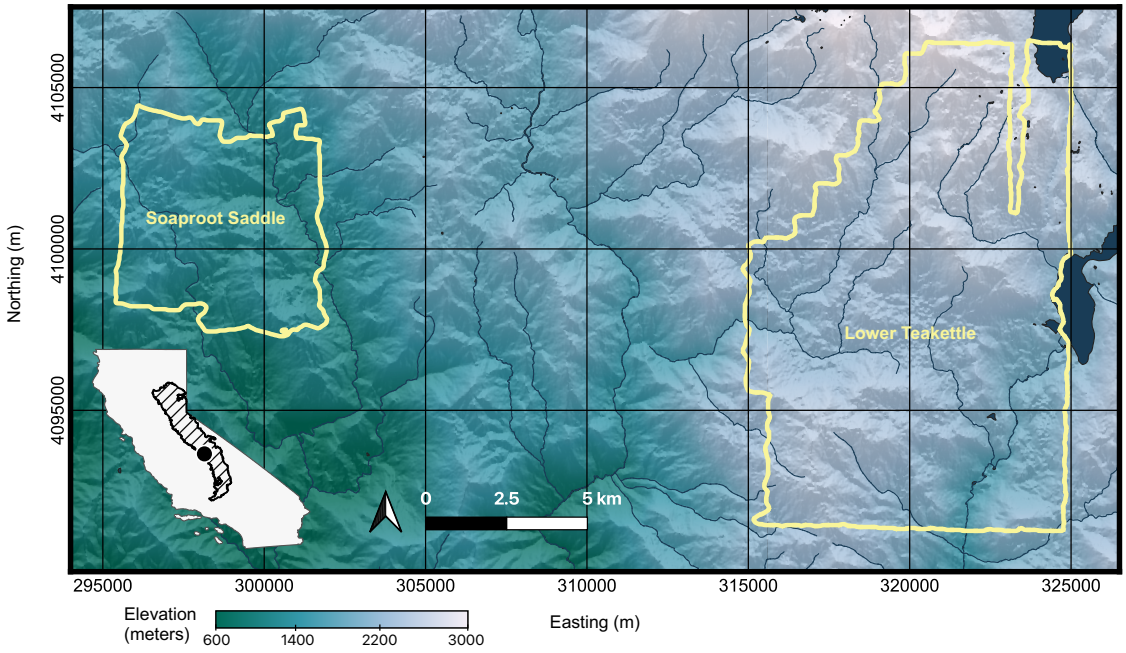

Figure 1.1. The study sites outlined in yellow show the regions of overlapping lidar data between 2013 and 2021 for NEON's Soaproot Saddle and Lower Teakettle sites. Dark blue lines and polygons are lakes and named rivers from the National Hydrography Dataset Plus High-Resolution product from the United States Geological Survey (USGS). The background raster image shows the NASA Shuttle Radar Topography Mission (SRTM) elevation from low elevation (dark green) to high elevation (light purple). We shaded the background from no shade (northern aspect) to light shade (southern aspect) to illustrate the topography of the region. The center location of the zoomed-in image is shown with a black point in the California inset inside the Sierra Nevada ecoregion (black hashed polygon).

The yellow polygons in Figure 1.1 show the overlap of high-resolution lidar from NEON's airborne observation platform for the years 2013, 2017, 2018, 2019, and 2021 (NEON, 2021). NEON also has co-registered, high-resolution (1 meter by 1 meter)

hyperspectral data (NEON, 2022b) corrected for atmospheric effects and topography (Karpowics & Kampe, 2022) over the same spatial extent and years except for some missing data in 2013. These rich datasets provide the opportunity to make detailed tree mortality maps over time.

1.2.2 Datasets

1.2.2.1 Lidar and canopy height models

We downloaded lidar point cloud tiles collected from NEON's airborne observation platform (NEON, 2021). The lidar resolution for 2013 is 3 points/m^2 and increases to 6 points/m^2 in 2017-2019 and 18 points/m^2 in 2021. Each lidar point is described by an x-, y-, and z-coordinate. The empirically-derived horizontal error is 3cm–5cm, while the vertical error is estimated to be 5cm–10cm (Krause & Goulden, 2022). To reduce noise, we filtered out any altimeter values more than five standard deviations from the mean. We also removed altimeter values greater than 120 meters in height after normalizing the lidar to account for the surface topography.

In addition, we downloaded NEON's canopy height model raster data (NEON, 2022a) to help create vegetation masks to aid in the granite-detection algorithm described in section 1.2.4 and to compute the canopy cover fraction. A recent study found significant effects of canopy cover on microclimate using plots ranging from 5 to 1300 m^2 in area (Zellweger et al., 2020). We chose to use the 2013 canopy cover fraction within a 20-meter radius of each tree's location (an area of about 1260 m^2) to investigate tree mortality

related to initial microclimate conditions. Because we find a polygon describing the crown perimeter of each tree for each year of data (section 1.2.2.5), we exclude the individual tree's 2013 crown perimeter from the 20-meter buffer when calculating the canopy cover fraction.

We also downloaded NEON's digital terrain model raster data (NEON, 2023) to derive slope and aspect to assess potential tree mortality drivers. Both the canopy height and digital terrain model data have a resolution of 1 meter by 1 meter. To reduce potential noise in the high-resolution digital terrain model, we calculated the median slope and aspect within a 10-meter radius of each tree. The algorithm we used to find each tree location is described in section 1.2.2.5.

1.2.2.2 Hyperspectral reflectance data

We downloaded mosaic orthorectified hyperspectral reflectance data from NEON's airborne observation platform for the years 2013, 2017, 2018, 2019, and 2021 (NEON, 2022b). The surface radiance collected by the airborne observation platform is transformed to surface reflectance by removing atmospheric effects from haze, cloud shadow, waterbodies, water vapor, and topography and applying a bidirectional reflectance distribution function (Karpowics & Kampe, 2022). This dataset includes 426 spectral bands and is distributed at a spatial resolution of 1 meter by 1 meter. The wavelength interval for band lengths between 340 nm and 1349 nm is 0.40 nm, while the interval from 1351 nm through 1449 nm is 0.98 nm. We approximated the Landsat 8 spectral bands by taking the mean surface reflectance for blue (452-512 nm), green (533-

590 nm), red (636-673 nm), near infrared (851-879 nm), and shortwave infrared 1 (1566-1651 nm) to create spectral indices and natural-color (red-green-blue or RGB) images for training labels. We also selected two water absorption bands, water band 1 (w_1) from 1440 through 1460 nm, and water band 2 (w_2) from 1935 through 1955 nm, because the water in leaves is a strong absorber of radiation in these wavelengths (Carter, 1991). Reflectance in the water bands should be lower where there is higher leaf water content. All wavelength ranges listed here are inclusive of the boundaries.

From the spectral bands, we created five vegetation indices including relative greenness, NDVI, normalized difference moisture index (NDMI), and two additional indices like NDMI created with the water absorption bands instead of the shortwave infrared 1 band. These indices are described in more detail in Table A.1. Because the relative greenness yielded the highest accuracy in the algorithm described in section 1.2.3 (Table 1.1), we focus on relative greenness for this study. Diurnal variations in canopy moisture during the air campaigns may have contributed to lower performance of the moisture indices.

1.2.2.3 National Hydrography Dataset

We downloaded the National Hydrography Dataset Plus High-Resolution for region 1803, which encompasses our study site from the USGS Earth Explorer website (U.S. Geological Survey, 2019). While we show named rivers within this dataset in Figure 1.1, we used all rivers in the dataset for our analysis of the relationship between distance to rivers and tree mortality.

1.2.2.4 Landsat surface reflectance

We retrieved Landsat Collection 2 Level 2 Tier 1 surface reflectance data from the USGS Earth Explorer website. Path 42, row 34 completely covers our region of interest, so we used this tile in our analysis. To compute NDMI, we collected the near infrared band (NIR; Band 4 for Landsat 5 and Band 5 for Landsat 8) and shortwave infrared 1 band (SWIR1; Band 5 for Landsat 5 and Band 6 for Landsat 8) and applied the formula

$$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1} \tag{1.1}$$

NDMI ranges from -1 to 1 where -1 represents dry conditions (low near infrared surface reflectance relative to shortwave infrared 1 surface reflectance) and 1 represents moist conditions. The change in dry season NDMI has been found to correspond to tree mortality (Goulden & Bales, 2019), so we compared our estimates of tree mortality against dry season NDMI. To obtain dry season NDMI, we selected the four Landsat tiles temporally nearest September 15th for a given year, filtered each tile for clouds, and took the median of the surface reflectance pixels for the surface reflectance bands before computing NDMI. We aggregated 2013 and 2017 tree mortality to the Landsat resolution and took the difference to compare to the corresponding change in NDMI from Landsat between late summer 2013 and 2017. Because some hyperspectral data is missing in 2013, this analysis does not span the entire region of interest shown in Figure 1.1.

1.2.2.5 Tree polygon algorithm

Using a 96-meter-by-96-meter square moving window with a 24-meter buffer, we combined all the lidar point clouds from 2013, 2017, 2018, 2019, and 2021 for the Soaproot Saddle and Lower Teakettle sites into one combined point cloud of x, y, and z locations. We combined the point clouds so that we would detect any trees that were standing at any point between 2013 and 2021. If a tree fell or was logged between 2013 and 2017, for example, we would have the treetop location from the combined point cloud and use the corresponding change in crown area from one year to the next to classify it as dead. To filter noise in the lidar point clouds, we dropped any z-values greater or less than 5 standard deviations from the mean. Following *lidR* package documentation, we first rasterized the canopy height to 1-meter resolution and smoothed the canopy height raster using a 3 × 3 median filter. Elsewhere in our methods, we used the NEON-processed canopy height models; however, for this portion of the methods, we generated the canopy height model corresponding to our moving window as we processed the lidar point cloud. We used the *find_trees* function (R package*: lidR*) to identify the treetops of the combined point cloud using the local maxima filter *lmf* (R package: *lidR*). The *lmf* function uses its own moving window on the smoothed canopy height raster to identify local maxima and classify them as treetops. We wanted the length of the local maxima filter window to grow with the height of the trees as suggested by the *lidR* package documentation, so we adjusted the window size according to the following function:

$$w(h) = \begin{cases} 4m & if\ h < 32m \\ \frac{1}{8}h & if\ 32m \leq h \leq 80m \\ 10m & if\ h > 80m \end{cases} \quad (1.2)$$

where *w* is the width of the window and *h* is the height of the canopy, both in meters. We decided on the parameters for this function through trial and error on a subset of trees to reduce the number of trees that are counted more than once when the window is too small and that are missed when the window is too large.

After finding the locations of the trees from the whole point cloud, we went through each year of lidar data (2013, 2017, 2018, 2019, and 2021) to assign lidar points to each tree for each year. To do this, we used the function *segment_trees* with the *dalponte2016* algorithm (R package: *lidR*) which has three adjustable parameters (Dalponte & Coomes, 2016). We selected a minimum threshold of 3 meters and adjusted the rules for adding returns to a given tree so that returns need to be at least 65% of the height and 75% of the mean collection of returns assigned to a given tree. Because there are partial trees around the edges of the 96-meter-by-96-meter window, we filtered out any trees with treetop locations in the 24-meter buffer to prevent partial trees or duplications in the segmentation process. We used the *concaveman* function to draw polygons around the x- and y-coordinates of each tree for each year (*R package: concaveman*) and adjusted the height of each tree using the canopy height model for each lidar year. The resulting attributes for each tree are the crown perimeter polygons and height for each year surveyed by NEON air campaigns. All analyses were completed using R, version 4.1.2.

1.2.2.6 Training data

We created an evenly spaced training grid. First, we drew a bounding box around all the NEON lidar data files for all years (2013, 2017, 2018, 2019, and 2021). Then, we selected a reference Landsat 8 image for path 42, row 34, which completely covers our region of interest. We transformed the raster image to the same coordinate reference system as the NEON lidar data (UTM, Zone 11, CRS: 32611). The pixel size of the Landsat 8 images is 30 meters by 30 meters. Starting 315 meters from the edges of the bounding box, we created an evenly spaced grid, with intersections separated by 300 meters in both the x- and y-directions, resulting in 120 points in the x-direction and 72 points in the y-direction for a total of 8640 grid points. Each grid had a 30-meter by 30-meter square perimeter coinciding with Landsat data pixels.

NH and AG manually labeled trees mutually exclusive sets of trees inside these training squares using 2017 RGB imagery that we derived from the NEON hyperspectral data following our flowchart shown in Figure A.2. First, for any tree that had a crown area greater than 1 square meter that drops below 1 square meter in 2017, 2018, and 2019, we labeled the tree dead for 2017. Next, if a quarter or more of the tree crown polygon was taken up by road, ground, or part of another tree, we labeled it as uncertain. If neither of those criteria were met, we considered evidence that the tree was alive or dead in 2017. If at least one third of the crown area of the tree was red or gray, we labeled it as dead if the same was true in 2018 and 2019. If the tree appeared dead by crown area color in 2017 (i.e., if at least one third of the crown area was red or gray) but more than two-thirds of the crown area in 2018 or 2019 was green, we labeled it as alive. Alternatively, if the crown

area in 2017 was more than two-thirds green, we labeled it as alive. If we could not see that the tree was alive in 2017, we labeled it as live if at least two thirds of the crown area was green in 2018 or 2019 under the assumption that a tree alive in 2018 or 2019 must be alive in 2017.

We also recorded three levels of certainty (0, 0.5, and 1) where 0 refers to no certainty, 0.5 is less certain, and 1 is relatively certain. These certainties are defined in further detail in the flowchart shown in Figure A.2. While our analysis considered 0.5 and 1 acceptable levels of certainty, future analyses could filter out the trees with certainty values of 0.5. Additional detail for the assignment of labels and certainty values is included in Figure A.2. We used the flowchart to label 9025 trees and labeled 8897 as dead or alive using our protocol. The life status of 128 trees was unclear, and these trees were not included in fitting parameters for the tree mortality algorithm. We labeled 6377 trees (71.7%) as alive and 2520 trees (28.3%) as dead. When we filtered out trees with no corresponding hyperspectral data, we had a total of 8885 labeled trees for our algorithm.

1.2.2.7 Wildfire perimeters

We downloaded the National United States Forest Service Final Fire Perimeters to quantify the number of trees that died after wildfires in our region of interest (USDA Forest Service National Forest System Lands GIS and Fire personnel, 2023). The 2020 Creek Fire and 2021 Blue Fire both burned through the Soaproot Saddle site, affecting tree mortality between 2019 and 2021.

1.2.3 Tree mortality classification algorithm

Our algorithm for generating tree crown perimeter polygons and classifying the trees and live or dead is shown in Figure 1.2. We explored five vegetation indices (relative greenness, NDVI, NDMI and two additional normalized difference moisture indices derived from two water absorption bands) described in Table A.1 to classify trees as dead or alive. Previously, Stovall et al. (2019) masked NAIP imagery where the canopy heights were less than 5 meters to isolate tree canopies and limit tree height to 5 meters and greater. We followed the 5-meter minimum height for our tree dataset, but we masked the hyperspectral data corresponding to canopy height model values less than 4 meters so that we could still collect some lower canopy returns within the crown perimeters, particularly for trees near 5 meters in height. In addition, we found that snow confounded tree mortality estimates in high elevation regions in early summer, so we also masked pixels where the mean luminosity (mean of the red, green, and blue bands) was greater than 0.2. For reference, we set the minimum reflectance to 0 and and the maximum to 0.2 when viewing the RGB imagery in the Quantum Geographic Information System (QGIS) software for labeling trees for training data. Reflectance values greater than 0.2 for the three visible light bands were bright surfaces such as snow or granite.

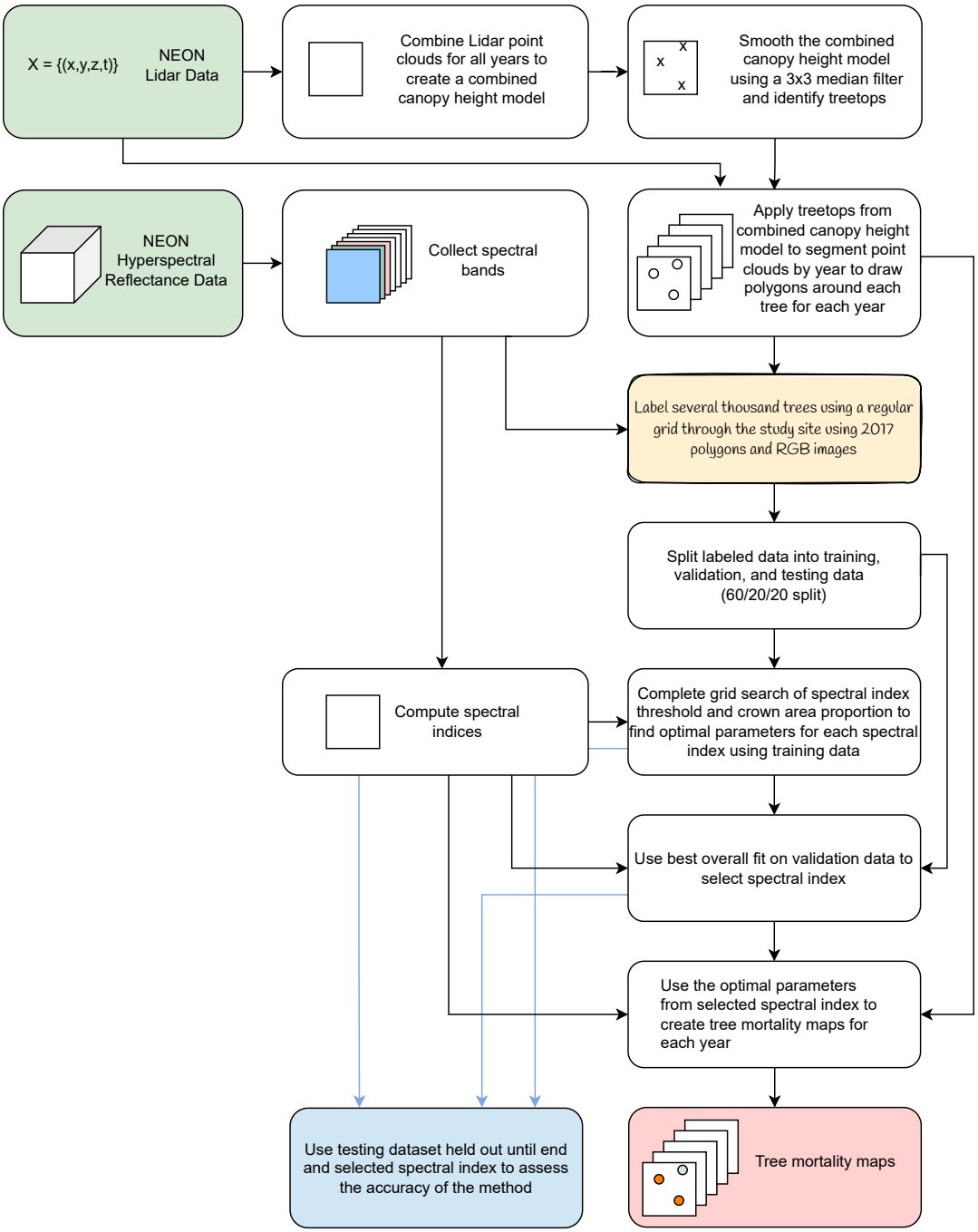

Figure 1.2. This diagram shows the starting datasets (green boxes at top-left), R scripts (white boxes), manual methods in QGIS (yellow box in the middle), the resulting tree mortality maps (red box at bottom-right), and accuracy (blue box at bottom-left).

We partitioned our 8885 labeled trees into a training, validation, and test dataset by randomly sampling 60% of the labels (5331 trees) for the training set, 20% of the labels for the validation dataset (1777 trees), and held out the remaining 20% of the labels (1777 trees) for the testing dataset. We used both a validation dataset and testing dataset, so that we could use the validation dataset to select the best spectral index while setting the test dataset aside for the final accuracy assessment. We set a seed for the random number generator within R to ensure reproducibility for the training, validation, and testing split.

Our model for classifying dead and live trees has two parameters that use an individual vegetation index derived from NEON's hyperspectral reflectance data. For all five indices we considered, increasing values of the index towards 1 indicate healthier vegetation (e.g. more green or moist), so we used the same algorithm to explore all five indices. The first parameter is a vegetation index threshold below which we label a pixel within a tree's crown perimeter dead. For example, healthy tree canopies typically have high NDVI values for which near infrared surface reflectance is high relative to red surface reflectance. The second parameter follows the methods of Stovall et al. (2019) and is a threshold of dead pixels within a tree crown perimeter over which the tree is classified as dead. For example, Stovall et al. (2019) classified a tree as dead if more than 37.5% of its pixels within the crown area were classified as dead with respect to falling below an NDVI threshold.

We considered vegetation index thresholds from 0 to 1 with an increment of 0.01 and crown area fraction thresholds of 0 to 1 with an increment of 0.01. We computed the accuracy on the training dataset for every combination of these two parameters. The grid search for the vegetation indices with the highest training dataset accuracy (NDVI and

relative greenness) are shown in Figure 1.3. Here, accuracy is defined to be the number of correct classifications of a tree as dead or alive over the total number of classifications. The dark blue regions of the parameter space maps indicate an accuracy resulting from guessing that most of the trees are dead, while the green regions result from guessing that most of the trees are alive. When we classify trees with at least 37% of pixels under a relative greenness of 0.37 as dead, we achieve an accuracy of 93.3% on the training dataset.

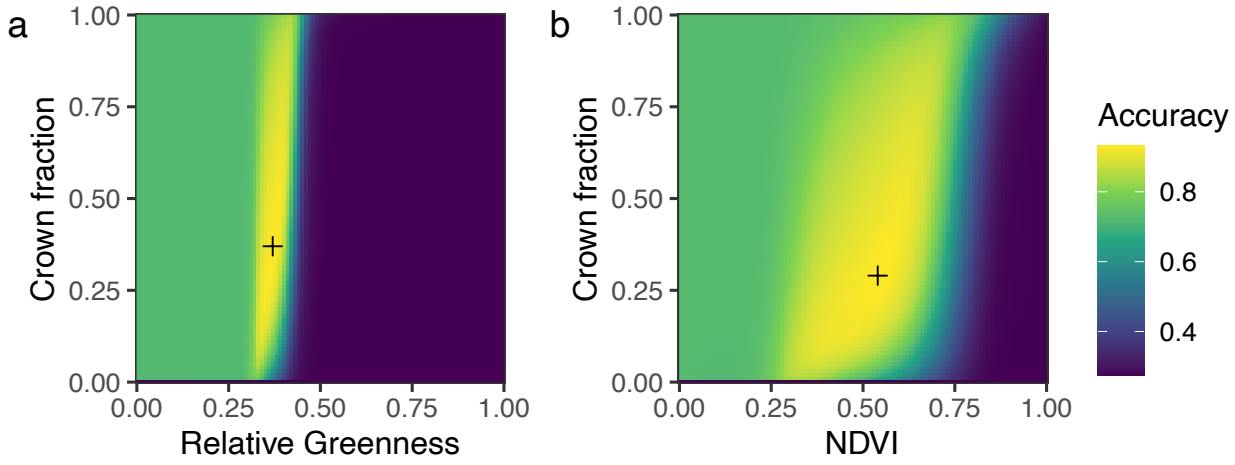

Figure 1.3. Accuracy levels for different combinations of vegetation index and crown area fraction parameters used to identify live or dead trees. In each panel, the x-axis shows the vegetation index threshold used to identify a pixel of a tree as live or dead. **a.** The x-axis shows the relative greenness thresholds from 0 to 1 with an increment of 0.01. The y-axis indicates crown area fraction from 0 to 1 with an increment of 0.01. Dark blue shows the lowest accuracies, while yellow shows the highest accuracies. The highest accuracy on the training dataset (93.3% shown on the map by a black cross) occurs when we label trees as dead when at least 37% of their crown area is under a relative greenness of 0.37. **b.** The x-axis shows the NDVI thresholds from 0 to 1 with an increment of 0.01. The y-axis indicates

crown area fraction from 0 to 1 with an increment of 0.01. The highest accuracy for NDVI on the training dataset (92.6%) occurs when we label trees with at least 29% of their crown area has an NDVI less than 0.54.

1.2.4 Application, validation, and testing of method

Because the number of labels is much smaller than the total number of trees within the domain, we analyzed the accuracy of the parameters in our validation dataset in relation to the number of labeled training data values used to fit the parameters to see if we labeled enough trees. We set a seed in R and randomly sorted the training labels. Then we selected the first *n* labels to use as training data where *n* begins at 10 and increases by increments of 10. We ran the whole algorithm apart from evaluating the performance on the testing dataset for each quantity of cumulative labels. The outcome from this analysis is shown in Figure 1.4. We found that the accuracies on the training and validation dataset converge after including just a couple thousand trees for training data.

The optimal parameters and accuracy analysis of the indices explored in this study are arranged in order of training dataset accuracy in Table 1.1. We consider a true positive to be an instance where a dead tree is correctly classified as a dead tree. Sensitivity refers to proportion of true positives, which are the total number of accurate classifications of dead trees over the total number of dead trees. Specificity refers to the number of correctly identified live trees over the total number of live trees. Of the five indices explored, relative greenness yielded the highest accuracy on the training and validation datasets and also had the highest sensitivity and specificity on the validation dataset.

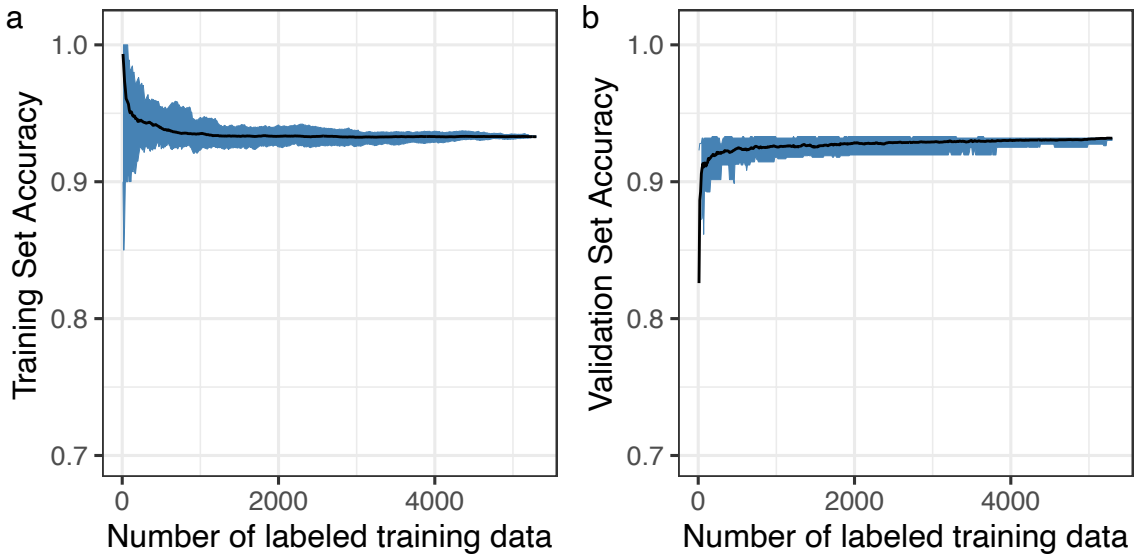

Figure 1.4. The accuracy of the of the parameters generated by a given number of labeled training data is assessed for the **a)** training data used and **b)** the 20% of data held out for validation (1777 values).

The confusion matrices for the training, validation, and testing datasets are shown for relative greenness in Table 1.2. The accuracy of our relative greenness model was 92.1% on the dataset held out for testing. A sample image of trees labeled by our algorithm is shown in Figure A.3. Shadows visible in the RGB image (Figure A.3a) appear to be removed by the normalization in our relative greenness metric (Figure A.3b).

Table 1.1

Performance of the Vegetation Indices Explored in Chapter 1

Vegetation index	Vegetation index threshold	Crown area fraction threshold	Training accuracy	Validation accuracy	Validation sensitivity (correctly classified dead trees divided by total dead trees)	Validation specificity (correctly classified live trees divided by the total live trees)
Relative greenness	0.37	0.37	**93.3%**	**93.2%**	**96.7%**	**84.8%**
NDVI	0.54	0.29	92.6%	91.7%	96.2%	81.1%
NDMI1	0.38	0.32	91.9%	91.3%	95.5%	81.3%
NDMI	0.21	0.36	91.9%	91.4%	96.1%	80.2%
NDMI2	0.59	0.40	91.0%	89.8%	93.6%	80.8%

Table 1.2

The Confusion Matrices for Relative Greenness for Training, Validation, and Testing Datasets

	Training accuracy: 93.3% (N=5331)		Validation accuracy: 93.2% (N=1777)		Testing accuracy: 92.1% (N=1777)	
	Estimated dead	Estimated live	Estimated dead	Estimated live	Estimated dead	Estimated Live
Labeled dead	1234	234	445	80	429	93
Labeled live	124	3739	41	1211	47	1208

1.2.5 Granite-detection algorithm

We hypothesized that higher granite cover fraction near a tree would be associated with negative health outcomes for trees based on soil quality as in a previous study (Paz-Kagan et al., 2017) and due to higher relative sensible heating and less latent cooling among trees surrounded by granite outcrops compared with those surrounded by

vegetation. To explore this hypothesis, we developed a granite-detection algorithm. We classified all pixels with a canopy height less than 0.5 meters and a luminosity greater than 0.11 as granite pixels. Lastly, we rasterized National Hydrography Dataset waterbody polygons and USFS roads polygons to filter out any pixels that are associated with waterbodies or roads. To investigate the role of granite outcrops in tree mortality, we calculated the fraction of ground pixels within 20 meters of a tree that are classified as granite.

1.2.6 Tree matching algorithm

We downloaded the Stovall (2019) dataset from figshare and matched the trees to trees in our dataset to enable a comparison of our method with the previous work it builds upon. For each tree in our dataset, we found the distance from the x- and y-coordinates of the 2013 treetop location and each tree in the Stovall dataset. If the closest tree in the Stovall dataset fell inside the 2013 tree crown perimeter for our tree, we considered it a match. We matched 714,656 trees out of 1,011,634 trees located in our region of interest with overlapping NEON spectral data. Our comparison is based on this set of 714,656 trees.

We used the matched dataset to compare the distribution of the mean NDVI within the crown area of trees in 2017 and over time for each of the following categories: 1) trees where both studies label the tree as alive, 2) trees where both studies label the tree as dead, and 3) trees we labeled as live in 2017 but Stovall labeled as dead in 2016. We use the first two categories as reference datasets to determine whether the trees in the third category are more likely to be alive or dead in 2017.

1.3 Results

1.3.1 Tree mortality during and after the 2012-2016 drought

We estimate that 324,301 of a total 1,011,577 trees (32.1%) were dead in the study domain by 2017 with 25.4% of the total trees dying between 2013 and 2017 (Figure 1.5). Tree mortality was considerably higher at Soaproot Saddle than Lower Teakettle (Figure 1.5a and 1.5b for cumulative numbers and Figure 1.6 for spatial maps). Of the 275,103 trees we mapped at Soaproot Saddle, we classified 49.6% as dead by 2017 (Figure 1.5a). In contrast, of the total 736,474 trees we mapped at Lower Teakettle, only 25.5% were dead by 2017 (Figure 1.5b). The differences between the two sites were further amplified when considering the change in mortality between 2013 and 2017. During this interval, tree mortality increased by 45.8% at Soaproot Saddle and by 17.4% at Lower Teakettle.

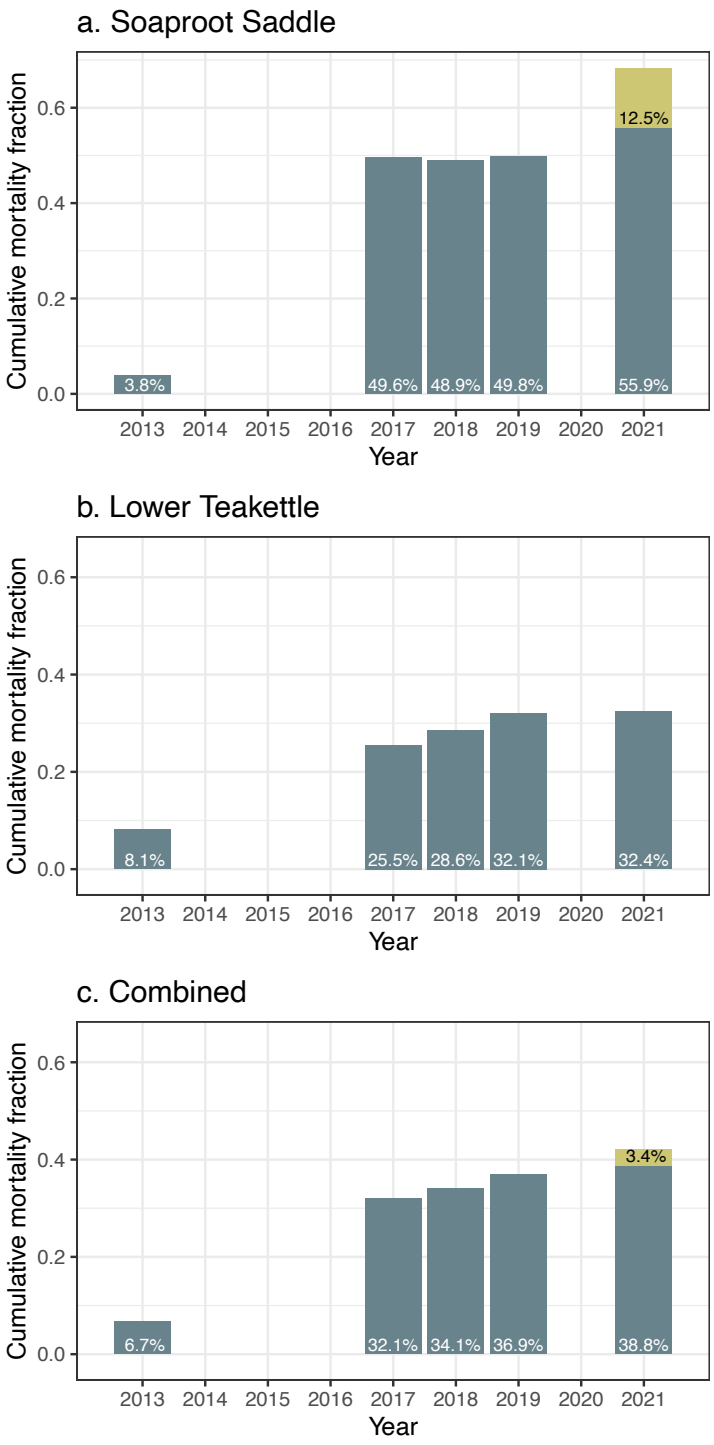

Figure 1.5. The cumulative tree mortality fraction for trees in the overlapping region of the lidar point clouds for each year of available data (2013, 2017, 2018, 2019, 2021) is shown for **a)** Soaproot Saddle, **b)** Lower Teakettle, and **c)** the combined dataset. The light gold in panels a and c represents the mortality between 2019 and 2021 within the perimeters of the 2020 Creek Fire and 2021 Blue Fire. For each year of data, we included trees that had crown areas greater than 1 square meter in 2013, a 2013 height greater than 5 meters, and spectral data for the years 2017, 2018, 2019, and 2021. Due to missing spectral data in 2013, we were only able to classify 744,727 trees from 2013 in the study region as dead or alive using our algorithm, so the mortality fraction shown for 2013 is based on those 744,727 trees. For the later years, there were 1,011,577 trees that met our criteria.

Maps of the temporal evolution of tree mortality are shown for Soaproot Saddle in Figure A.4 and for Lower Teakettle in Figure A.5. We found very little additional tree mortality at Soaproot Saddle between 2017 and 2019; however, additional mortality was visible by 2021 when absolute tree mortality reached a cumulative level in non-fire affected areas of 55.9% (Figure 1.5, Figure 1.6). In Lower Teakettle, visible increases in tree mortality occurred each year between 2017 and 2019, with little additional change in the final interval between 2019 and 2021. There are some areas that appear misclassified in 2019 because trees labeled dead in 2019 are labeled live later in 2021. We explore some key areas of inconsistent classification in our 2019 tree mortality estimates in Figure A.6. One region in Lower Teakettle (Figure A.6a) appears to have high levels of snow cover and low luminosity, while the second region (Figure A.6b) has trees that were brown in 2019 but green in 2021.

The Creek Fire in 2020 and the Blue Fire in 2021 contributed to significant additional tree mortality at Soaproot Saddle. While most of the Creek Fire affected areas of the Sierra National Forest to the northwest of our region of interest, part of the wildfire extent intersected our study domain, affecting 76,644 trees (Figure A.4, panel d). Within the Creek Fire perimeter intersecting our study domain, 45.2% of the trees were dead in 2019 prior to the start of the fire. After the fire in 2021, 85.5% of the trees within the fire perimeter were dead (an increase by 89%). We found that tree mortality within the 2021 Blue Fire perimeter increased from 64.3% to 95.6% between 2019 and 2021 (an increase of 49%) among the 11,068 trees within the Blue Fire's perimeter. We show a zoomed in image of the 2021 Blue Fire burn area in Figure A.7.

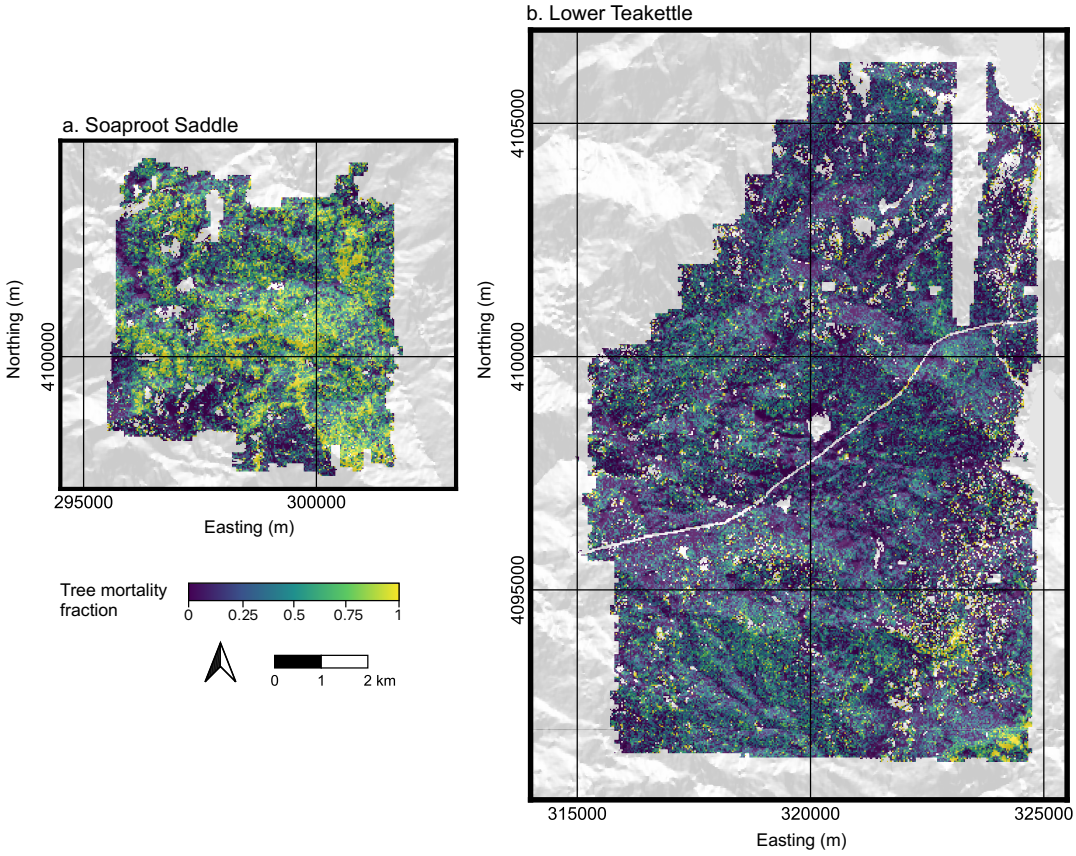

Figure 1.6. The spatial pattern of tree mortality in 2017 averaged on a 30-meter-by-30-meter grid for **a)** Soaproot Saddle and **b)** Lower Teakettle. Light yellow pixels indicate 100% tree mortality for the trees in that pixel, while dark blue indicates 0% tree mortality. The aspect is converted from 0 degrees to 360 degrees to a scaled value reflecting North (1) to South (0) and overlayed in white to gray to show the underlying topography.

We plotted the mean relative greenness and mean NDVI density distributions for our estimates of live and dead trees in Figure 1.7. The relative distributions for mean relative greenness reach a maximum of 0.42 for live trees and 0.33 for dead trees. The mode for mean NDVI is 0.76 for live trees and 0.37 for dead trees.

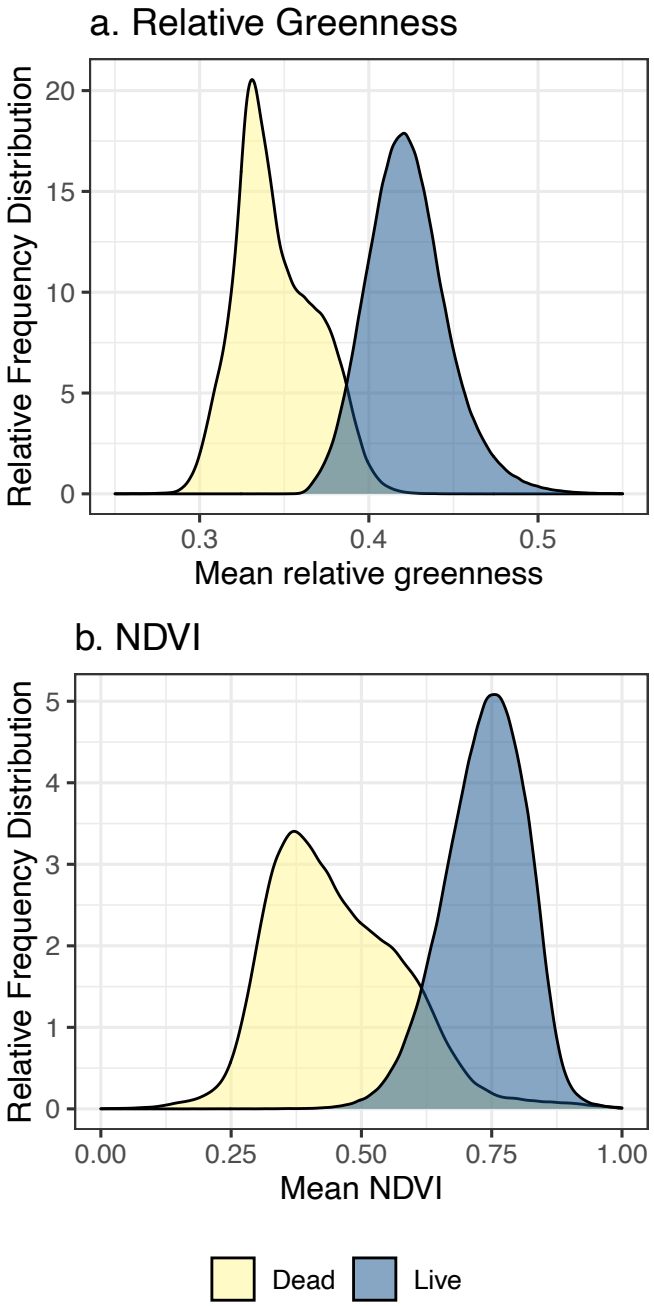

Figure 1.7. The vegetation index distributions of live (dark blue) and dead (light yellow) trees is shown for **a.** mean relative greenness and **b.** mean NDVI during 2017 for all the individual trees in the study domain shown in Figure 1.1.

1.3.2 Biophysical drivers of tree mortality

Next, we consider biophysical drivers of tree mortality within our study region. We explored tree height, canopy cover fraction within 20 meters (excluding an individual tree's canopy within its crown perimeter), distance from the nearest river, trees per hectare, the mean distance of the 10 nearest trees, slope, aspect, and the fraction of granite within 20 meters to investigate the role of tree height and canopy cover, proximity to rivers, tree population density, and topography. We used the tree locations derived from the combined point cloud to compute the tree population density, mean distance of the 10 nearest trees,

distance from rivers, and fraction of granite within 30 meters. For tree height and canopy cover, we used the tree heights and crown perimeters from 2013. Figure 1.8 shows our analysis of each of these feature variables for Soaproot Saddle and Lower Teakettle individually, while the analyses for the combined dataset are shown in Figure A.8. The slopes, y-intercepts, and mortality ranges for each feature variable are shown in Table A.2.

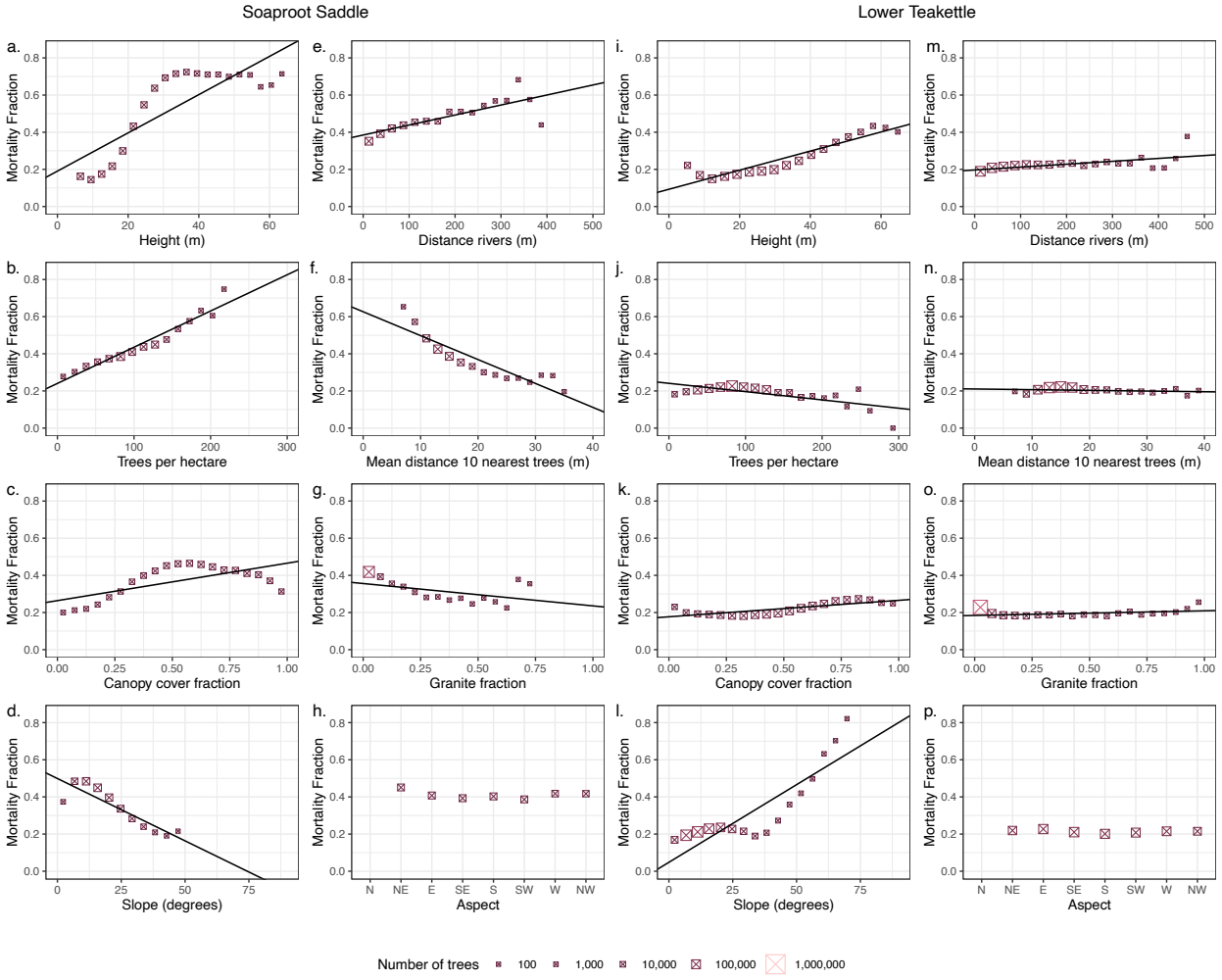

Figure 1.8. We selected bins spanning the domain of each feature variable (the variables on the x-axes of the panel images) for live and dead trees and computed the likelihood of mortality for each bin of each feature variable. The size and color of the points indicate the

order of magnitude of the number of trees in each bin. We computed a trend line through the points for any point representing a bin containing at least 50 trees.

To assess the correlations among feature variables and our target variable, we computed the correlations between each pair of feature variables and the change in mean relative greenness between 2013 and 2017 for each tree (Figure A.9). We chose the change in mean relative greenness within the crown perimeter of the tree to have a continuous variable for the tree's health status rather than a binary 0 or 1 for dead or alive when computing the correlations among variables. We found that tree height has the strongest correlation with likelihood of tree mortality at the site level (Figure A.9) and ranged from a 8.4% increase in mortality probability for every 10-meter increase in height at Lower Teakettle to a 10% increase in mortality likelihood for every 10-meter increase in height at Soaproot Saddle (Table A.2). The drier, lower elevation site (Soaproot Saddle) also had stronger positive relationships between tree mortality and trees per hectare, distance to rivers, and canopy cover fraction than at Lower Teakettle (Table A.2). While tree mortality fraction decreases with increasing slope at Soaproot Saddle, tree mortality fraction increases with increasing slope at Lower Teakettle.

1.3.3 Comparison with previous work

To make a fair comparison with Stovall et al. (2019), we only compare trees from Stovall et al. (2019) that intersect a crown perimeter of one of our trees. If multiple tree top locations intersect our crown perimeter, we chose the one with the closest tree height as

ours. Both studies only consider trees above 5 meters in height. The number of trees we found in the overlapping domain of the two studies, the fraction of those trees we were able to match, and the corresponding mortality fraction of each group is shown in Table 1.3. While the tree top locations relative to the lidar point clouds in the two studies are similar in a high-resolution example (Figure A10), Stovall et al. (2019) found more trees overall. The spatial patterns and differences in tree top locations Soaproot Saddle and Lower Teakettle are shown for representative regions in Figure A11.

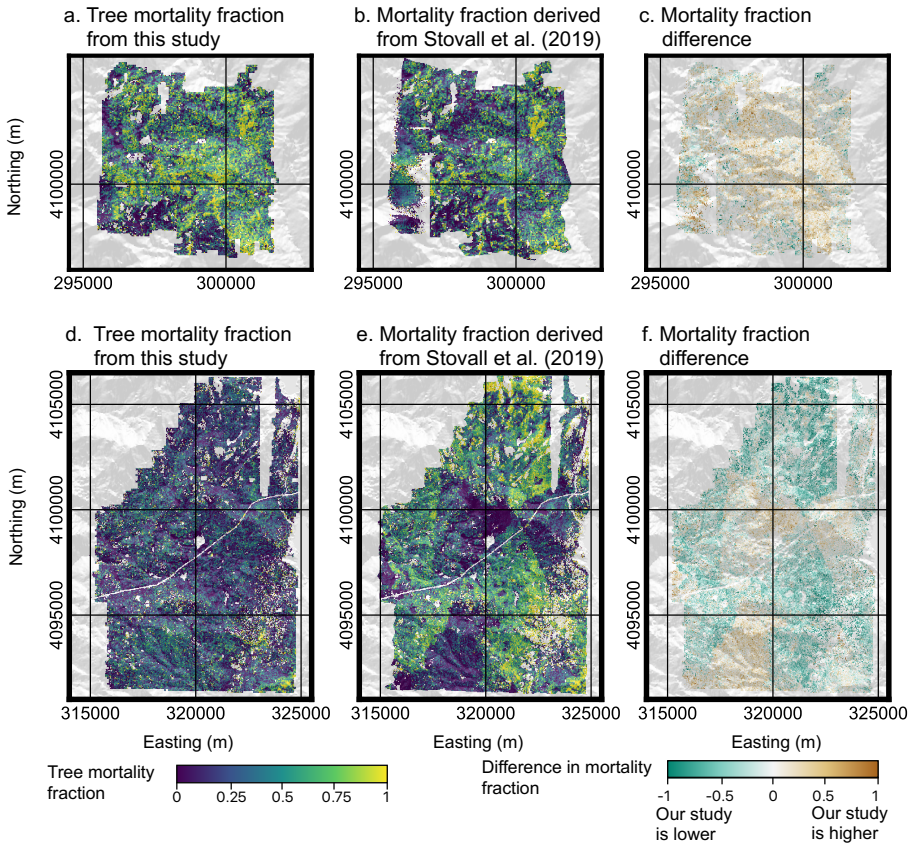

Figure 1.9. Fractional mortality comparison. a. The mortality fraction at Soaproot Saddle for our study, b. for the study by Stovall et al. (2019), and c. the difference between the two. d.-f. The same as a.-c. for Lower Teakettle. All pixels are 30m x 30m Landsat pixels. For our study, the mortality fraction is computed for the year 2017, while mortality fraction is

computed for 2016 for Stovall et al. (2019) due to differences in methods and corresponding data availability. The aspect is overlayed in white (north) to gray (south) to show the underlying topography.

Because different years were studied by Stovall et al. (2019) compared with this study, the closest comparison of tree mortality after the drought is the cumulative tree mortality in 2016 from Stovall et al. (2019) compared to our 2017 mortality. While Stovall et al. (2019) found a cumulative mortality of 41% for 2016 compared with our 32% from 2017, we can compare only the region of overlap. Using the values from Table 1.3, we find that our cumulative tree mortality fraction in 2017 in the overlapping domain of the two studies is 33% vs. 40% for Stovall et al. (2019).

To consider spatial differences in mortality between the two studies, we rasterized the Stovall et al. (2019) dataset to the Landsat grid. Figure 1.9 shows the mortality fraction from our study, that of Stovall et al. (2019), and the difference between the two studies. In the difference maps (Figure 1.9c and 1.9f), brown areas represent pixels where our tree mortality estimates are higher than those of Stovall et al (2019), while green areas represent pixels where our estimates of tree mortality are lower.

Table 1.3

Confusion matrix for Stovall et al. (2019) label comparison

	Labeled dead in Stovall et al.	Labeled live in Stovall et al.
Labeled dead in this study	151,393 trees (21.2%)	68,732 trees (9.6%)
Labeled live in this study	**125,785 (17.6%)**	368,576 trees (51.6%)

Table 1.3 shows the confusion matrix or comparison table of our study labels from 2017 NEON imagery and lidar data and labels from Stovall et al. (2019) based on 2013 lidar data and 2016 NAIP imagery. Our method for matching the trees from our study to the Stovall et al. (2019) study is described in Section 1.2.5. For the 714,486 trees we were able to match, the studies have 72.8% agreement, though this agreement is primarily determined by trees that are alive, the larger of the two groups in the study. Both studies label the same 368,576 trees as live and the same 151,393 trees as dead. Because our study is based on surface reflectance one year later than that of Stovall et al. (2019), the trees that they labeled as live that we labeled as dead are not necessarily incompatible. These trees may have been alive in 2016 and then died by 2017. However, 125,785 of the trees that Stovall et al. (2019) labeled as dead in 2016 were labeled as alive in 2017 in this study. In other words, 45.4% of the matched trees labeled as dead in the Stovall et al. (2019) study are labeled as alive in our study.

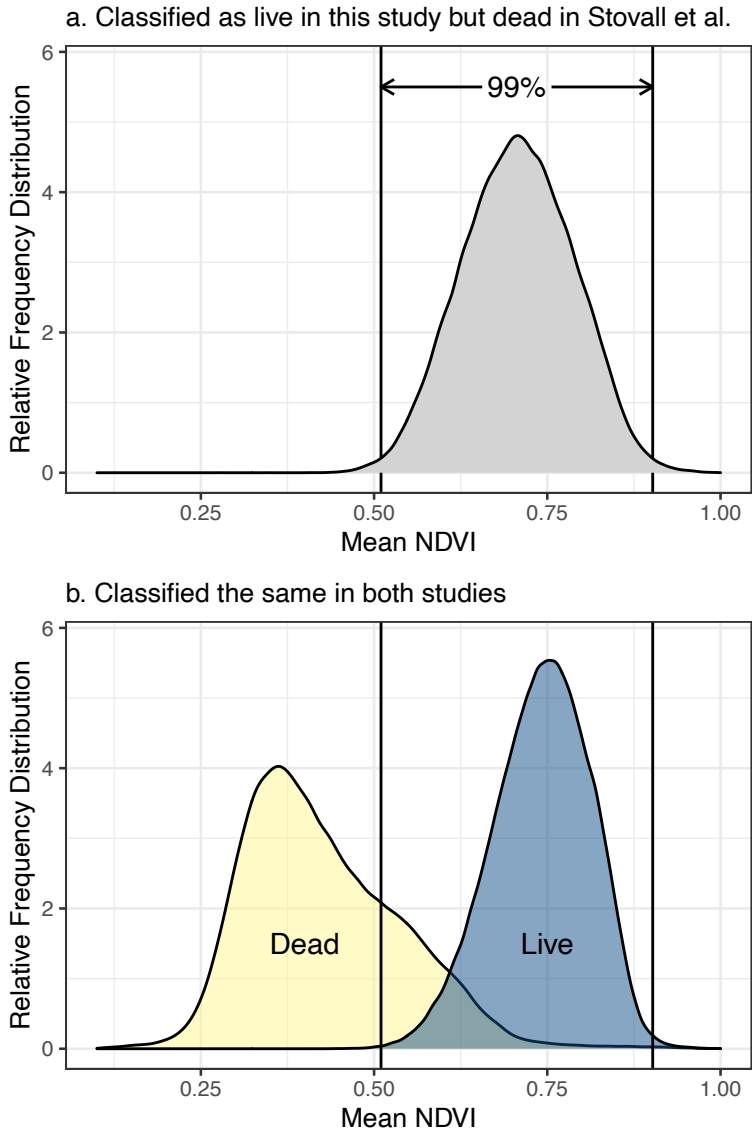

Figure 1.10. Comparison of mean crown area NDVI distributions for trees labeled differently and the same between our study and that of Stovall et al. (2019). The x-axis shows mean NDVI of individual trees, while the y-axis shows the probability density function. **a.** We show the mean NDVI distributions for trees that were identified as live in this study but dead in Stovall et al. (2019). The boundaries for the middle 99% of the mean NDVI values of these trees are shown by the black vertical bars that carry down into the next panel. **b.** The mean NDVI distributions of trees that both our study and Stovall et al. (2019) labeled as live are shown in blue, while the distributions for trees we both labeled dead are shown in light yellow.

To better understand the differences between the two approaches, we compared the mean NDVI within the crown perimeters of trees we labeled as live, but Stovall et al. (2019) labeled as dead (Figure 1.10). The trees where both studies agree are well-

separated by mean NDVI. Overall, the distribution of the mismatched trees shown in Figure 1.10a more closely matches the distribution of the live trees shown in Figure 1.10b.

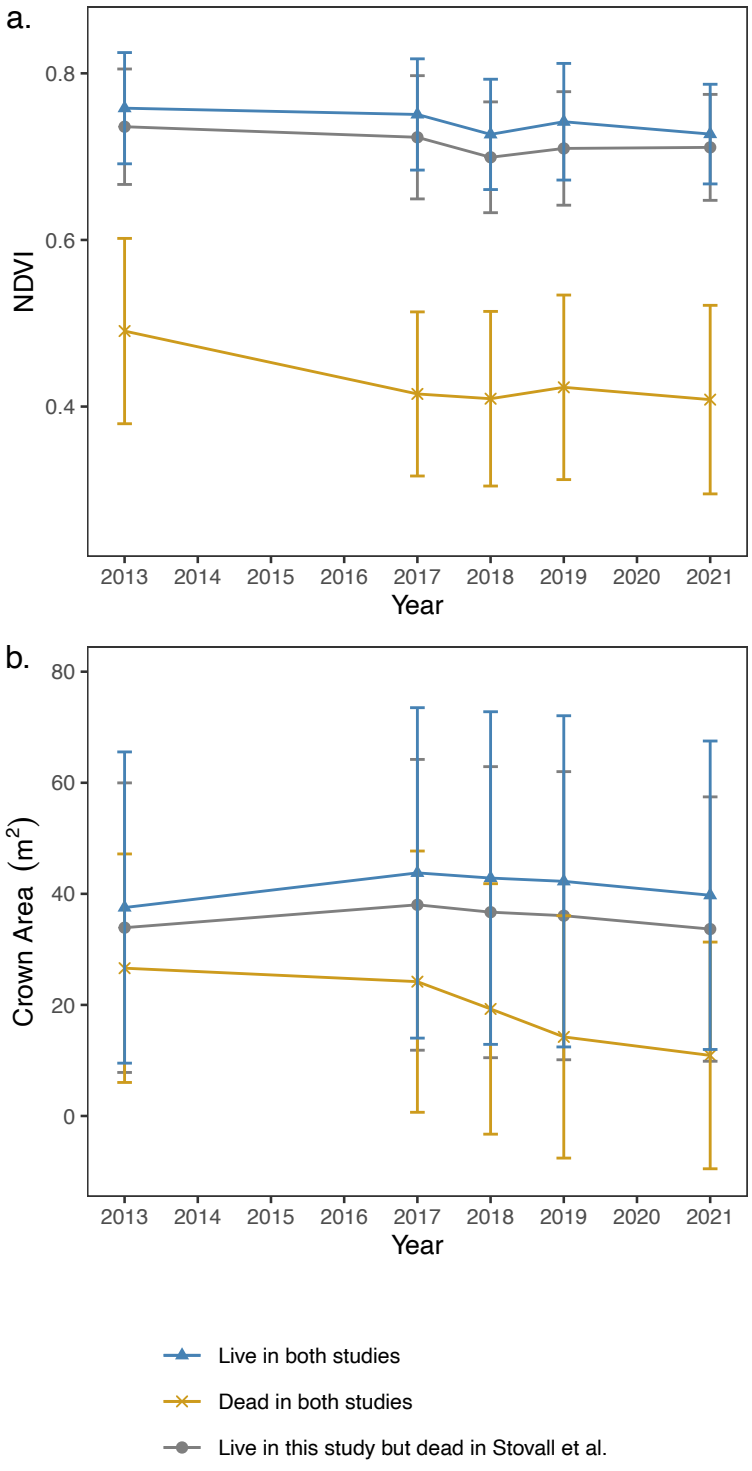

Figure 1.11. Comparison of NDVI and crown area for different sets of labeled trees from Stovall et al. (2019) and the approach developed here. **a.** NDVI time series with the crown perimeter of trees labeled as live in both studies (blue), dead in both studies (yellow), or as live in the current study but dead in Stovall et al. (gray). **b.** The same as for panel a but for crown areas derived solely from NEON lidar measurements.

Because we used the combined point cloud to collect our tree locations, we can track each individual tree over time. We plotted the mean NDVI within the crown perimeters between 2013 and 2021 and the change in the crown area over the same time frame (Figure 1.11). The mean relative greenness for the trees we labeled as live and Stovall et al. labeled dead more closely resemble the trees both studies labeled as live than those labeled dead. In Figure 1.11b, we see higher variance in crown area. However, the trees that both studies agree are dead by 2017 show a decline in crown area, with a crown area of 0 m^2 falling within one standard deviation of the mean. The other two categories (where both studies agree the trees are live in blue, and where only our study labels trees as live in gray) decline more slowly. We would still expect to see some decline from 2013 to 2021, since this group of trees includes trees that die between 2017 and 2021.

We compared the tree mortality between 2013 and 2017 rasterized to the Landsat grid to the change in late summer NDMI over the same years. We used the change between 2013 and 2017 to capture the large tree mortality event at the end of the 2012-2016 drought and due to the years of data availability of the NEON datasets (2013, 2017, 2018, 2019, and 2021). The result is shown in Figure A.12. We found that there is a negative linear association with an R^2 value of 0.36. This indicates that higher tree mortality is associated with a decrease in late summer NDMI consistent with work by Goulden and Bales (2019).

1.4 Discussion

We mapped the spatial patterns of tree mortality under drought conditions in the Sierra Nevada. We loosely followed the methodology charted by Stovall et al. (2019) but made several key changes with respect to the datasets used and the ways in which trees were segmented and labeled. We found evidence that fewer trees may have died by 2016 than previously estimated, although there were broadly similar patterns with higher levels of mortality at the lower elevation Soaproot Saddle. Our lower estimate of cumulative tree mortality (32% vs. 41%) could be due to differences in the total number of trees analyzed (we analyzed half the number of trees) and proportion of low-to-high elevation regions that we considered. For example, trees from the lower elevation site (Soaproot Saddle) were more likely to die, so considering fewer trees from that region will lead to a lower tree mortality estimate. However, within the region where our study overlaps with previous work, our method – which had 93% accuracy on our training dataset – disagreed with 45.4% of the labels for trees labeled dead by Stovall et al. (2019). A comparison of NDVI and crown area time series for this set of mismatches revealed a pattern that was more consistent with the characteristics of live trees identified by both studies.

Differences in spatial patterns from previous research efforts may be due to differences in the preprocessing of the spectral reflectance data. Spectral reflectance data are sensitive to the time of day that the returns are collected. While NEON hyperspectral data is corrected for atmospheric effects and topography (Karpowics & Kampe, 2022), there is not enough information provided with the NAIP imagery to apply this correction. Previous work correcting NAIP imagery using Landsat data revealed significant differences in reflectance indices such as NDVI before and after correction (Zhang et al., 2019). The differences in reflectance in the NAIP imagery at different times of day and different angles

of reflection in the highly variable terrain in the Sierra Nevada mountains may account for many of the differences between the tree mortality map presented in this study and previous results by Stovall et al. (2019).

Consistent with previous work, we found that tree mortality risk increases with height. We found a steep increase in mortality risk with height (about 10% per 10 meters of height) at the lower elevation site (Soaproot Saddle) up to a tree height of 30 meters. Trees taller than 30 meters maintained a 70% likelihood of mortality after the drought. At the higher elevation site (Lower Teakettle), we found a more moderate increase in mortality risk with height (about 8.4% per 10 meters of height). There was much more tree mortality at the lower elevation Soaproot Saddle site, consistent with previous work on drought-associated tree mortality (Byer & Jin, 2017; Stovall et al., 2019).

After the drought, 2.0% of the trees died between 2017 and 2018, and 2.8% died between 2018 and 2019. Legacy effects from the 2012-2016 drought may have kept mortality levels elevated in subsequent years. Most of these additional increases occurred at Lower Teakettle. When we mapped additional tree mortality for each year of data, we found some inconsistencies, particularly between 2019 and 2021 where trees that we labeled as dead in 2019 were later labeled as live in 2021. On closer examination in QGIS, one region is in an area where there is both visible snow and low luminosity in the surface reflectance, potentially due to a late evening or early morning data collection flight. The low luminosity poses a challenge to removing snow-filled pixels by our methods which impose a luminosity threshold. Other areas in the eastern regions of Lower Teakettle appear brown in 2019 but look green and healthy in spectral imagery from July 2021. Because our

algorithm only looks at the spectral signature of the trees for a given year, we would not expect these trees to be accurately classified without additional processing.

Because biophysical drivers of tree mortality vary significantly by species (Stephenson & Das, 2020), improvements in species classification models from high-resolution spectral and lidar data may help to create accurate maps of tree mortality risk. Future directions for this work include using the individual tree mortality dataset to build tree mortality models that incorporate non-linear relationships and interactions among biophysical features. Raster data provided with this study may be helpful to scientists working to improve Landsat-based estimates of tree mortality. In addition, these estimates of the impact of drought and subsequent wildfire may assist land managers in forest conservation efforts and may provide helpful parameters for dynamic vegetation models that include these effects.

1.5 Conclusion

Our study provides an open-source methodological framework for estimating tree mortality from high-resolution lidar and spectral data that can be applied to other forests using hand-labeled training data and parameters specific to the trees in new study sites. We estimate that 49.6% of trees were dead by the end of the 2012-2016 California drought at Soaproot Saddle and 25.5% were dead at Lower Teakettle. While tree mortality at Soaproot Saddle appears to saturate in 2017, we found that trees at higher elevations continue to die at a rate of about 2% per year after the drought subsided. Moreover, we found that subsequent wildfire at Soaproot Saddle increased tree mortality fraction by 49-

89% within burned areas. The datasets provided here may help to constrain dynamic vegetation models to improve our understanding of forest disturbance now and in the future.

1.6 Data Availability Statement

The scripts for the methods and datasets we created for this study are archived at Zenodo at the following website: https://doi.org/10.5281/zenodo.7812035 (Hemming-Schroeder et al., 2023a). Our algorithm, analyses, and several figures were created using R Statistical Software, version 4.1.2 (R Core Team, 2023). We completed the hand-labeling of training data and additional figures using QGIS, version 3.18.2-Zürich (QGIS Development Team, 2023). We ran the R scripts for the algorithm on the High Performance Community Computing Cluster (HPC3) at the University of California, Irvine. The batch scripts for running the R scripts on a high performance computing cluster are written for a SLURM scheduler.

CHAPTER 2

Retrospectively modeling individual tree mortality risk after a severe drought in California

Adapted from an unpublished manuscript with permission from coauthors Carl Norlen, Markelle Kelly, Padhraic Smyth, Steven Allison, and James Randerson

2.1 Introduction

Our climate is warming with increasingly frequent and severe droughts in the western United States over the last two decades (Crockett & Westerling, 2018). Prolonged warm and dry droughts are associated with large tree mortality events in temperate conifer forests (Adams et al., 2009; Allen et al., 2010). These large tree mortality events are caused by a combination of increasing hydraulic stress and growing bark beetle populations during severe droughts (Anderegg et al., 2015b). During the 2012-2016 drought in California, soil water was depleted to a depth of 5-15 m and an estimated 129 million trees died (USDA, 2017). In hotter and drier low elevation forests, about half of the trees died with higher levels of mortality observed in areas farther away from streams (Hemming-Schroeder et al., 2023b; Paz-Kagan et al., 2017) and for taller trees (Hemming-Schroeder et al., 2023b; Paz-Kagan et al., 2017; Stovall et al., 2019).

A key driver of tree mortality among taller trees at low elevations is preferential selection of large trees by bark beetles (Fettig et al., 2019; Koontz et al., 2021; Stephenson et al., 2019). One study showed that combined effects of drought and bark beetles killed about 48% of sugar pine (*Pinus lambertiana*) and 90% of Ponderosa pine (*Pinus ponderosa*) during the 2012-2016 drought (Fettig et al., 2019). Tree mortality among Ponderosa pine during the drought is primarily attributed to western pine beetle (Robbins et al., 2021). Under baseline conditions, background tree mortality in western forests of the United States is primarily associated with endemic bark beetle attack (Das et al., 2016). Initially, western pine beetle populations begin to grow beyond endemic numbers as warm and dry drought conditions stress smaller host trees and allow more beetles to survive milder

winters (Robbins et al., 2021). As bark beetle populations grow over a drought period, they can coordinate attacks on larger trees (Raffa et al., 2008). The large loss of Ponderosa pine during this last drought may prevent a similarly large tree mortality event from occurring again until the host population recovers which may not happen until much later this century (Robbins et al., 2023).

Another mechanism that may reduce the impact of future droughts is mechanical thinning. Mechanical thinning of understory and overstory trees at the Teakettle Experimental Forest fifteen years prior to the 2012-2016 drought increased tree growth rates before and during the drought (Zald et al., 2022). Similarly, a study by Keen et al. (2023) found evidence that mechanical thinning may protect trees that would otherwise be vulnerable to drought and bark beetles. This may be an important explanatory mechanism behind a finding by Norlen and Goulden (2023) that die-off during a drought may reduce mortality in subsequent droughts in semi-arid conifer forests like the Sierra Nevada.

While mechanical thinning or smaller-scale drought-induced tree mortality may help protect a forest against future drought, large-scale die-off and wildfire, particularly in lower elevation regions of the Sierra Nevada, may drive shifts to new biomes as the climate continues to change. A recent study by Hill et al. (2023) found that almost 20% of the Sierra Nevada's vegetation may no longer match the current climate conditions. Moreover, most of the area with the mismatch is located under 2400 m (Hill et al., 2023). Since the elevation range of Ponderosa pine is from 400 m to 2200 m (Robbins et al., 2023), stands with Ponderosa pine are particularly vulnerable to drought and western pine beetles in the Sierra Nevada. In the years following the drought, forested areas of the Sierra Nevada dominated by Ponderosa pine had the highest number of snags (Vilanova et al., 2023).

Previous studies investigating drivers of tree mortality in the Sierra Nevada have found that tree mortality risk increases with tree height (Hemming-Schroeder et al., 2023b; Stovall et al., 2019). While there is evidence to suggest that taller trees may experience vascular damage from the hydraulic demands of drought conditions (Anderegg et al., 2015a), the correlation between tree height and mortality may be driven by preferential selection of larger host trees by western pine beetles (Restaino et al., 2019). A study by Stephenson and Das (2020) showed that among just under 6000 trees in Sequoia National Park, the only species to show an increasing trend of mortality with height were trees of the *Pinus* genus. They found that the relationship was so strong among the Pinus genus of trees which represented only 10% of trees in their study region, that it overshadowed the weaker trend of declining mortality with increasing height among other tree species. In contrast, Restaino et al. (2019) found a positive association between basal area and tree mortality for incense cedar and white fir.

In addition to tree size, topographic features have been associated with tree mortality risk. Paz-Kagan et al. (2017) conducted a study investigating tree mortality in Sequoia National Forest after the 2012-2016 drought and found elevated tree mortality at low elevations, shallower slopes, and greater distance from rivers. In addition, they found that western and southwestern aspects were associated with higher tree mortality (Paz-Kagan et al., 2017). In a related study to the south of Sequoia National Forest, Hemming-Schroeder et al. (2023b) also found a positive relationship between tree mortality and shallower slopes as well as distance from rivers for trees at low elevations. However, trees at high elevations showed the opposite relationship with slope and a much weaker relationship with distance from rivers (Hemming-Schroeder et al., 2023b). Neither site

showed a relationship between tree mortality and aspect (Hemming-Schroeder et al., 2023b).

Models struggle to predict tree mortality in advance (Trugman et al., 2021). Process-based models need to account for many factors such as demography, hydraulic stress, and interactions with bark beetles which introduce high levels of model complexity and uncertainty (Trugman, 2021). Machine learning models may offer an alternative strategy for modeling challenging problems like tree mortality risk during an extreme drought. For example, Paz-Kagan et al. (2017) used support vector machines to estimate tree species and tree mortality for 2-meter hyperspectral reflectance and canopy height model pixels collected in 2015 at Sequoia National Park. Following this effort, they used random forests with height, soil type, topography, and fire history feature data to estimate tree mortality likelihood and found that elevation, distance from rivers, and rock cover were the top three predictor variables (Paz-Kagan et al., 2017).

More recently, Koontz et al. (2021) segmented lidar point clouds and images from drones taken in 2018 before aggregating spectral information from pixels within individual tree crowns to estimate tree species and mortality. Because the pixel information is aggregated for individual trees, this may allow for more independent examples from the individual trees for training and validation in any modeling efforts. Several segmentation algorithms for delineating the crowns of individual trees are explored and compared to ground data to select the best one (Koontz et al., 2021). However, because the data is collected in 2018 after many trees have died, it can be difficult to accurately collect traits for dead trees (Koontz et al., 2021) or account for fallen dead trees.

In our study, we use a data set of more than one million trees collected from high-resolution aircraft lidar and multispectral data between 2013 and 2021 to analyze risk factors for tree mortality in the Southern Sierra Nevada in response to the 2012-2016 California drought. Because the earliest data set that we use is in 2013, we can collect the location and height of trees that died after 2013 and may also be able to capture tree mortality prior to 2013 before trees lose their branches and become more difficult to detect by lidar. We use extreme gradient boosting to model individual tree mortality risk. Because extreme gradient boosting is a tree-based method like random forests, it can capture non-linear relationships between the feature variables and target variable. While random forests use a popular vote of a given number of trees to make a classification, extreme gradient boosting uses gradient descent to learn the model's parameters. In general, gradient descent refers to moving in the direction within the parameter space that minimizes your loss function. Because of this, extreme gradient boosting may be more accurate than other decision tree methods. For our feature variables, we only use information from vegetation traits, topography, climate variables prior to the drought and the perturbation in precipitation and temperature during the drought. The question that guides our investigation is the following: To what extent can we model individual tree mortality risk using information available prior to observed die-off events? Tree mortality risk models in regions vulnerable to biome shifts may help inform conservation efforts as these forests respond to a climate increasingly characterized by heat waves and drought.

2.2 Methods

2.2.1 Study site and tree mortality data

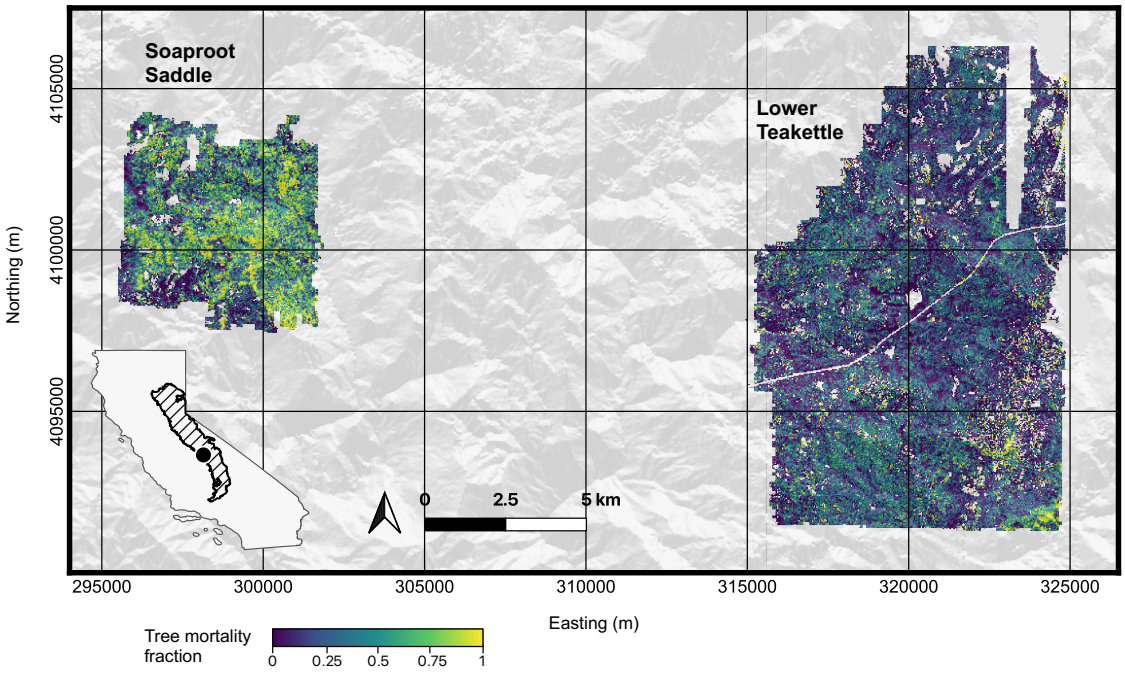

Figure 2.1. The 2017 tree mortality fraction is shown for Soaproot Saddle (to the west) and Lower Teakettle (to the east) on the 30-meter resolution Landsat grid. Dark blue represents no tree mortality, whereas yellow represents 100% tree mortality in a pixel. The semi-transparent topography shows the north-to-south (white-to-gray) aspect calculated from the National Aeronautics and Space Administration's (NASA) Shuttle Range and Topography Mission (SRTM) data set. The inset shows the outline of state of California from the Topologically Integrated Geographic Encoding and Referencing (TIGER) database with the Sierra Nevada Level 3 Ecoregion (hashed polygon), and the location of the study sites marked with a black point. This figure is adapted from the information shown in Figures 1 and 6 in Hemming-Schroeder et al. (2023b).

For this study, we used a data set of more than one million trees classified as dead or alive in 2017 from Hemming-Schroeder et al. (2023b). This tree mortality data set is derived from the National Ecological Observatory Network's (NEON) airborne observation platform's lidar and hyperspectral data sets at the Soaproot Saddle and Lower Teakettle sites in the Sierra National Forest of California shown in Figure 2.1. Soaproot Saddle has a lower elevation between 1000 m and 1400 m with warmer temperatures and precipitation of about 900 mm per year (Krauss, 2018). The most prominent tree species at Soaproot Saddle include Ponderosa pine (*Pinus ponderosa*), incense cedar (*Calocedrus decurrens*), canyon live oak (*Quercus chrysolepis*), and California black oak (*Quercus kelloggii*) (Krauss, 2018). Consistent with previous studies, NEON reported high mortality for Ponderosa pine due to bark beetle infestation during the drought (Krauss, 2018). The site had an estimated cumulative mortality of about 50% by 2017 after the 2012-2016 drought (Hemming-Schroeder et al., 2023b). Meanwhile, Lower Teakettle is at higher elevation (2000 m to 2800 m) with higher precipitation (1220 mm per year) and proportion of precipitation falling and accumulating as snow (Krauss, 2018). Lower Teakettle is also characterized by cooler temperatures and had an estimated 26% cumulative tree mortality by 2017 after the 2012-2016 drought (Hemming-Schroeder et al., 2023b). The most prominent tree species at Lower Teakettle include red fir (*Abies magnifica*), white fir (*Abies concolor*), Jeffrey pine (*Pinus jeffreyi*), and lodgepole pine (*Pinus contorta*) (Krauss, 2018). About a quarter of the trees (275,103) in the individual tree mortality data set are from Soaproot Saddle, while the remaining three quarters (736,474 trees) are from Lower Teakettle.

2.2.2 Feature Variables

2.2.2.1 Vegetation traits

Trees in the individual tree data set were segmented by combining all lidar data available at the time of publication (from flight campaigns in 2013, 2017, 2018, 2019, and 2021) to identify individual trees. Each lidar point has an x-, y-, and z-coordinate. Hemming-Schroeder et al. (2023b) appended a year and used the entire data set to identify the tree top locations using the algorithm of Dalponte and Coomes (2016). In a following step, the data frame for each lidar point cloud was filtered by year to segment the tree with a given treetop to obtain its approximate point cloud for each year. Then, concave polygons were drawn around the x- and y-locations of the segmented point clouds to draw a crown perimeter for each tree for each year of data. Hemming-Schroeder et al. (2023) created a canopy height model for each portion of data in the study region (the data was processed using a 96 m by 96 m window with a 24 m buffer). The canopy height model generated from the lidar point cloud was used to assign the height of each tree for each year of available data. The data were filtered to include only trees that had crown information for all five years of flight campaigns to allow for an assessment of tree status based on changes in lidar-derived crown area and surface reflectance for each year of data. A total of 1,011,577 trees are in the data set. More detail is provided in Hemming-Schroeder et al. (2023b).

For each tree location, the canopy fraction is derived from filtering the canopy height model from NEON (2022a) to exclude values under 5 meters and then computing the area of the remaining canopy within a 20 m radius of each individual tree. The crown

perimeter of the individual tree was removed from the circular buffer to exclude the tree's own canopy from the canopy cover fraction of its neighborhood. The trees per hectare variable is calculated from all the trees greater than 5 meters in height that were detected using the combined lidar point cloud.

2.2.2.2 Vegetation indices

We downloaded Landsat Collection 2 Level 2 Tier 1 surface reflectance Landsat data from the United States Geological Survey (USGS) Earth Explorer website for path 42 and row 34. We computed the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Moisture Index (NDMI) for the four Landsat scenes temporally nearest September 15, 2011, to obtain initial vegetation conditions in late summer before the 2012-2016 drought. NDVI is computed by taking the difference of Landsat's near infrared and red bands and dividing by their sum. The domain of NDVI is between -1 and 1, inclusive, where 1 represents healthier vegetation with high reflectance in the near infrared and low reflectance in the red band. Similarly, NDMI is computed by taking the difference of the near infrared and short infrared 1 bands and diving by their sum. The domain of NDMI is also between -1 and 1, inclusive, where 1 represents low reflectance in the shortwave infrared 1 band which corresponds to high moisture.

2.2.2.3 Evapotranspiration

We generated 30-m gridded evapotranspiration (ET) data for each water-year. The water-year begins in October of the previous calendar year and goes through the end of September of the given water-year. For example, the 2009 water-year runs from October 2008 through September 2009. To create the gridded ET product, we collected annual Landsat NDVI retrieved from Google Earth Engine (Gorelick et al., 2017) for each water-year and annual ET from a network of 10 eddy covariance flux towers distributed across California (retrieved from https://www.ess.uci.edu/~california/). For each of the 10 eddy covariance sites, we extracted annual NDVI values for 9 upwind Landsat pixels. We then created a regression between annual NDVI and annual ET which we used to extrapolate ET across the landscape and through time (R^2 = 0.692). This method for generating gridded ET from NDVI was originally described in Goulden (2012) and revised in Norlen and Goulden (2023).

2.2.2.4 Granite fraction

We computed the granite fraction within 20 meters of each individual treetop. First, we collected the NEON canopy height model raster data and RGB raster data derived from NEON's hyperspectral data (both at 1 meter resolution). We assigned a pixel as granite if the canopy height was less than 0.5 m and the luminosity of the pixel was greater than 0.20. For each individual tree, we generated a 20-m circular buffer. The granite fraction within each buffer was used for the granite fraction variable.

2.2.2.5 Distance from rivers

We downloaded the High-Resolution National Hydrography Dataset from U.S. Geological Survey (2019) for region 1803 and computed the distance between each individual tree and the nearest river.

2.2.2.6 Topography

The aspect and slope variables were computed from the NEON digital terrain model (1 m resolution). We used the *raster* package in R to compute the aspect and the slope from the digital terrain raster data. To address noise, we smoothed the aspect and slope calculations as follows. First, we generated 10-m circular buffers around each treetop location. Then, we calculated the median aspect and slope within each 10-m buffer. Since aspect begins at 0 radians pointing north and ends again at North at 2π radians, we transformed the aspect to create an index (1) from north to south, A_{NS}, and (2) from east to west, A_{EW}. We computed A_{NS} and A_{EW} using the following equations where x is the aspect in radians:

$$A_{NS} = 0.5 \cos x + 0.5$$
$$A_{EW} = 0.5 \sin x + 0.5 \quad (2.1)$$

2.2.2.7 Climate Variables

We downloaded Daymet data which is a monthly-averaged 1-km resolution climate data product. We computed mean annual precipitation before (2009-2011) and percent change during the drought (2012-2016) by water-year. Next, we computed the mean dry season temperature before and difference during the drought using the same water-years as for mean annual precipitation.

Finally, we reasoned that warmer winter temperatures may melt the winter snowpack, decreasing the available water during the dry season. In addition, warmer winter temperatures may allow for more endemic bark beetles to survive the winter, increasing their populations. We calculated the minimum winter temperature during and difference during the drought to account for these effects.

2.2.3 Model Setup

2.2.3.1 Target variable

Ideally, we would choose whether a tree died during the drought between 2012 and 2016 to capture drought-induced mortality. However, our data set begins in 2013 after the drought has started, and the 2013 hyperspectral dataset is incomplete for a large portion of Lower Teakettle. Our next available time point is in 2017 which has complete data for both sites. Because of this, we use whether a tree is dead in 2017 as our target variable. The implication is that we assume that tree mortality prior to the drought is low enough to ignore.

2.2.3.2 Feature variables

To model tree mortality risk, we use 6 metrics of vegetation traits, 5 metrics corresponding to topography, and 6 metrics of climate data (Table 2.1). For the combined dataset (trees from both Soaproot Saddle and Lower Teakettle), we computed the 2nd- and 98th-percentile of each feature variable. Those represent the lower and upper limits of the scalebar in Figures 2.2 and 2.3 which show the mapped feature variables for Soaproot Saddle and Lower Teakettle, respectively. For example, panel (a) showing tree height for Soaproot Saddle in Figure 2.2 and Lower Teakettle in Figure 2.3 uses the same scale bar. We calculated the lower boundary and upper boundary of the domain for each variable using the 2nd and 98th percentile of each feature from the entire dataset. These percentiles for each site and the combined dataset are shown in Table B.1. The original resolution of the feature variables varies, so we aggregated each one from the scale of the individual tree to the Landsat coordinate reference system and resolution (30 meters) for visualization.

Table 2.1

The feature variables used to model tree mortality risk in Chapter 2.

Feature variable	Resolution	Source	Ecological rationale
Tree height (2013)	< 1 m	Derived from NEON lidar	Beetles select for larger trees later in the drought. Larger trees have higher water demands.

Canopy cover within 20 meters (2013)	1 m	Derived from NEON canopy height models	Higher canopy cover may offer increased latent cooling.
Trees per hectare (2013)	30 m	Derived from NEON lidar	Closer trees may compete, and high tree population density may increase risk of bark beetle infestation.
NDVI before the drought (2011)	30 m	Derived from Landsat	Healthier vegetation prior to the drought may be more likely to survive.
NDMI before the drought (2011)	30 m	Derived from Landsat	Vegetation with more leaf moisture and more moisture in the surrounding soil may be more sensitive to water deficits.
Evapotranspiration before the drought (2009-2011)	30 m	Derived from Landsat by Norlen and Goulden (2023)	Evapotranspiration prior to the drought may indicate the moisture overdraft during water deficits.
Granite fraction within 20 m	1 m	Derived from NEON canopy height models and hyperspectral data	Granite outcrops may indicate poor soil quality and higher sensible heating.
Distance to rivers (m)	< 1 m	Derived from NEON lidar and the National Hydrography Dataset	Trees further from rivers may be more water-stressed but may also be more adapted to drier conditions.
Slope	1 m	Derived from NEON lidar digital terrain model	Steep slopes may be associated with relatively cooler conditions at low elevations but poorer soil quality and higher runoff.
North-south aspect	1 m	Derived from NEON lidar digital terrain model	Southern slopes receive more solar radiation during the day and may become more water stressed. However, vegetation on southern slopes may also be adapted for drier conditions.
East-west aspect	1 m	Derived from NEON lidar digital terrain model	Because of the westerlies, storm systems moving over the Sierra Nevada typically move from west to east. As air parcels are pushed up

			mountains, they condense and rain on western slopes leaving a rain shadow (or dry conditions) on eastern slopes.
Annual mean precipitation before and fraction of baseline annual mean precipitation during the drought	1 km	Daymet	The typical amount of water inputs to the system may drive species distributions. The magnitude of the drought perturbation may affect areas differently depending on the baseline water inputs.
Dry season (summer and fall) mean temperature before and difference during the drought	1 km	Daymet	Dry season temperatures exacerbate water stress during drought conditions.
Winter mean monthly minimum temperature before and difference during the drought	1 km	Daymet	Cooler winter temperatures lead to increased fraction of precipitation that is snow and increases the longevity of the winter snowpack which provides dry season moisture. Warmer winter temperatures allow for greater numbers of bark beetles to survive.

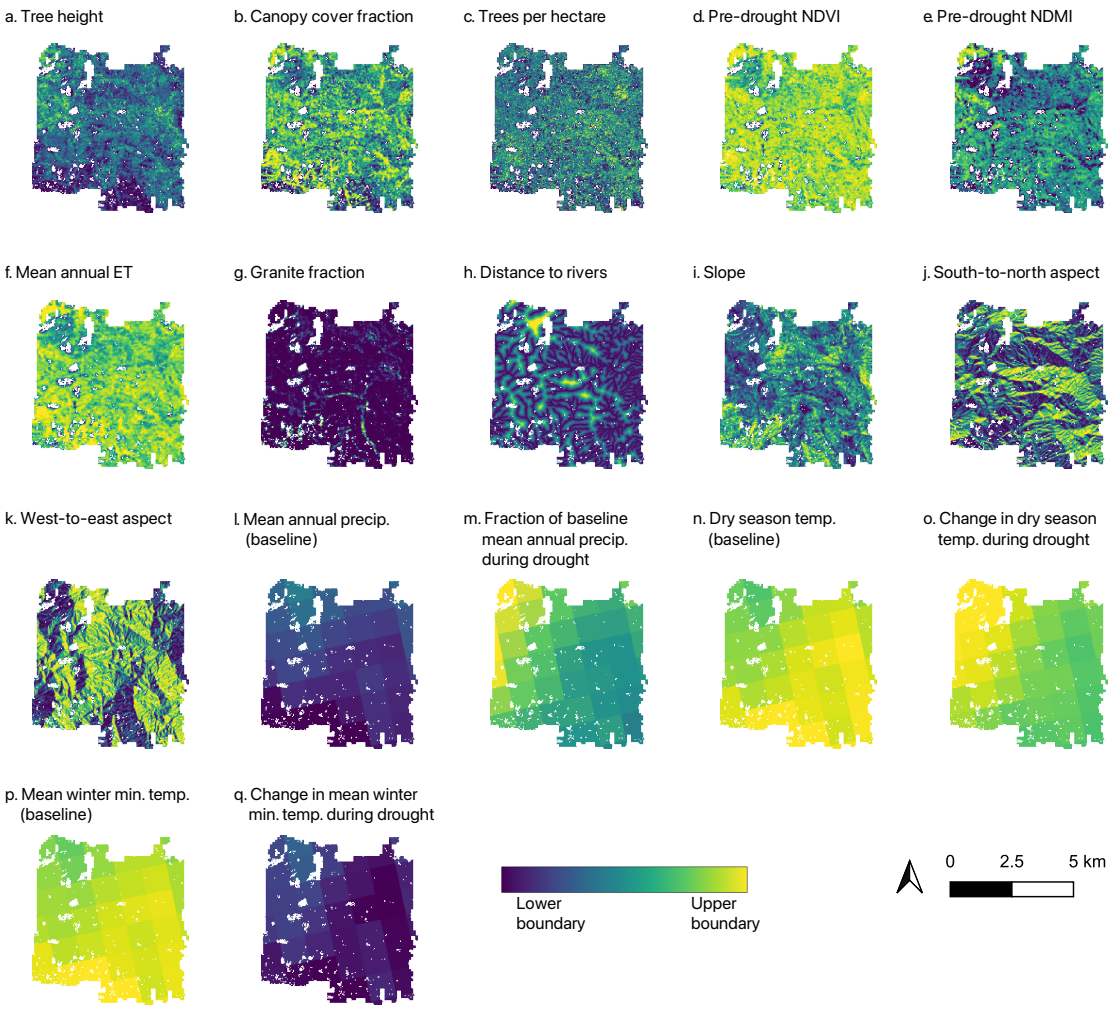

Figure 2.2. Feature variables for Soaproot Saddle. We show the 17 feature variables used in our model to predict tree mortality risk. We computed the 2nd- and 98th-percentiles of each variable from the full feature data set including both Soaproot Saddle and Lower Teakettle at the level of the individual tree to choose the lower and upper boundary cut-offs.

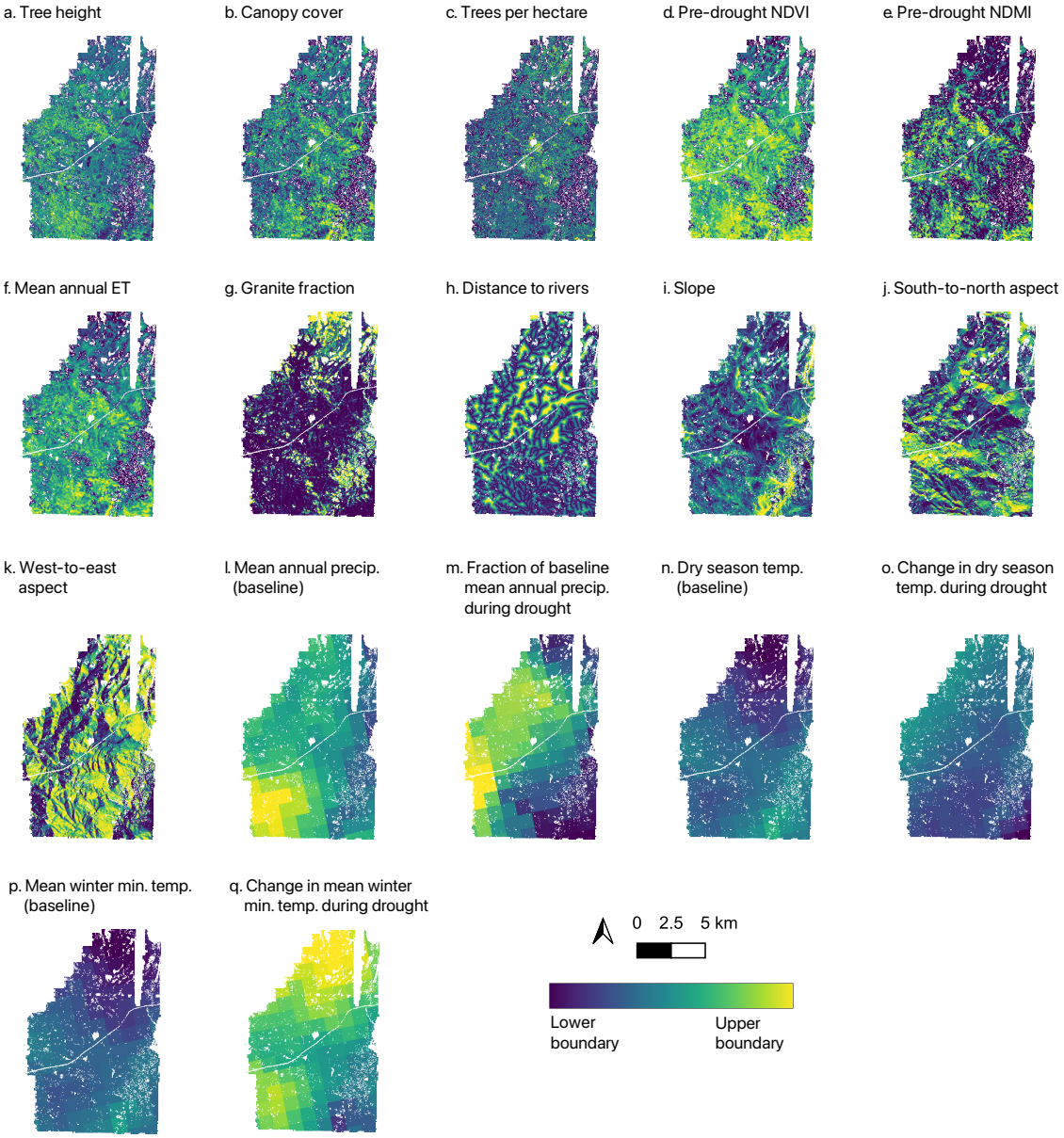

Figure 2.3. Feature variables for Lower Teakettle. We show the 17 feature variables used in our model to predict tree mortality risk. We computed the 2nd- and 98th-percentiles of each variable from the full feature data set including both Soaproot Saddle and Lower Teakettle at the level of the individual tree to choose the lower and upper boundary cut-offs.

2.2.3.3 Training, validation, and test data

We set a seed for reproducibility and randomly split our data set into an 60/20/20 training, validation, and testing data sets. Some trees in the test data set may be near trees in the training and validation data sets and may yield overly optimistic results on the test data due to spatial autocorrelation. Next, because the classes are not evenly represented in the original data set, a random split yields a training data set with uneven live and dead classes. Uneven training classes may lead to a bias in the model whereby the model favors the larger class and performs poorly on the smaller class. Because of the imbalance in the live and dead classes (about 68% compared to 32%), a model may overfit to the live class if the training and validation data are not resampled. To explore potential biases from class imbalances and their remedies, we completed two resampling strategies including (1) no resampling and (2) resampling to equalize the number of live and dead trees. To weight each example in each class as evenly as possible, we replicated the class examples as many times as needed before randomly sampling for the remaining amount under 100%. For example, if a class needed to be increased by 150%, we replicated the data set once (for the 100%) and then randomly sampled without replacement for the remaining 50%.

We completed both resampling schemes at three levels of data inclusion: (A) the full data set, (B) data from Soaproot Saddle only, and (C) data from Lower Teakettle only. The two resampling strategies across three levels of site inclusion make for a total of six resampled data sets. The number of trees without resampling and with resampling for even classes are shown in Table 2.2 for each site and the combined dataset. The original split at Soaproot Saddle is already approximately equal in number for both classes. At Lower

Teakettle, the number of trees in the dead class is about one quarter of the total number of trees. In the combined dataset, about one third of the trees are dead. The numbers in parentheses indicate the class fraction.

Table 2.2

The total number of trees in the training data set for each resampling strategy and site

	No resampling			Resampling live and dead classes only		
Class	Dead	Live	Total	Dead	Live	Total
Both sites	194,052 (0.32)	409,675 (0.68)	603,727	409,675 (0.50)	409,675 (0.50)	819,350
Soaproot Saddle only	81,508 (0.50)	81,051 (0.50)	162,559	81,508 (0.50)	81,508 (0.50)	163,016
Lower Teakettle only	112,544 (0.26)	328,624 (0.74)	441,168	328,624 (0.50)	328,624 (0.50)	657,248

2.2.3.4 Null models

To help assess model skill, we introduce two null models. In the first null model (Null Model I), we guess that all the trees belong to the larger class (either live or dead). In the second model, we guess that trees belong to this larger class with probability p corresponding to the fraction of the total represented by the larger class. Since the same reasoning holds if the classes are equal, the resulting null model class and overall accuracies apply to any of the training datasets shown in Table 2.2.

Let $0.5 < p \leq 1$ be the fraction of the larger class which means $1 - p < 0.5$ is the fraction of the smaller class. Let the total number of trees in both classes be N. Then, the larger class has Np trees and the smaller class has $N(1 - p)$. In Null Model I, we consider a

case where we guess that all the trees belong to the larger of the two classes. The theoretical confusion matrix is shown in Table 2.3. We choose the live class to be in the position of the larger class because the number of trees in the live class is equal or greater than that of the dead class in this study.

Table 2.3

Theoretical confusion matrix and class accuracies for guessing all examples belong to the larger class

	Modeled dead	Modeled live	Class accuracy
Labeled dead	0	$N(1-p)$	$\frac{0}{0+N(1-p)}=0$
Labeled live	0	Np	$\frac{Np}{0+Np}=1$

The class accuracy for the dead class is 0 or 0% while the class accuracy for the live class (the class that is equal or greater in number) is 1 or 100%. The overall accuracy of this model is represented by

$$\frac{0 \cdot N(1-p)+1 \cdot Np}{N}=p \tag{2.2}$$

If we guess all the trees are live (the class equal or greater in number), we can expect that the class accuracy for dead trees will be 0% and the class accuracy for the live trees will be

100%. This yields an overall accuracy of $p \cdot 100\%$. By the same reasoning, this argument also holds if $p = 0.5$ and therefore $1 - p = 0.5$ and the classes are equal.

Next, we introduce Null Model II. In Null Model II, we guess a tree belongs to the equal or larger class $p \cdot 100\%$ of the time. Table 2.4 shows the confusion matrix and corresponding class accuracy for Null Model II.

Table 2.4

Theoretical confusion matrix and class accuracies for guessing examples belong to the equal or larger class with probability p

	Modeled dead	Modeled live	Class accuracy
Labeled dead	$N(1-p)(1-p)$	$N(1-p)p$	$\frac{N(1-p)(1-p)}{N(1-p)(1-p)+N(1-p)p} = 1-p$
Labeled live	$Np(1-p)$	$Np \cdot p$	$\frac{Np \cdot p}{Np(1-p)+Np \cdot p} = p$

The theoretical class accuracy for dead trees would be $1 - p$ and for live trees, p. The overall accuracy expected in Null Model II is given by the following:

$$\frac{N(1-p)(1-p) + Np \cdot p}{N} = (1-p)^2 + p^2 \tag{2.3}$$

If the classes are balanced so that $p = 0.5$, the class accuracies would both be 0.50 and the overall accuracy would also be 0.50.

Using the right-hand side of Equations (2.2) and (2.3), we show the theoretical class and overall accuracies that would be achieved from Null Model I and Null Model II. Our models need to outperform these null models to be considered better than only guessing the majority class or randomly guessing based on the fraction of dead trees.

Table 2.5

The class accuracies and total accuracy for Null Model I and Null Model II

	Null Model I: Guess all trees belong to the live class (the equal or larger of the two classes)			Null Model II: Guess a tree belongs to the live class with probability p where p is the fraction corresponding to the live class		
	Class accuracy for dead trees	Class accuracy for live trees	Total accuracy	Class accuracy for dead trees	Class accuracy for live trees	Total accuracy
Both sites, no resampling	0.00	1.00	0.68	0.32	0.68	0.56
Soaproot Saddle, no resampling	0.00	1.00	0.50	0.50	0.50	0.50
Lower Teakettle, no resampling	0.00	1.00	0.74	0.26	0.74	0.62
Both sites, even classes	0.00	1.00	0.50	0.50	0.50	0.50
Soaproot Saddle, even classes	0.00	1.00	0.50	0.50	0.50	0.50
Lower Teakettle, even classes	0.00	1.00	0.50	0.50	0.50	0.50

2.2.3.5 Model and hyperparameters

We used the extreme gradient boosting package (xgboost) and R, version 4.1.2, to model tree mortality as a binary variable from our feature variables. Extreme gradient boosting is a method first described by Chen and Guestrin (2015). Subsequent decision trees are used in an iterative process to fit the model using a loss function (Chen & Guestrin, 2015). Different activation functions can be used for the final classification. We used the softmax function which gives an estimated likelihood of an example belonging to each class. The class likelihoods are nonnegative and must sum to one.

There are several model hyperparameters to select when running an extreme gradient boosting model. We held the learning rate constant at 0.01 and the number of rounds constant at 1000 for all the models. For the remaining hyperparameters, we completed a grid search and varied the following:

a. Maximum tree depth from a value of 1 through 12, increasing by 1,
b. The number of features sampled in each step of the model (0.6, 0.8, 1.0),
c. The fraction of examples from the training set that were used in each step of the model (0.4, 0.6, and 0.8),
d. The value of gamma (minimum loss reduction required to add another node to a branch of a decision tree) for values 4, 8, and 12,
e. The value of the L2 regularization term for values 0, 1, and 2,
f. The value of the L1 regularization term for values 0, 1, and 2.

Since the L1 and L2 regularization terms are used in a similar way by adding a term to the loss function to prevent overfitting, we only varied one of them at a time. This means that we explored 12 maximum depths, 3 fractions of feature sampling, 3 fractions of examples used in successive steps of the model, 3 terms for gamma, and 6 regularization terms. In total, we explored 1296 model hyperparameter combinations for each of the six resampling and data inclusion strategies.

To choose the best model from the hyperparameter search, we computed the Matthew's correlation coefficient (MCC) on the confusion matrix values of the validation data set. The MCC is computed as follows:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{2.4}$$

where *TP* is the number of true positives (accurately identified dead trees), *TN* is the number of true negatives (accurately identified live trees), *FP* are the false positives (inaccurately classified as dead trees), and *FN* are the false negatives (inaccurately classified as live trees) and the domain of the *MCC* is between -1 and 1, inclusive. Using the confusion matrices of the null models, it can be shown that the MCC for Null Model I and Null Model II are both 0. For the purposes of computing the MCC, the choice of whether to call the live class the positives or the dead class the positives is arbitrary.

We found that when we chose the model with the highest MCC on the validation data set (MCC_{valid}) , the training accuracy could still be as much as 5% higher than the

validation dataset accuracy. This indicates that our model was overfitting to the training data. To account for this, we filtered out any results from the hyperparameter search where the difference between the training accuracy and validation accuracy was 0.5% or greater before selecting the model hyperparameter combination with the highest MCC score on the validation dataset. We show the spread of the model training and validation accuracy across the hyperparameter search for each maximum depth in Figure 2.4. Each point and error bar in the panels of Figure 2.4 represents 108 models from the hyperparameter search.

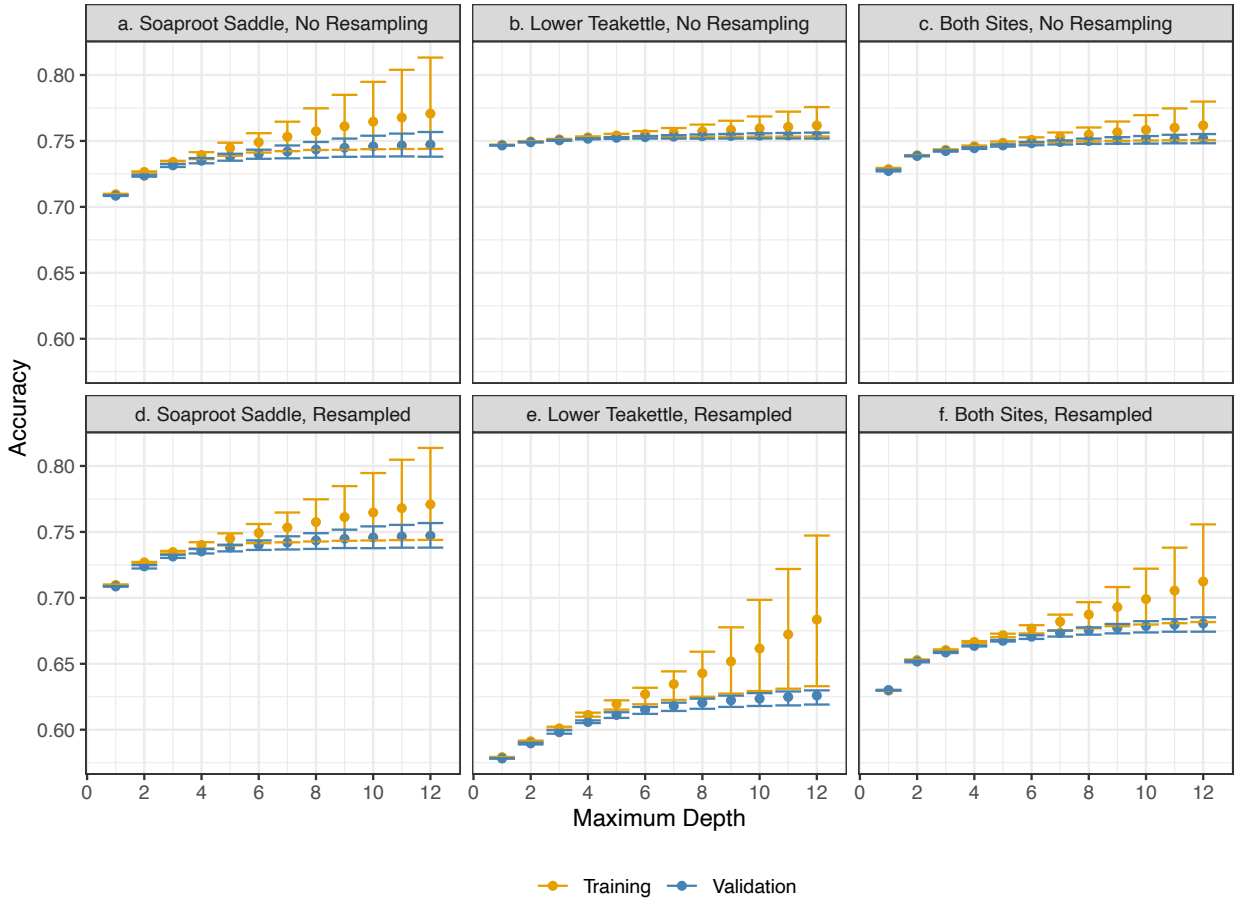

Figure 2.4. The training and validation accuracies for the hyperparameter search are shown for maximum tree depths from 1 to 12. The training accuracies are shown in gold,

and the validation accuracies are shown in dark blue. Error bars represent the middle 95% of the accuracies for the hyperparameter search. We performed the same hyperparameter search for each resampling and data inclusion option including **a.** data from Soaproot Saddle only and no resampling, **b.** data from Lower Teakettle only and no resampling, **c.** data from both sites and no resampling, **d.** data from Soaproot Saddle only and resampling to even the number of live and dead trees, **e.** data from Lower Teakettle only and resampling to even the number of live and dead trees, and **f.** data from both sites and resampling to even the number of live and dead trees.

2.2.3.6 Model Selection

In Table 2.6, we show the training and validation accuracies for the best model of each of the 6 experiments. We also compute the MCC for the training and validation datasets and compute the relative gain in validation accuracy over the theoretical accuracies of the null models. To compute the relative gain in validation accuracy, we computed the following:

$$G = \frac{Acc_{valid} - Acc_{null}}{Acc_{null}} \tag{2.5}$$

where G is the relative gain, Acc_{valid} is the validation accuracy, and Acc_{null} is the theoretical accuracy of the null model.

Table 2.6

Accuracy metrics for each of the six resampling and data inclusion strategies.

	Training accuracy	Training MCC	Validation accuracy	Validation MCC	Relative gain over Null Model 1	Relative gain over Null Model II
Soaproot Saddle, not resampled	0.74	0.49	0.74	0.48	0.47	0.47
Soaproot Saddle, resampled to even the classes	0.74	0.48	0.74	0.48	0.48	0.47
Both sites, not resampled	0.76	0.39	0.75	0.38	0.11	0.33
Both sites, resampled to even the classes	0.68	0.36	0.67	0.36	0.34	0.34
Lower Teakettle, resampled to even the classes	0.61	0.22	0.61	0.22	0.21	0.21
Lower Teakettle, not resampled	0.76	0.20	0.75	0.17	0.01	0.22

The models that only used training and validation data from Soaproot Saddle performed the best. The training data for Soaproot Saddle before and after resampling are very similar, since the classes are almost even in number before resampling. The models achieve a 74% accuracy on both the training and validation datasets which is a 47-48% gain over the null models.

The model experiment with the next best MCC score on the validation dataset is the combined site with no resampling. The training and validation accuracies for this model are both 74% compared with null model accuracies of 68% (guessing all trees are live) and 56% (guessing live trees with probability p) which corresponds to a gain of 9-31%. However, this model underpredicts the dead class of trees with a class accuracy of 35% compared with a class accuracy for live trees of 95% (shown in Table 2.7). The pattern of having a low class accuracy for the smaller class and very high class accuracy for the larger class is similar to Null Model I (guessing all trees are live).

We attempted to avoid the class bias of Null Model I with the next best sampling scheme as rated by the validation MCC. In the dataset with both sites where we resampled to even the classes, we found a training and validation accuracy of 63%. Since the theoretical null model accuracies for this setup are both 50%, this is a gain of 26%. The confusion matrix for this model setup is shown in Table 2.7, part b, where we show the true positives and true negatives both outweigh the errors. The class accuracies are closer together with a class accuracy for dead trees of 55% and an accuracy of 80% for live trees.

The worst-performing models only used data from Lower Teakettle. Our resampled model to even the classes of the Lower Teakettle data achieved an accuracy of 58% (a gain of 16% over the null models) and shows class accuracies over 50% for both classes (55% and 68% for dead and live, respectively). While the Lower Teakettle model with the original data achieved an accuracy of 74%, the class accuracy for dead trees is low (10%) and is very similar to Null Model I.

Table 2.7

Confusion matrices and class accuracies for a. no resampling and b. resampling to even the number of live and dead trees for the training and validation datasets for both sites, data from Soaproot Saddle only, and data from Lower Teakettle only

a. No resampling		Training data			Validation data		
		Modeled dead	Modeled live	Class accuracy	Modeled dead	Modeled live	Class accuracy
Both sites	Labeled dead	68,344	125,708	0.35	22,372	42,651	0.34
	Labeled live	21,763	387,912	0.95	7,494	128,725	0.94
Soaproot Saddle only	Labeled dead	57,453	24,055	0.70	19,072	8,241	0.70
	Labeled live	17,917	63,134	0.78	5,938	20,723	0.77
Lower Teakettle only	Labeled dead	11,639	100,905	0.10	3,528	34,182	0.09
	Labeled live	5,210	323,414	0.98	1,976	107,582	0.98
b. Resampling to even classes		Training data			Validation data		
		Modeled dead	Modeled live	Class accuracy	Modeled dead	Modeled live	Class accuracy
Both sites	Labeled dead	229,114	180,531	0.56	75,914	61,360	0.55
	Labeled live	84,159	325,516	0.79	28,325	107,894	0.79
Soaproot Saddle only	Labeled dead	57,405	24,103	0.70	19,042	8,271	0.70
	Labeled live	17,988	63,520	0.78	5,962	20,849	0.78
Lower Teakettle only	Labeled dead	180,286	148,338	0.55	59,623	50,489	0.54
	Labeled live	107,063	221,561	0.67	35,955	73,603	0.67

We combined the training and validation datasets for each of the six resampling and data inclusion experiments and trained a model on this combined dataset representing 80% of the total data in the study. We used the unique set of optimal hyperparameters that we found for each of the six resampling schemes. The optimal hyperparameters for each resampling scheme are shown in Table 2.8.

Table 2.8

The optimal hyperparameters for each model resampling scheme

	Not resampled			Resampled for even classes		
Hyperparameter	Both sites	Soaproot Saddle	Lower Teakettle	Both sites	Soaproot Saddle	Lower Teakettle
Maximum depth	9	5	11	8	4	4
Fraction of feature variables	0.6	0.6	1.0	0.6	1.0	1.0
Fraction of training data	0.8	0.4	0.8	0.4	0.4	0.6
Gamma	8	8	8	12	4	8
L1 regularization (alpha)	0	0	0	0	0	0
L2 regularization (lambda)	1	1	1	1	1	1

2.2.3.7 Mortality fraction and expected mortality fraction maps

To visualize the modeled mortality risk, we aggregated the tree mortality estimates to the coordinate reference system and resolution of Landsat data (30-meter resolution). We computed mortality fraction in two ways. The first way, which we refer to as the

mortality fraction, is to compute the number of dead trees from the total number of trees in each pixel. The second way, which we refer to as the expected mortality fraction, incorporates the estimated likelihood of belonging to each class produced by the model's softmax function. For each pixel, we sum the estimated likelihood of belonging to the dead class over the total number of trees. In other words, we compute the expected mortality fraction for pixel (i, j),

$$\mu_{i,j} = \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} x_{i,j,k} \tag{6}$$

where i and j are the indices for each pixel corresponding to the x- and y-direction, $N_{i,j}$ is the number of trees in the pixel (i, j), and $0 \leq x_{i,j,k} \leq 1$ refers to the estimated likelihood that the kth tree in pixel (i, j) is dead. In cases where the tree is classified as alive, the estimated likelihood of the tree being dead falls on the interval [0, 0.5).

2.3 Results

2.3.1 Best fit accuracy and confusion matrices

For the best fit models and maps shown in the results, we combined the training and validation dataset into one large training dataset representing 80% of the examples. We fit the models using the optimal parameters that we found for each resampling scheme and test them against the dataset set we held out for testing (20%) which was not used in

earlier parts of our analysis. The testing dataset is not resampled. The training and test dataset accuracies and MCC values are shown in Table 2.9 along with the relative gain over the null models with respect to the test accuracy.

Table 2.9

The training (80% using the combined training and validation datasets) and test (20%) accuracy is shown for the best fit model for each resampling scheme. The relative gain over the null models is computed relative to the test accuracy.

	Training accuracy	Training MCC	Test accuracy	Test MCC	Relative gain over Null Model 1	Relative gain over Null Model II
Soaproot Saddle, not resampled	0.74	0.49	0.74	0.47	0.47	0.47
Soaproot Saddle, resampled to even the classes	0.74	0.48	0.74	0.47	0.47	0.47
Both sites, not resampled	0.76	0.39	0.75	0.38	0.11	0.33
Both sites, resampled to even the classes	0.68	0.37	0.72	0.35	0.43	0.43
Lower Teakettle, resampled to even the classes	0.61	0.22	0.64	0.20	0.28	0.28
Lower Teakettle, not resampled	0.76	0.20	0.76	0.18	0.02	0.22

Table 2.10

Confusion matrices for the combined training and validation dataset used to train the final models (80%) and testing (20%) datasets.

a. No resampling		Combined training and validation data used for training (80%)			Testing data (held out dataset representing 20% of total)		
		Modeled dead	Modeled live	Class accuracy	Modeled dead	Modeled live	Class accuracy
Both sites	Labeled dead	91,801	167,274	0.35	22,355	42,198	0.35
	Labeled live	29,322	516,572	0.95	7,698	128,992	0.94
Soaproot Saddle only	Labeled dead	77,047	31,774	0.71	18,906	8,148	0.70
	Labeled live	23,947	83,765	0.78	6,093	20,944	0.77
Lower Teakettle only	Labeled dead	15,948	134,306	0.11	3,696	33,803	0.10
	Labeled live	7,162	431,020	0.98	2,082	107,571	0.98
b. Resampling to even classes		Combined training and validation data used for training (80%)			Testing data (held out dataset representing 20% of total)		
		Modeled dead	Modeled live	Class accuracy	Modeled dead	Modeled live	Class accuracy
Both sites	Labeled dead	306,659	240,290	0.56	35,822	28,731	0.55
	Labeled live	111,889	434,005	0.80	28,582	108,108	0.79
Soaproot Saddle only	Labeled dead	76,672	32,149	0.70	18,848	8,206	0.70
	Labeled live	24,102	84,217	0.78	6,078	20,959	0.78
Lower Teakettle only	Labeled dead	239,731	199,005	0.55	20,546	16,953	0.55
	Labeled live	142,278	295,904	0.68	35,840	73,813	0.67

In Table 2.10, we show the confusion matrices for each of the resampling schemes for the training and test dataset. The performance is comparable to what we saw for the

comparison between the training and validation datasets in section 2.2.3.6 for model selection.

2.3.2 Mortality fraction maps

The mortality fraction computed using the number of dead trees over the total number of trees in each pixel with at least one tree is shown for each resampling scheme in Figure 2.5 for Soaproot Saddle and Figure 2.6 for Lower Teakettle. In both Figures 2.5 and 2.6, the models are shown in panels a. through f., while the reference map computed from the observations is shown in panel g. The first column shows the result for training on data from Soaproot Saddle only, while the second and third column show the result of training on Lower Teakettle only and data from both sites in the third column. The first row shows the results for data which was not resampled, whereas the second row shows the results for data that was resampled to even the number of examples of live and dead trees. The difference maps for Soaproot Saddle and Lower Teakettle are shown in Figures B1 and B2, respectively.

The models that were trained on data from Soaproot Saddle only and the full data set with no resampling have very similar results at Soaproot Saddle. They also compare relatively well to the observations, though areas of higher modeled mortality are too high, and the areas of low modeled mortality are too low. However, when the training dataset from both sites is resampled to include even number of live and dead trees, we see greater overestimates of tree mortality at Soaproot Saddle (modeled data in Figure 2.5f compared to observations in Figure 2.5g). For the models trained only on Lower Teakettle, we find

almost no tree mortality at Soaproot Saddle using the training dataset that was not resampled (Figure 2.5b) and very little mortality from the model using the resampled training dataset (Figure 2.5e) which does not appear to coincide with the patterns of mortality in the observations.

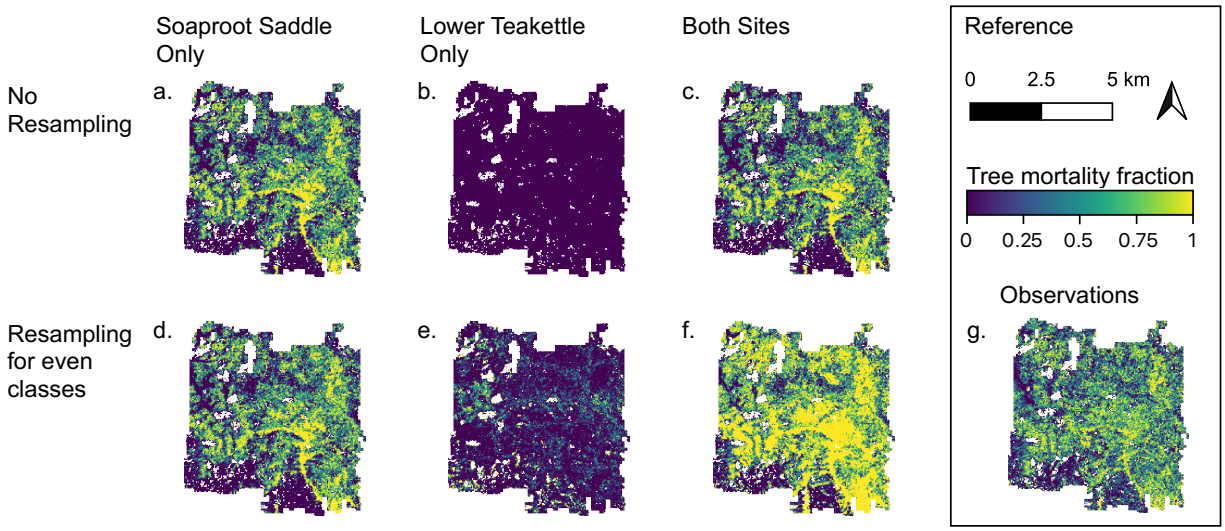

Figure 2.5. Modeled mortality fraction at Soaproot Saddle

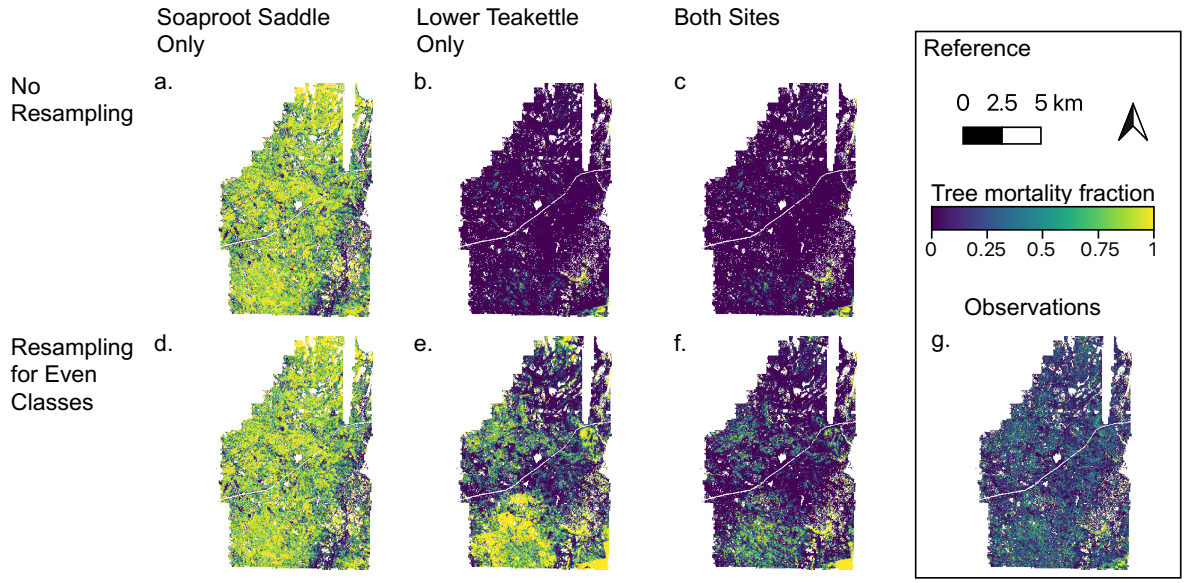

Figure 2.6. Modeled mortality fraction at Lower Teakettle

In contrast to Soaproot Saddle, the spatial pattern of tree mortality that best represents the observations at Lower Teakettle is the resampled dataset from both sites. Both training datasets that use only data from Soaproot Saddle overestimate tree mortality throughout Lower Teakettle. However, the training datasets that only used data from Lower Teakettle also produce poor alignment with the magnitude of tree mortality throughout Lower Teakettle. The training dataset that was not resampled estimates no tree mortality throughout large portions of the study region. However, the dataset that was resampled captures some of the pattern of tree mortality but shows extreme values. For the training datasets using data from both sites, the dataset that was not resampled yields a model that underestimates tree mortality with a similar pattern as the model trained on the data that was not resampled from Lower Teakettle only. The dataset from both sites

that was resampled for even classes captures the pattern and magnitude of mortality at Lower Teakettle the best from the resampling schemes we explored.

2.3.3 Expected mortality fraction maps

We incorporated the estimated likelihood that each tree is dead produced by the model's softmax function to compute an expected mortality fraction that incorporates the model's certainty about each classification. More details are found in section 2.2.3.7 of the methods. The modeled expected mortality fraction is shown in Figure 2.7 for Soaproot Saddle for the six resampling schemes. Qualitatively, the spatial pattern and magnitude of the expected mortality fraction appears highly similar among the models trained on datasets from Soaproot Saddle only (Figure 2.7a and 2.7d) as well as the training dataset from both sites that was not resampled (Figure 2.7c). The expected mortality fraction modeled from the training dataset from both sites that was resampled to even the classes overestimate mortality. The overestimate is apparent in the difference maps (Figure B.3) where there is a much stronger overestimate in mortality fraction throughout Soaproot Saddle in Figure B.3, panel f, than shown in panels a, c, and d. The expected mortality fraction modeled from data from Lower Teakettle underestimates tree mortality fraction and appears to show a different pattern of mortality than seen in the observations.

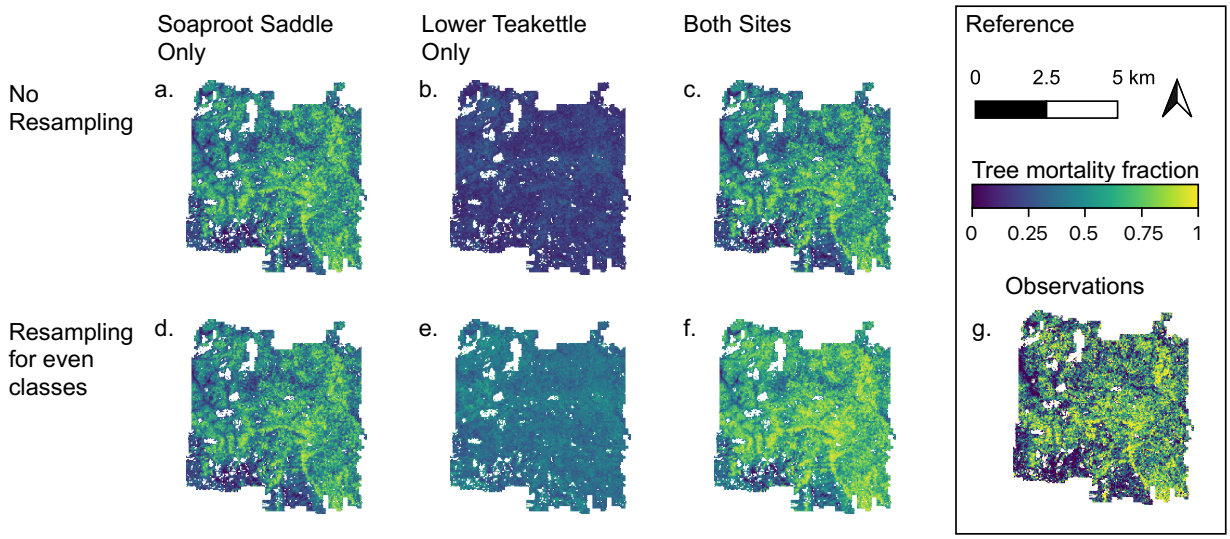

Figure 2.7. Modeled expected mortality fraction at Soaproot Saddle.

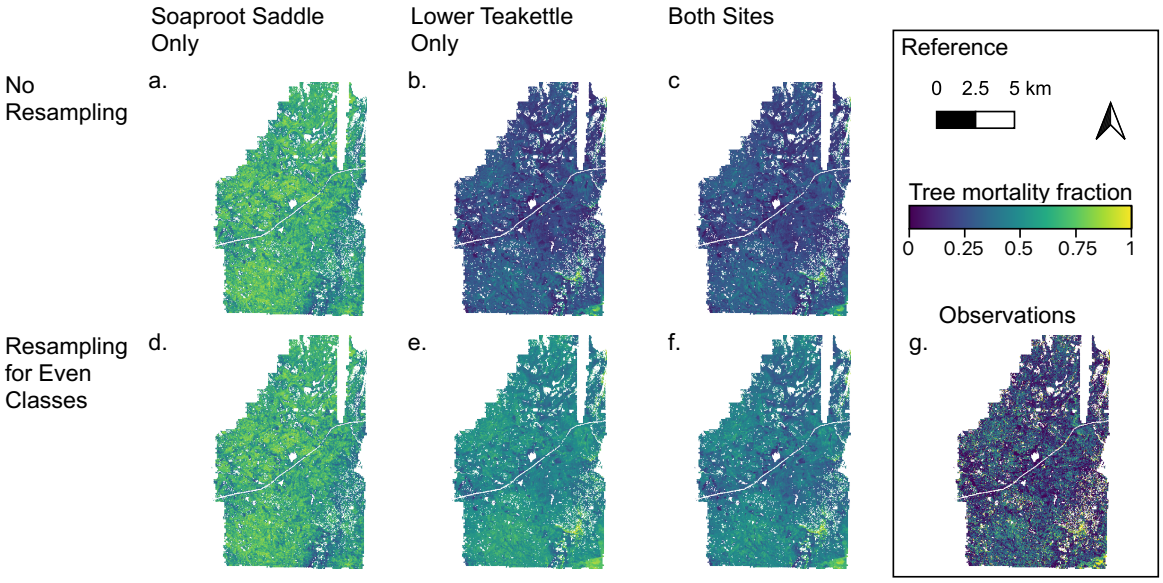

Figure 2.8. Modeled expected tree mortality fraction at Lower Teakettle.

The modeled expected mortality fraction for Lower Teakettle is shown in Figure 2.8, and the differences between the observed mortality fraction and modeled expected

mortality fraction are shown for Lower Teakettle in Figure B.4. The model at Lower Teakettle again overestimates tree mortality when only trained on data from the lower elevation site, Soaproot Saddle (Figure 2.8a and 2.8d). The dataset resampled using only data from Lower Teakettle overestimates tree mortality fraction throughout most of the study site (Figure 2.8e). To a lesser extent, the model trained on the full dataset also overestimates tree mortality throughout Lower Teakettle (Figure 2.8f). However, we see a difference in the spatial pattern and magnitude of tree mortality models trained on data from Lower Teakettle only and from both sites that are not resampled when estimated using the expected mortality fraction. The difference maps show a more balanced and diffuse error pattern for over- and under-estimates of tree mortality for these models (Figure B.4, panels b and c).

In Figure 2.9, we show the 1:1 plots of the modeled mortality fraction and expected mortality fraction for the 30-meter pixels against the observations for pixels with at least 4 trees. For the modeled mortality fraction (first two rows of Figure 2.9), models trained only on data from Soaproot Saddle performed relatively well on data from Soaproot Saddle (Figure 2.9a and 2.9d). In Figure 2.9a, we show the results for models trained on data that was not resampled (i.e., the original training and validation data) in Figure 2.9a and the resampled data in Figure 2.9d. In both cases, the medians of the box plots align well with the 1:1 line. In Figure 2.9b, we show the modeled mortality fraction for models trained on the original (not resampled) training and validation datasets at Lower Teakettle and only include the pixels at that site. The observed mortality fractions at Lower Teakettle between 0 and 0.5 are largely estimated as a mortality fraction of 0. More nonzero estimates of mortality occur for higher observed values of mortality, but the model still underestimates

for mortality fractions between 0.6 and 1.0. The corresponding result with trained on the resampled dataset shows a better fit with the 1:1 line (Figure 2.9e). In both Figures 2.9b and 2.9e, darker boxplot values for lower mortality fractions indicate more trees fall in these lower mortality fraction intervals compared to the light yellow boxes at higher mortality fractions. We see a similar trend with the combined dataset in Figures 2.9c and 2.9f. The accuracy of the model at Soaproot Saddle appears to improve the match to the 1:1 line at higher mortality fractions in Figure 2.9c. Figure 2.9f shows underestimates of mortality for mortality fractions between 0.2 and 0.5 and overestimates between 0.8 and 0.9.

The last two rows of Figure 2.9 show the expected mortality fraction. The expected mortality does poorly at the extremes. These models overestimate mortality fraction for pixels with low mortality and underestimate it for pixels with high mortality.

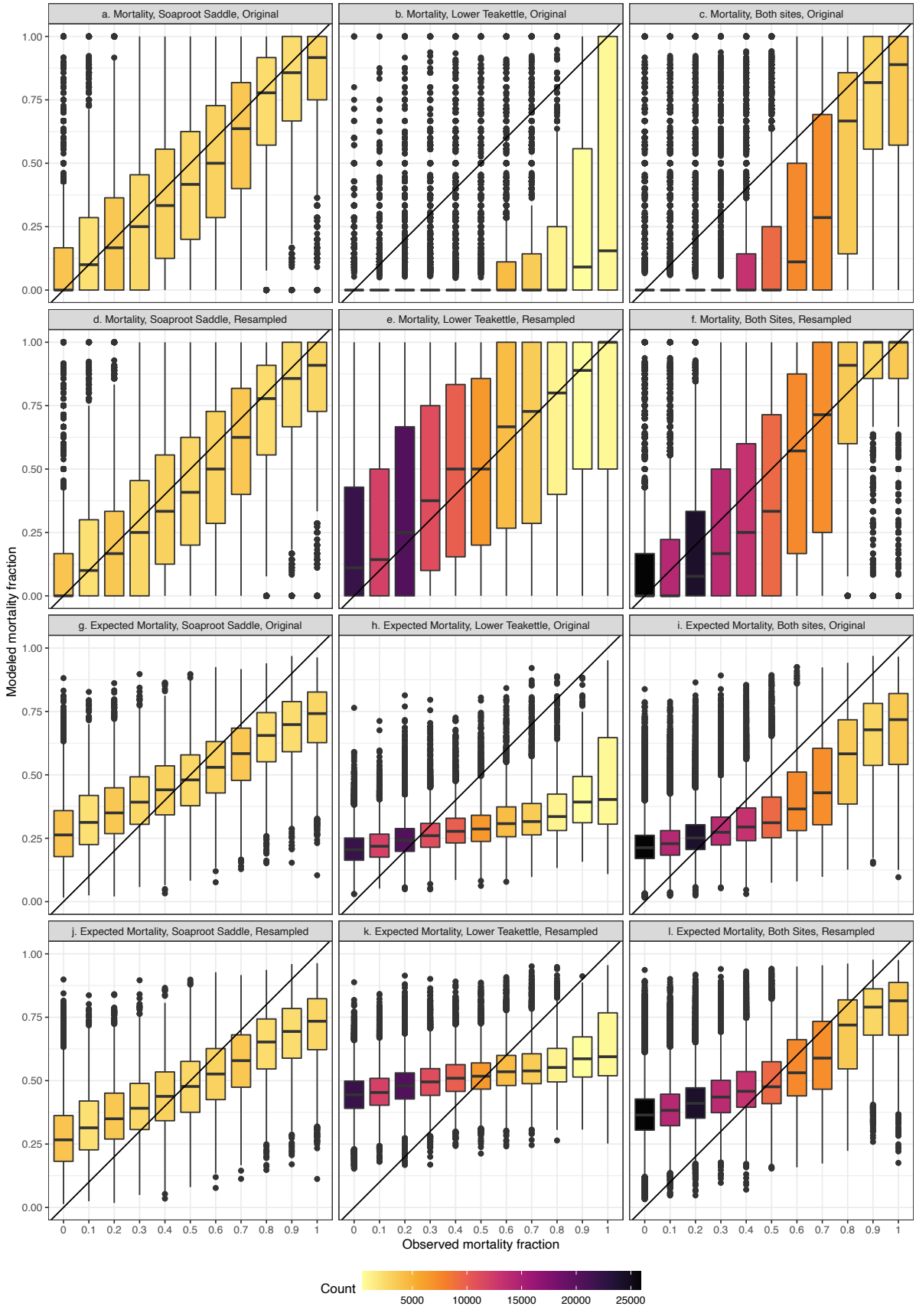
a. Mortality, Soaproot Saddle, Original
b. Mortality, Lower Teakettle, Original
c. Mortality, Both sites, Original
d. Mortality, Soaproot Saddle, Resampled
e. Mortality, Lower Teakettle, Resampled
f. Mortality, Both Sites, Resampled
g. Expected Mortality, Soaproot Saddle, Original
h. Expected Mortality, Lower Teakettle, Original
i. Expected Mortality, Both sites, Original
j. Expected Mortality, Soaproot Saddle, Resampled
k. Expected Mortality, Lower Teakettle, Resampled
l. Expected Mortality, Both Sites, Resampled
Modeled mortality fraction
0.00
0.25
0.50
0.75
1.00
Observed mortality fraction
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Count
5000
10000
15000
20000
25000

Figure 2.9. Plots of mortality fraction against observed mortality fraction for mortality estimates aggregated to Landsat resolution (30 meters) and filtered for pixels with at least four trees. The first two rows show the results for the mortality fraction predictions, while the last two rows show the results for the expected mortality fraction. The first column shows the results of models trained on Soaproot Saddle only and only represents pixels from Soaproot Saddle. The second column shows the same for Lower Teakettle, and the third column shows the results for models trained on data from both sites and shows pixels from both sites. Finally, the first and third rows show the results for models trained on the original training and validation datasets with no resampling. The second and fourth rows show results for models trained on the datasets that were resampled to even the number of live and dead trees.

2.3.4 Variable importance

We used the xgb.importance function within the xgboost package in R, version 4.1.2, to compute the fractional gain from each feature. We computed this for models trained on data sets that were resampled and that were not resampled to check for variations due to the resampling methods and found that models trained on a given site had comparable results regardless of resampling scheme. The results are shown in Table 2.11. Because the order of importance as measured by the fractional gain varies depending on the resampling scheme used, we arranged the table in the order corresponding to the resampling scheme with the best performance. The best model that we found was for the resampling scheme where we used only data from Soaproot Saddle but resampled to even the classes.

The most important variable for each of the resampling schemes was the height of the trees as estimated from NEON lidar in 2013. This feature is especially important for the models for Soaproot Saddle only where the fractional gain from the splits involving this feature among the decision trees contributes to 46-47% of the total gain in accuracy within the model. The next important features at Soaproot Saddle include the distance to rivers and the canopy cover fraction within 20 meters of a tree.

For the models trained and validated on examples from Lower Teakettle only, the baseline NDVI was one of the top three features. The next important feature was slope in the case of the data that was not resampled and the temperature perturbation during the drought in the case of the resampled dataset. For the models which attempted to characterize trends at both Soaproot Saddle and Lower Teakettle, the temperature perturbation during the drought was nearly as important as the height with respect to the fractional gain. The next important variable was slope. The north-to-south and east-to-west aspect were among some of the least important variables we explored.

Models trained on data from both sites depend on the difference in mean dry season temperature during the drought almost as much as tree height (21-24% vs. 28-29%). The third most important variable for both is baseline dry season temperature (8-9%). The elevated dependence on climate variables compared to the site-level models may reflect that model trained on data from both sites use the climate variables to distinguish between the mortality pattern of Soaproot Saddle and Lower Teakettle.

Table 2.11

Variable importance ranking for each resampling scheme shown in order of importance of the best which was Soaproot Saddle data only, resampled for even classes. The bolded values represent the top three feature variables for a given resampling scheme.

	Not resampled			Resampled for even classes		
Variable	Both sites	Soaproot Saddle	Lower Teakettle	Both sites	Soaproot Saddle	Lower Teakettle
Height	**0.27**	**0.47**	**0.23**	**0.26**	**0.50**	**0.35**
Distance to rivers	0.04	**0.08**	0.05	0.04	**0.09**	0.03
Canopy cover fraction	0.04	**0.06**	0.05	0.04	**0.07**	0.04
Slope	0.05	0.06	**0.09**	0.05	0.05	0.08
Trees per hectare	0.03	0.04	0.02	0.02	0.04	0.01
Baseline annual precipitation	0.04	0.04	0.05	0.04	0.03	0.04
Difference in mean dry season temperature during the drought	**0.18**	0.04	**0.08**	**0.21**	0.04	**0.11**
Baseline NDMI	0.03	0.04	0.04	0.03	0.03	0.02
Mean baseline dry season temperature	**0.10**	0.03	0.03	**0.09**	0.04	0.04
Granite fraction	0.02	0.03	0.03	0.02	0.03	0.01
Fraction of baseline annual precipitation during the drought	0.03	0.02	0.06	0.03	0.02	0.05
Baseline NDVI	0.03	0.02	0.07	0.03	0.02	**0.08**

Baseline minimum winter temperature	0.05	0.02	0.04	0.06	0.01	0.05
Difference in minimum winter temperature during the drought	0.04	0.02	0.04	0.04	0.02	0.03
North-south aspect	0.01	0.01	0.03	0.01	0.01	0.01
ET	0.02	0.01	0.05	0.02	0.01	0.03
East-west aspect	0.01	0.00	0.03	0.01	0.00	0.01

2.4 Discussion

Our aim in this study was to find out to what extent we can model individual tree mortality risk before a severe drought using the machine learning method extreme gradient boosting. To complete this objective, we selected seventeen feature variables which include vegetation traits, topography, and climate variables and explored six resampling schemes to train a model to predict whether or not a tree would die. We considered training data from Soaproot Saddle only, Lower Teakettle only, and from both sites. Compared with Lower Teakettle, we found a higher accuracy for tree mortality predictions at Soaproot Saddle and different sets of key drivers of tree mortality at the two sites. In addition, we found that using estimated likelihoods of a tree being dead improved our ability to create a general model of tree mortality risk at both sites.

Our tree mortality models at Soaproot Saddle were more accurate than at Lower Teakettle with an accuracy of 74%, a 48% gain over random guessing. This means that the feature variables we explored are more predictive of the tree mortality at the lower elevation sites. While tree height was the most important predictor variable at both sites, it accounted for 47-50% of the gain in accuracy at Soaproot Saddle compared with 23-35% at Lower Teakettle. In a previous study, Hemming-Schroeder et al. (2023b) showed that the change in mortality likelihood increased with height at Soaproot Saddle at a higher rate than at Lower Teakettle. Mortality likelihood increased by approximately 10% for every 10-meter gain in height at Soaproot Saddle compared to 5% for every 10-meter increase in height (Hemming-Schroeder et al., 2023b). The stronger trend in tree mortality with height at the lower elevation site is likely related to Ponderosa pines living at this elevation range because large Ponderosa pines are preferentially selected by bark beetles when their populations become large enough to successfully attack large trees during a prolonged and warm drought (Robbins et al., 2021). Previous work has found elevated mortality among Ponderosa pines after the 2012-2016 drought (Fettig et al., 2019). Moreover, while Ponderosa pines have an increasing tree mortality likelihood trend with height, trees of other genera show the opposite (Stephenson & Das, 2020) which may help explain why tree height is a less clear predictor of tree mortality at Lower Teakettle and why tree mortality there is harder to predict.

Additional drivers of tree mortality predictions at Soaproot Saddle may also be important due to the interaction between drought stress and bark beetle behavior at low elevations. Previous work on tree mortality at Soaproot Saddle showed that tree mortality risk increased with increasing distance to rivers and with increasing canopy cover fraction

(Hemming-Schroeder et al., 2023b). Being far from a river and in competition with neighboring trees for available soil water may increase water stress for a given tree and therefore likelihood of that tree being attacked by bark beetles. Moreover, a tree's likelihood of being attacked by bark beetles may also increase in more densely populated neighborhoods of host trees. Higher elevation sites with cooler conditions and different species may require different feature variables to accurately represent patterns of mortality.

However, we found that the predicted spatial tree mortality pattern at both sites improved when we considered the expected mortality fraction. Several machine learning methods including extreme gradient boosting use an activation function for classification which produces an estimated likelihood of belonging to one class or another. In a binary classification problem, a tree with an estimated mortality likelihood of 0.51 can have the same classification as a tree with an estimated mortality likelihood of 0.99, even though one classification is much more certain than the other which has almost no certainty. The expected mortality fraction incorporates the uncertainty. A limitation of this metric is that mortality fractions near 0 and 1 are challenging to predict, because they would require high model certainty to reach. Conversely, the binary classification approach yields more extreme values (near 0 and 1) than seen in the data. Future work may explore improving the magnitude of the predictions.

We found relatively high accuracy at Soaproot Saddle, a warmer and dry low elevation site in the Sierra Nevada. Improving tree mortality risk modeling in this region may help to inform stewardship efforts for forests which may no longer match the current climate and are at risk of shifting to new biome types after a large disturbance. Soaproot

Saddle also represents an elevation range conducive to Ponderosa pine which were killed in large numbers during the 2012-2016 drought because of the combined effects of drought stress and growing western pine beetle populations. Representing these models effectively with mechanistic models may require a relatively high number of parameters and model complexity. Our results show that an off-the-shelf machine learning method like extreme gradient boosting may be an efficient way to estimate mortality risk in transitional zones like the lower elevations of the Sierra Nevada which are susceptible to the combined effects of drought and insect attack.

2.5 Conclusion

We used a recently available dataset of over one million trees to retrospectively model individual and aggregated tree mortality risk during a severe drought in the Sierra Nevada ecoregion of California. We found that tree mortality risk was easier to predict at a lower elevation site which experienced 50% tree mortality during the 2012-2016 drought. Our study shows that some of the complex interactions between drought and insect attack may be captured by an off-the-shelf machine learning model. While tree mortality at lower elevations was more challenging to predict, we found that incorporating model-estimated likelihood of a tree being dead improved the spatial structure and magnitude of our mortality predictions, particularly for the model trained on the original data which has class imbalances. Modeling tree mortality risk in forests most vulnerable to large-scale disturbance may help inform forest conservation efforts and estimates of carbon losses from these ecosystems as the climate changes.

CHAPTER 3

Estimating tree mortality fraction after the 2012-2016 California drought using convolutional neural networks and Landsat time series data

Adapted from an unpublished manuscript with permission from coauthors Luciana Chavez Rodriguez, Padhraic Smyth, Steven Allison, and James Randerson

3.1 Introduction

California experienced a severe drought between 2012 and 2016. Between the years 2012 and 2015, a precipitation deficit led to an estimated 1.5-meter overdraft in the water budget in the Sierra Nevada (Goulden & Bales, 2019). Moreover, subsurface water may have been depleted to a subsurface depth ranging between 5 and 15 meters (Goulden & Bales, 2019). Canopy water losses in some areas exceeded 30% (Asner et al., 2016). The drought was also unusually warm with temperatures increasing throughout the drought period (Crockett & Westerling, 2018). During drought conditions, endemic bark beetle populations can grow to epidemic numbers through a combination of surviving warmer winters and attacking drought-stressed trees (Robbins et al., 2021). The combined effects of the warm drought and insect attack led to mortality fraction as high as 50% by the end of the drought in warmer, lower elevation regions of the Sierra Nevada (Hemming-Schroeder et al., 2023b).

Regional estimates of tree mortality following the drought used high-resolution lidar data from flight campaigns and meter-to-submeter scale imagery to identify trees and classify them as dead or alive (Hemming-Schroeder et al., 2023b; Stovall et al., 2019). These studies span an area of about 160 km^2 coinciding with the National Ecological Observatory Network's Soaproot Saddle and Lower Teakettle sites in the Sierra Nevada east of Fresno, California, in the Sierra National Forest. Estimates of tree mortality in Hemming-Schroeder et al. (2023b) range from 25 to 50% with the highest areas of tree mortality occurring at low elevations and among the tallest trees. However, a key limitation of high-resolution data is that it is often limited to a smaller domain.

A key large scale data set for investigating tree mortality during the 2012-2016 California drought are the annual Aerial Detection Survey (ADS) data sets generated through the United States Forest Service's Aerial Survey Program (2023). This data set is hand-drawn on tablets and collected while viewing a three-kilometer swath of a forested areas from small aircraft (Coleman et al., 2018; United States Forest Service Aerial Survey Program, 2023). A research study investigating the accuracy of the polygons collected for the ADS data sets showed that damage type is correctly classified 97% of the time relative to a ground truth data set primarily collected in California (Coleman et al., 2018). Several studies have used this dataset to train models of tree mortality from satellite data. For example, Young et al. (2017) fit a statistical model based on modeled climatic water deficit and basal area from LEMMA to the ADS data and report an adjusted R^2 of 0.20 between their model and observations. In another study, Byer and Jin (2017) used random forests in a two-stage model to estimate tree mortality severity class at a resolution of 250 meters from topography and time series indices derived from MODIS and report an accuracy of

96.3% with class accuracies for the non-background mortality classes ranging from 73-81%. Their regression model which also used random forests estimated the trees dead per acre with a reported root mean square error of 7 trees on the validation data set (Byer & Jin, 2017).

Additional studies have used satellite time series data characterize disturbance from insect attack. The Landsat-based Detection of Trends in Disturbance and Recovery (LandTrendr) algorithm draws linear segments creating a linear piecewise function through the time series to detect change disturbance, recovery and stable ecosystem states (Kennedy et al., 2010). In a subsequent study, Liang et al. (2014) added a classification component to the LandTrendr algorithm to further classify disturbed pixels with mountain pine beetle outbreaks in Colorado using National Agricultural Imagery Program images to generate training data. Classification accuracies on the validation data set ranges from 87-94% (Liang et al., 2014).

In 2014, Zhu and Woodcock published the Continuous Change Disturbance Classification (CCDC) algorithm. The CCDC algorithm also segments the time series but fits harmonic terms in addition to a degree one polynomial to capture seasonal patterns as well as underlying trends. If a pixel's surface reflectance exceeds the prediction of the harmonic model for three consecutive time points, a disturbance is identified. A random forest model is used to classify the disturbance and had a 92% accuracy for detecting whether or not there was a change on the validation data set (Zhu & Woodcock, 2014). Some of the omission errors include pixels with a partial disturbance (Zhu & Woodcock, 2014). To address challenges the CCDC algorithm has in identifying more gradual disturbance types like beetle kill and drought stress, Zhu et al. (2020) built upon the CCDC algorithm by

optimizing the surface reflectance bands used and threshold parameterizations to create the Continuous Monitoring of Land Disturbance algorithm which has a producer's accuracy of 60% for class which included insect attack. They also found that using key surface reflectance bands including green, red, near infrared, and shortwave infrared 1 and 2 was more important than using indices derived from these bands and that using a multispectral approach was better than a single band alone (Zhu et al., 2020). Ye et al. (2021) build on the Continuous Monitoring of Land Disturbance algorithm using a state space model and achieved a producer's accuracy of 85% for detecting insect attack in Colorado (Ye et al., 2021).

While the previous classification methods perform particularly well in cases where a disturbance affects an entire pixel, the classification of disturbances like insect attack have been more challenging (Zhu et al., 2020). Moreover, much of this work has focused on making a binary or multiclass classification rather than predicting a continuous variable for each pixel. A recent study by (Schiefer et al., 2023) linked high-resolution tree mortality fraction from unmanned aerial vehicles with under 4 cm resolution to the relatively high resolution Sentinel-1 and Sentinel-2 timeseries using long short term memory networks. They found a correlation between the observed and predicted mortality fraction of 0.66. Similarly to Zhu et al. (2020), they found that using several bands yielded the best outcome. Similar methods may help link high resolution data sets from flight campaigns available over a wider domain and the 40-year Landsat time series.

In this study, we pose the following question: Can convolutional neural networks be used to predict cumulative tree mortality fraction within a Landsat pixel after the 2012-2016 drought from Landsat time series data? We use the time series of surface reflectance

from Landsat 5 and Landsat 8 as our feature variables and use a one-dimensional convolutional neural network through the time series to estimate tree mortality fraction in 2017. We start with a convolutional neural network rather than a recurrent neural network like the long short term memory network to begin with a simpler model. We used 30-meter 2017 tree mortality fraction raster data derived from an individual tree mortality data set described in Hemming-Schroeder et al. (2023b) and publicly available (Hemming-Schroeder et al., 2023a). By training a model which can be applied to a larger domain than the individual tree mortality data set, we aim to provide a tool to characterize the impact of the drought on the carbon cycle at a larger scale.

3.2 Methods

3.2.1 Study site

The individual tree mortality data set that we are using comes from the Soaproot Saddle and Lower Teakettle National Ecological Observatory Network sites in the Sierra National Forest. Figure 3.1 shows the spatial extent of the individual tree mortality data set. During the 2012-2016 California drought, these sites experienced 25-50% tree mortality (Hemming-Schroeder et al., 2023b). Higher tree mortality (nearly 50%) occurred at the lower elevation site, Soaproot Saddle. The elevation range at Soaproot Saddle is from 1000-1400 m (Krauss, 2018). It is also the drier of the two sites and has a mean annual precipitation of 900 mm per year (Krauss, 2018). The most common tree species at Soaproot Saddle include Ponderosa pine (*Pinus ponderosa*), incense cedar (*Calocedrus*

decurrens), canyon live oak (*Quercus chrysolepis*), and California black oak (*Quercus kelloggii*) (Krauss, 2018). Among these tree species, Ponderosa pine may have experienced the highest mortality rates due to bark beetle attack during the drought (Krauss, 2018).

In contrast, Lower Teakettle has a higher elevation range of 2000 to 2800 meters and a higher average annual rainfall of 1200 mm per year (Krauss, 2018). Whereas snow at Soaproot Saddle melts quickly, the snowpack at Lower Teakettle has a 30-year average depth of 1140 mm (Krauss, 2018). Dominant tree species at Lower Teakettle include red fir (*Abies magnifica*), white fir (*Abies concolor*), Jeffrey pine (*Pinus jeffreyi*), and lodgepole pine (*Pinus contorta*) (Krauss, 2018).

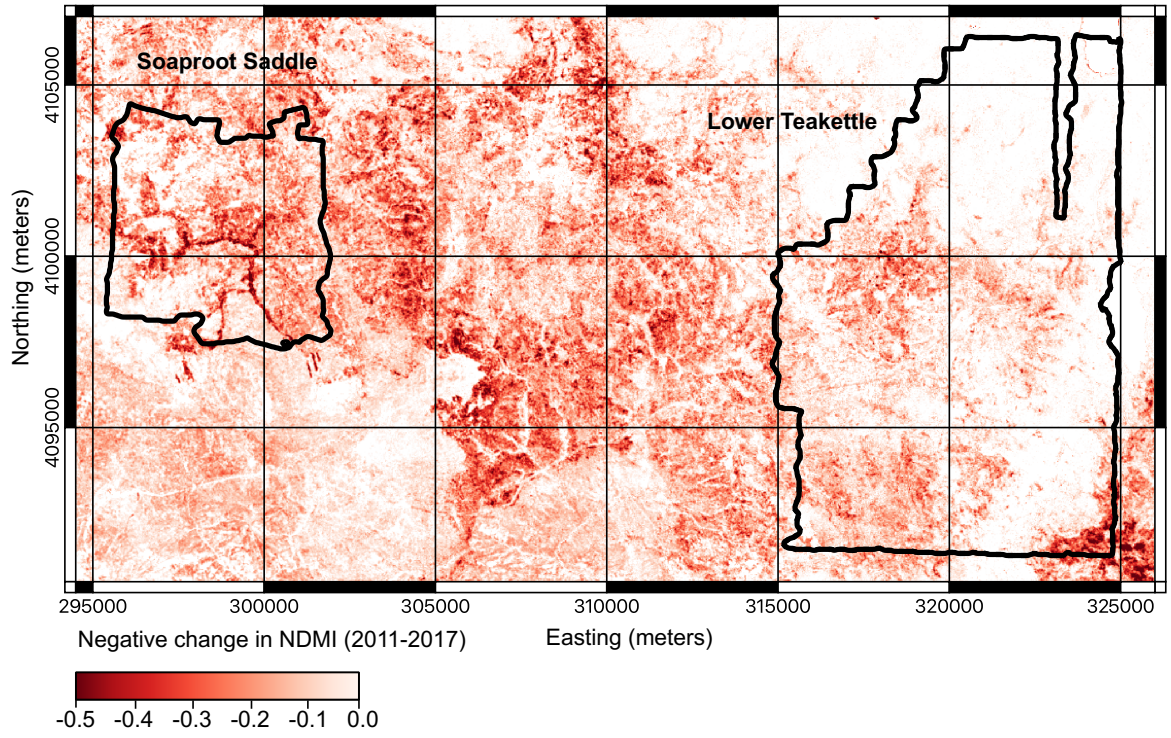

Figure 3.1. The decrease in NDMI between 2011 (before the drought) and 2017 (after the drought) for our study region. The outline of Soaproot Saddle is shown to the west and Lower Teakettle to the east. For visual clarity, we show the negative change (decrease in

NDMI during the drought) separate from the positive change (increase in NDMI during the drought). The positive change is shown in Figure C.1.

3.2.2 Data Sets

3.2.2.1 Landsat surface reflectance

We downloaded Landsat 5 Thematic Mapper, Landsat 7 Enhanced Thematic Mapper Plus, and Landsat 8 Operational Land Imager data sets from the Collection 2 Level 2 Tier 1 surface reflectance data inventory on the United States Geological Survey (USGS) Earth Explorer website. The Landsat 5 data we downloaded begins in April of 1984 and ends in October of 2011. Landsat 7 data covers some of the gap between Landsat 5 and Landsat 8 in 2012 from May 2012 through October 2012, and we use Landsat 8 data from May 2013 through June of 2023.

After converting the Landsat digital numbers to surface reflectance and filtering for clouds, cloud shadows, snow, and water using the QA Pixel later, we computed 6 bands. We found the relative red, green, and blue surface reflectance by taking those surface reflectance bands and dividing by their sum. Next, we computed mean luminosity by taking the average of the three visible bands. Finally, we computed two vegetation indices, the normalized difference vegetation index (NDVI) and the normalized difference moisture index (NDMI). NDVI is computed by taking the difference between the near infrared band (band 3 for Landsat 5 and Landsat 7 and band 4 for Landsat 8) and the red band (band 4 for Landsat 5 and Landsat 7; band 5 for Landsat 8) and diving by their sum. Similarly, NDMI

is computed by taking the difference between the near infrared band and first shortwave infrared band (band 5 for Landsat 5 and Landsat 7 and band 6 for Landsat 8) and dividing by their sum. The domains of the relative visible bands and luminosity are between 0 and 1, inclusive, whereas the domains of NDVI and NDMI are between -1 and 1, inclusive. For NDVI and NDMI, values closer to 1 indicate vegetation that is greener and moister, respectively.

In Figures 3.2 and 3.3, we show the starting NDMI in 2011 before the drought and the progressive decreases in NDMI during the drought period for Soaproot Saddle and Lower Teakettle, respectively. Decreases in NDMI indicate drying. For visual clarity irrespective of color, we plotted the decreases in NDMI separately from the increases. Figures C2 and C3 show the baseline 2011 NDMI and increases in NDMI throughout the drought period. These are much less pronounced than the progressive drying. While the drought ends in 2016, we still see an overall decline in NDMI in 2017 for both sites.

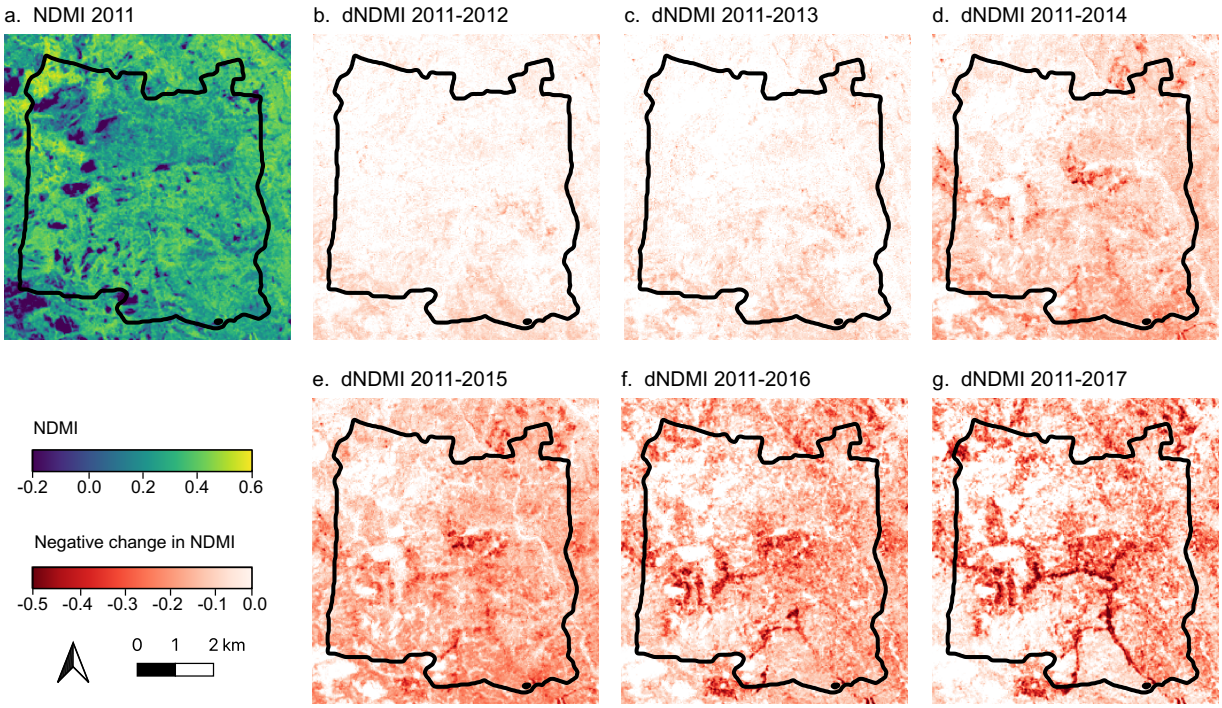

Figure 3.2. Progressive decrease in NDMI during the 2012-2016 drought at Soaproot Saddle. a. The baseline NDMI in 2011. b. The decrease in NDMI between 2011 and 2012, c. 2011 and 2013, d. 2011 and 2014, e. 2011 and 2015, f. 2011 and 2015, g. 2011 and 2016, and h. 2011 and 2017. The respective increases in NDMI at Soaproot Saddle are shown in Figure C.2.

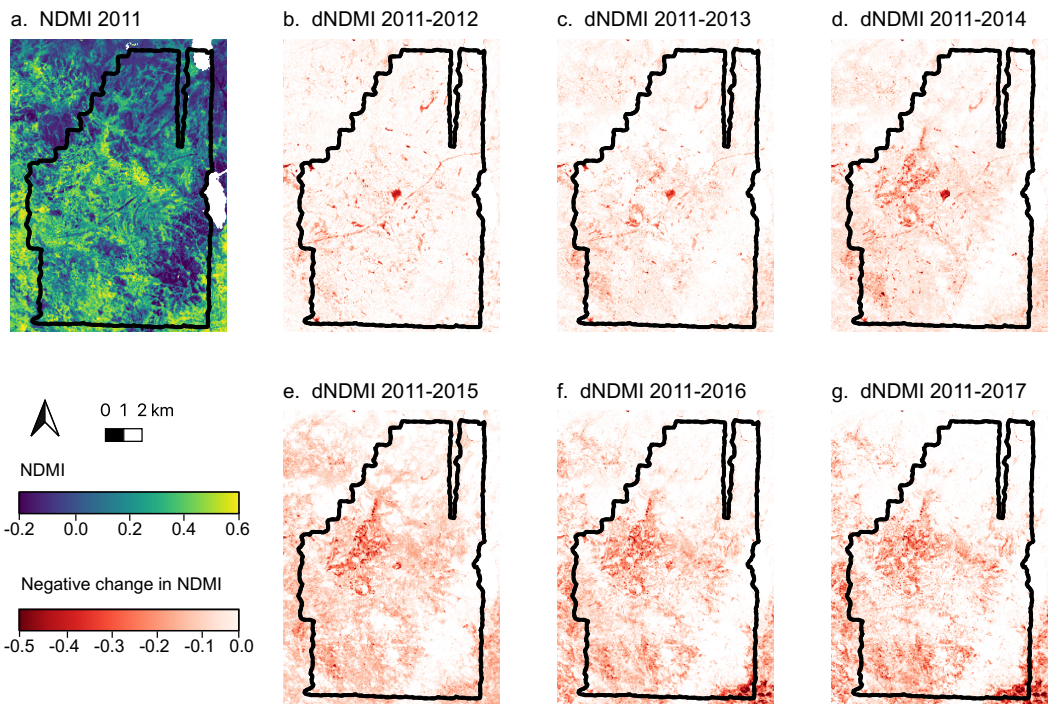

Figure 3.3. Progressive decrease in NDMI during the 2012-2016 drought at Lower Teakettle. a. The baseline NDMI in 2011. b. The decrease in NDMI between 2011 and 2012, c. 2011 and 2013, d. 2011 and 2014, e. 2011 and 2015, f. 2011 and 2015, g. 2011 and 2016, and h. 2011 and 2017. The respective increases in NDMI at Lower Teakettle are shown in Figure C.3.

3.2.2.3 Landsat Time Series

The Landsat satellites have a return interval of 16 days. We explored each possible 16-day interval sequence of time points between 1984 and 2023 and selected the one which coincided with the most Landsat scenes. Each Landsat scene was matched to its nearest date. While Landsat 7 is offset by 8 days from Landsat 5 and 8, we dropped Landsat

7 due to striping in the imagery for the time series models. We obtained the 2020 Creek Fire perimeter from the National United States Forest Service Final Fire Perimeters (USDA Forest Service National Forest System Lands GIS and Fire personnel, 2023), and plotted the time series for three pixels randomly selected from within the fire perimeter to validate our time series. The time series for NDMI is plotted in Figure 3.4 for these three pixels. The points prior to the drought period (orange rectangle from 2012 to 2016 water-years) are from Landsat 5, while the data beginning in 2013 is from Landsat 8. While some of the declines in NDMI during the drought are more subtle, there is a marked decline during the burning period of the 2021 Creek Fire. For our analysis of 2017 tree mortality, we only include time series data through September of 2017.

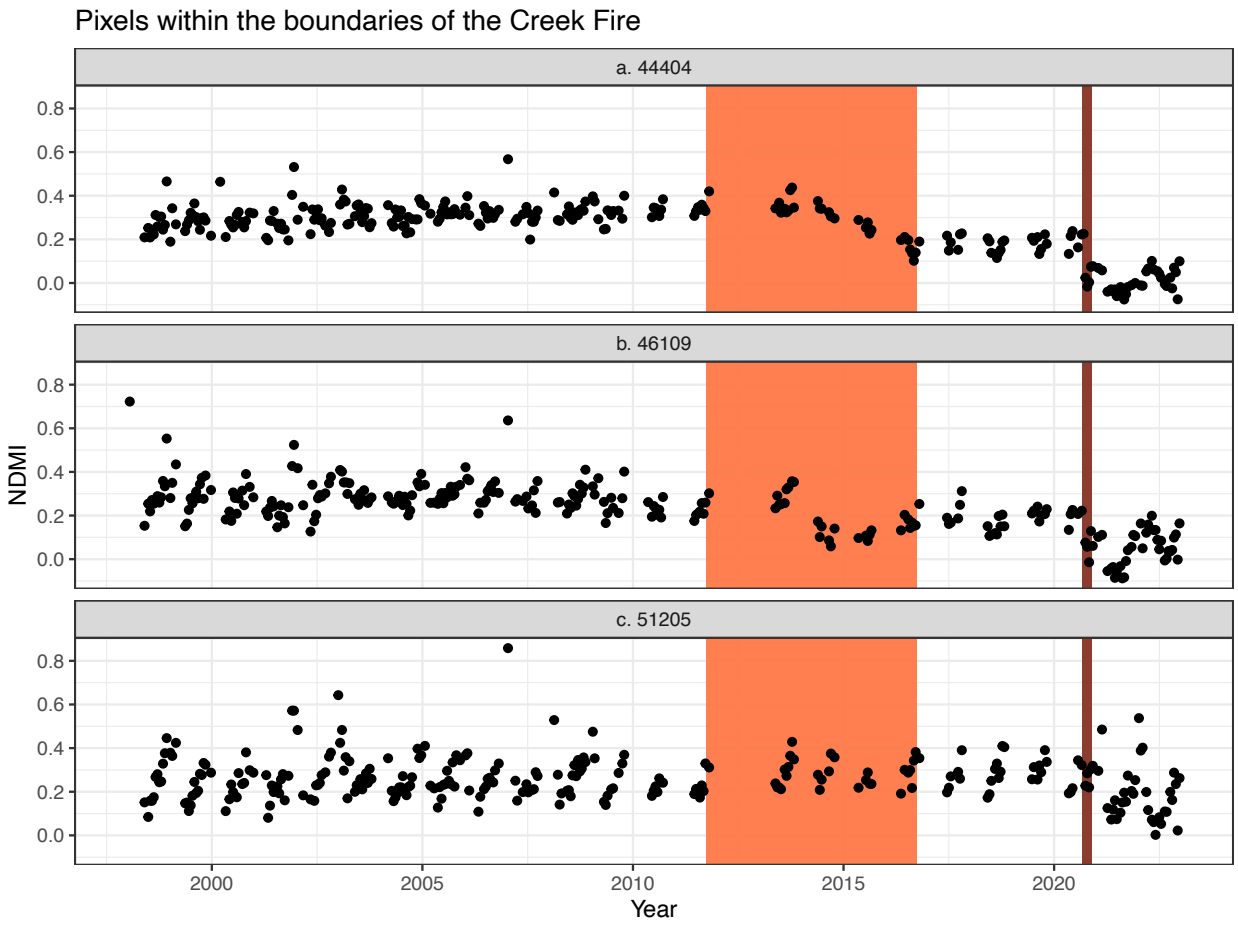

Figure 3.4. Pixels within the 2020 Creek Fire perimeter. We randomly selected three pixels with at least 4 trees to validate the time series data. While our algorithm only uses the time series through September of 2017 to model the mortality fraction in 2017, we wanted to ensure that the time series appeared to be accurately assembled. The orange rectangle represents the drought period from 2012 through 2016, while the dark orange rectangle in late 2020 represents the burning period of the Creek Fire. There is no data in 2012, because we cut Landsat 7 data from our analysis. Other missing data largely reflects time points that were filtered out due to cloud cover, cloud shadow, or snow.

3.2.2.3 Individual Tree Mortality

For our tree mortality observations, we used an individual tree mortality data set described in Hemming-Schroeder et al. (2023b) and published on Zenodo (Hemming-Schroeder et al., 2023a). This data set is primarily derived from lidar and hyperspectral data from the Soaproot Saddle and Lower Teakettle sites from the NEON. The lidar and hyperspectral data are collected using NEON's airborne observation platform for the years 2013, 2017, 2018, 2019, and 2021. Since the algorithm uses the combined lidar point cloud to detect treetops, a tree present in 2013 which falls by 2017 will still be included in tree mortality estimates. Any trees that are not at least 5 meters tall in 2013 are filtered out of the data set. For the mortality fraction data set, we used the mortality fraction in 2017 for all trees within a given Landsat pixel.

The published data set from Hemming-Schroeder et al. (2023a) includes a trees per pixel raster file at the spatial resolution of the Landsat data (30 meters). However, some of the pixels with a value of zero trees may have some trees in them which were not detected in the treetop algorithm. More details about our masking strategy to obtain valid pixels with zero trees are included in the next section.

3.2.2.4 National Ecological Observatory Network canopy height models

We downloaded all available canopy height data for Soaproot Saddle and Lower Teakettle from NEON which includes the years 2013, 2017, 2018, 2019, and 2021. The intersection of these data at the two sites over these five years is shown in the outlines in Figure 3.1. The canopy height models are derived from lidar point clouds collected on

NEON's airborne observation platform. The spatial resolution of the canopy height models is 1 meter.

We aggregated this data by taking the maximum canopy height model value for each Landsat pixel in our study region. If our trees per pixel raster data from section 3.2.2.2 showed zero trees but the aggregated canopy height model showed a maximum height in the Landsat pixel over 5 meters, we concluded that we were missing data in that pixel. These pixels were assigned a value of NA.

3.2.3 Model

3.2.3.1 Feature and target variables

We used the time series data for each of our Landsat bands (relative blue, relative green, relative red, luminosity, NDVI, and NDMI) as our feature data. If data was missing due to missing or dropped scenes (e.g. Landsat 7 data), we filled those time points with 0 for bands with a domain between 0 and 1 and -1 for indices with a domain from -1 to 1. We then transformed NDVI and NDMI to fall between 0 and 1 by adding 1 and dividing by two. For each band, we train parameters for the following model:

$$f_i(X_i) = \vec{y} \tag{3.1}$$

where i refers to one of the six bands, f_i is the one-dimensional convolutional neural network trained on data from band i, X_i is the time series data for band i, and $\vec{y}$ is the 2017

tree mortality fraction of each pixel. In this equation, X_i has dimensions N by M where N is the number of pixels and M is the number of time points, and $\vec{y}$ has dimensions N by 1. For pixels with zero trees, we assigned a mortality fraction of zero. Our rationale is that an ideal model based on the time series will be able to recognize the signature of pixels with no trees and assign a mortality fraction of zero.

In addition to fitting a model for each band separately, we explore a one-dimensional convolutional neural network on the combined bands. In this case, we have the same overall structure as Equation 3.1, but X_i is now a 3D array with dimensions N by M by 6 where the last dimension represents the number of channels or bands.

3.2.3.2 Convolutional neural network architecture

We used the *keras* library in R, version 4.1.2, for the convolutional neural networks. In all the models, we use a one-dimensional convolutional neural network, because we only convolve a kernel in one dimension, time. In the case of training a model on a single band, we convolve a 1 by j kernel through the time series for each pixel where j is the number of time points covered by our kernel. Meanwhile, for the case with all the bands, we convolve a 1 by j by 6 kernel through the time series for each pixel where 6 represents the number of channels or bands.

The architecture of the one-dimensional convolutional neural network is described in Table 3.1 for the model trained on a single band. For the model trained on all the bands, the key difference is that the kernel sizes are 3x3 with 2x2 pooling. To select an architecture, we began with a base model described in Chollet and Allaire (2018) that

performs well for classification on the modified National Institute of Standards and Technology handwriting data sets (Deng, 2012). We kept the length three kernel in Step 2 which represents a time period of about 6 weeks to first collect sub-seasonal patterns in the time series. We alternate between pooling with a size of two which halves the number of terms and using a relatively small kernel (length 3) to try to tease out key patterns in the data while reducing the number of terms prior to flattening the array in Step 11. We include two dropout regularization steps after two of the network layers to reduce overfitting. The dropout rate represents the fraction of inputs that are randomly set to zero during model training. Additional analysis on the sensitivity of the results to the network architecture would help inform how well this model may perform in other use-cases.

Table 3.1

Architecture of convolutional neural network for training on single bands

Step	Description
1	Initiate a sequential model using the *keras* library in R, version 4.1.2
2	Convolve a length 3 kernel through the time series and generate 32 filters with ReLu activation function
3	Pool the result with pool size of 2
4	Convolve a length 3 kernel through the time series and generate 64 filters with ReLu activation function
5	Pool the result with pool size of 2
6	Convolve a length 3 kernel through the time series and generate 64 filters with ReLu activation function
7	Pool the result with pool size of 2
8	Convolve a length 3 kernel through the time series and generate 64 filters with ReLu activation function
9	Dropout regularization with rate 0.25
10	Flatten the filters to a one-dimensional array
11	Dense layer with length 64 using ReLu activation function
12	Dropout regularization with rate 0.50
13	Output single value between 0 and 1 with sigmoid function

3.3 Results

3.3.1 Model Performance

We set a seed and generated a partition of the pixels in our study region to 60% training, 20% validation, and 20% held out until the end for testing. The same training pixels are used to train each model for the models trained on each band and on all the bands. Similarly, we use the same validation pixels for all the models to select a best model. The results on our training and validation data sets at each training epoch are shown in Figure 3.5. While the *keras* library generates its own validation data set from the training data internally, we used our pre-selected validation data set to ensure we were comparing model performance on the same training and validation data sets.

The mean square error for both the training and validation data sets falls below 0.17 for the model trained on all the bands, outperforming the cases of each band considered separately. While training the models, we saved each model after each epoch, so that we could choose the model with the optimal number of epochs. Because the performance on the validation data set stopped improving after the 22nd epoch for our best model, we used that version of the model for our tree mortality predictions. In practice, this means that we can run our combined data set (training, validation, and test) of pixels through the saved model and rasterize the output to view our mortality fraction maps.

The performance of the model on the training and validation data sets after 22 epochs is shown in Table 3.2. Our best model had a mean square error of 0.158 and 0.165 on the training and validation data sets, respectively. On our test data set we held out until

the end of our model selection process for assessing the model performance, the model had a mean square error of 0.164.

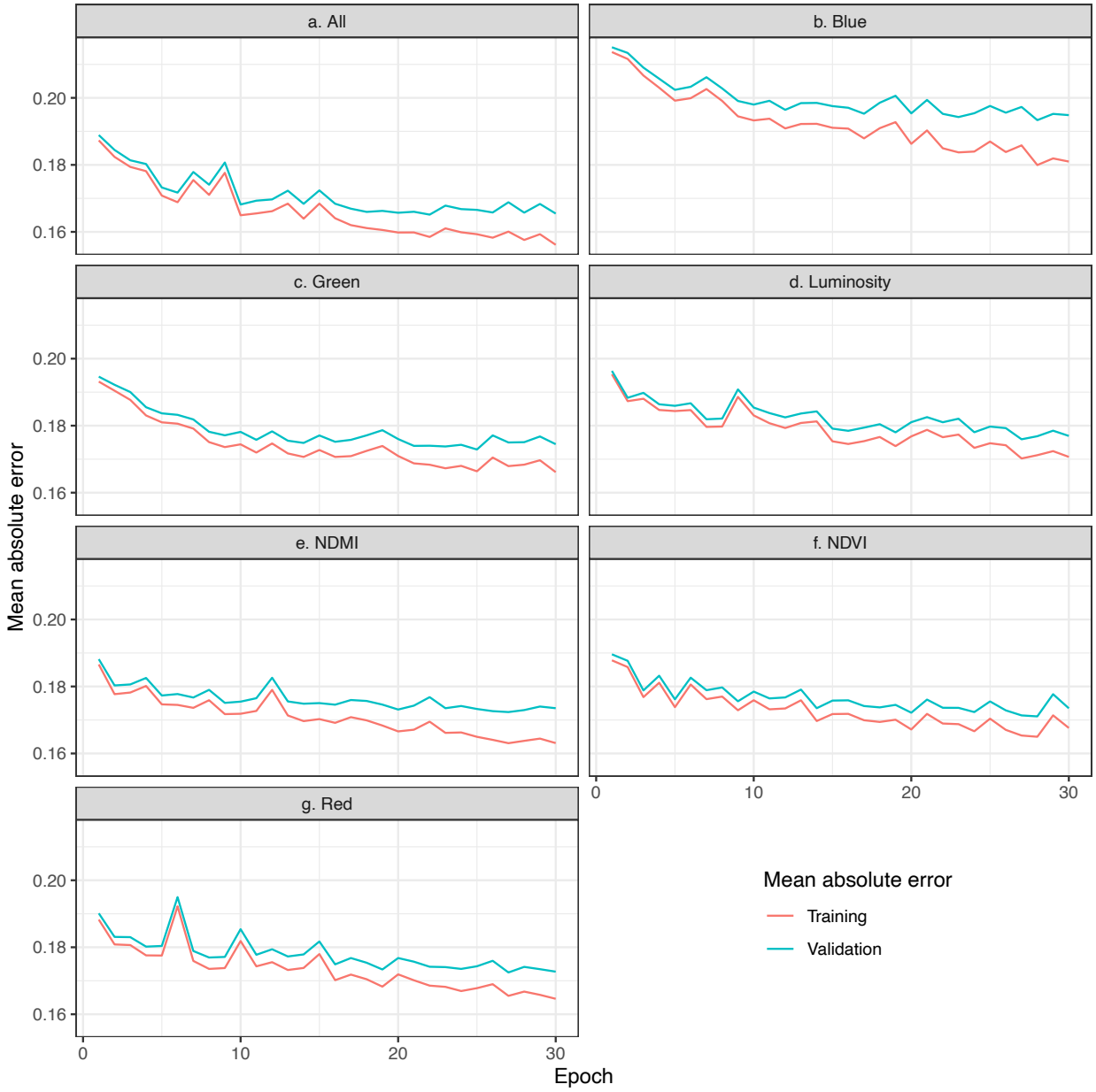

Figure 3.5. The mean absolute error on the training and validation data sets while training the model for a. all the bands, b. relative blue, c. relative green, d. luminosity, e. NDMI, f.

NDVI, and g. relative red. The performance on the training data set is shown in orange, while the performance on the validation data set is shown in aqua.

Table 3.2

Mean absolute error on the training (60%) and validation (20%) data sets for 22 epochs arranged by performance on the validation data set.

Band	Training	Validation
All	0.158	0.165
Red	0.169	0.174
NDMI	0.168	0.174
Green	0.169	0.174
NDVI	0.169	0.177
Luminosity	0.177	0.181
Blue	0.185	0.195

We found that the CNN that we trained on all 6 Landsat-derived bands achieved the best results with a mean absolute error of 0.158 on the training data and 0.165 on the validation and test data.

3.3.2 The spatial distribution of the tree mortality fraction predictions

We show a comparison of the tree mortality observations and predictions in Figure 3.6. Much of the spatial structure from the observations is present in the CNN predictions, though some of the finer detail is lost. The errors are relatively diffuse with little spatial structure which shows some evidence that there is little spatial bias in the modeled result.

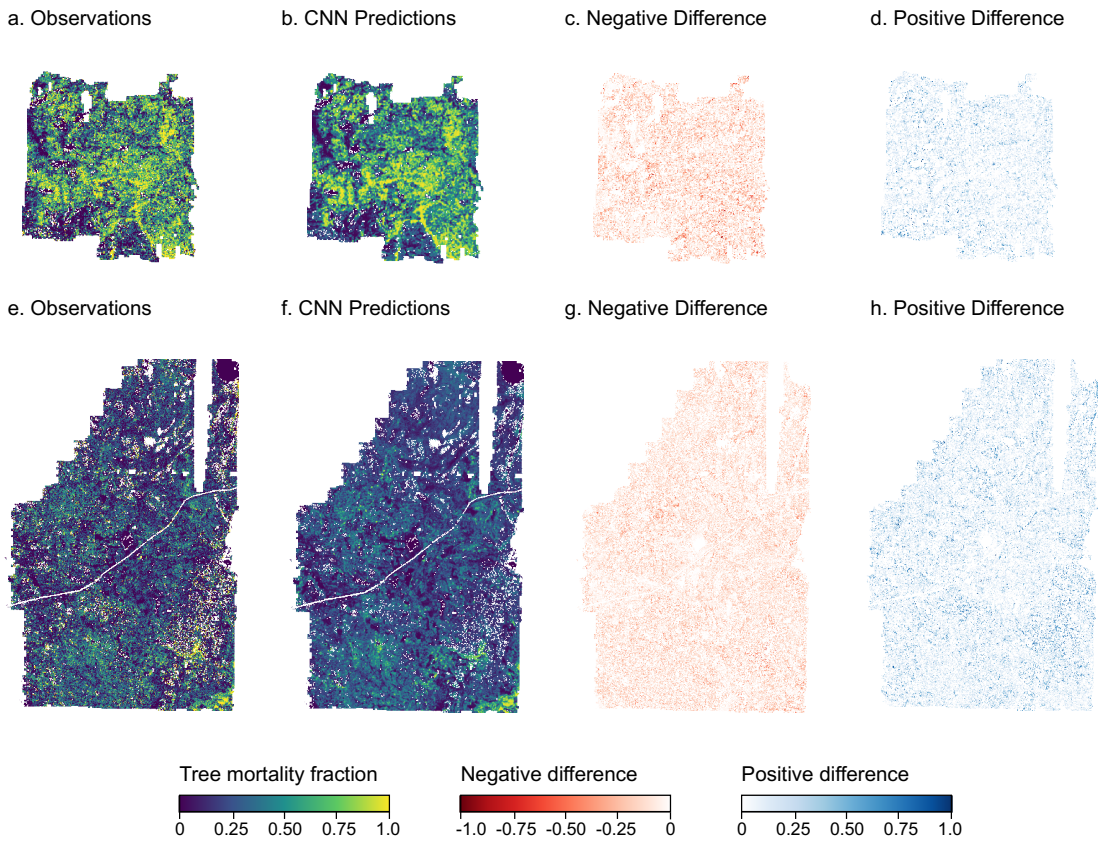

Figure 3.6. Modeled mortality fraction at Soaproot Saddle and Lower Teakettle. a. The observed mortality fraction at Soaproot Saddle. b. The predicted mortality fraction at Soaproot Saddle. c. The negative component of the difference between the observations and the predictions (i.e. overpredictions) at Soaproot Saddle. d. The positive component of the difference between the observations and the predictions (i.e. underpredictions) at Soaproot Saddle. e. The observed mortality fraction at Lower Teakettle. f. The predicted mortality fraction at Lower Teakettle. g. The negative component of the difference between the observations and the predictions (i.e. overpredictions) at Lower Teakettle. h. The positive component of the difference between the observations and the predictions (i.e. underpredictions) at Lower Teakettle.

In Figure 3.7, we plot the spread of the predicted mortality fraction for each observed mortality fraction bin for pixels with at least 4 trees. The bin with the most trees in it has a mortality fraction of zero. At low observed mortality fractions (0.2 or less), the model tends to overpredict mortality fraction. In contrast, the model underpredicts mortality for moderate to high tree mortality bins. The correlation between the mortality fraction observations and predicts is 0.66 (an R^2 of 0.44) for all the pixels and rises to 0.75 (an R^2 of 0.57) when we only consider pixels with at least four trees.

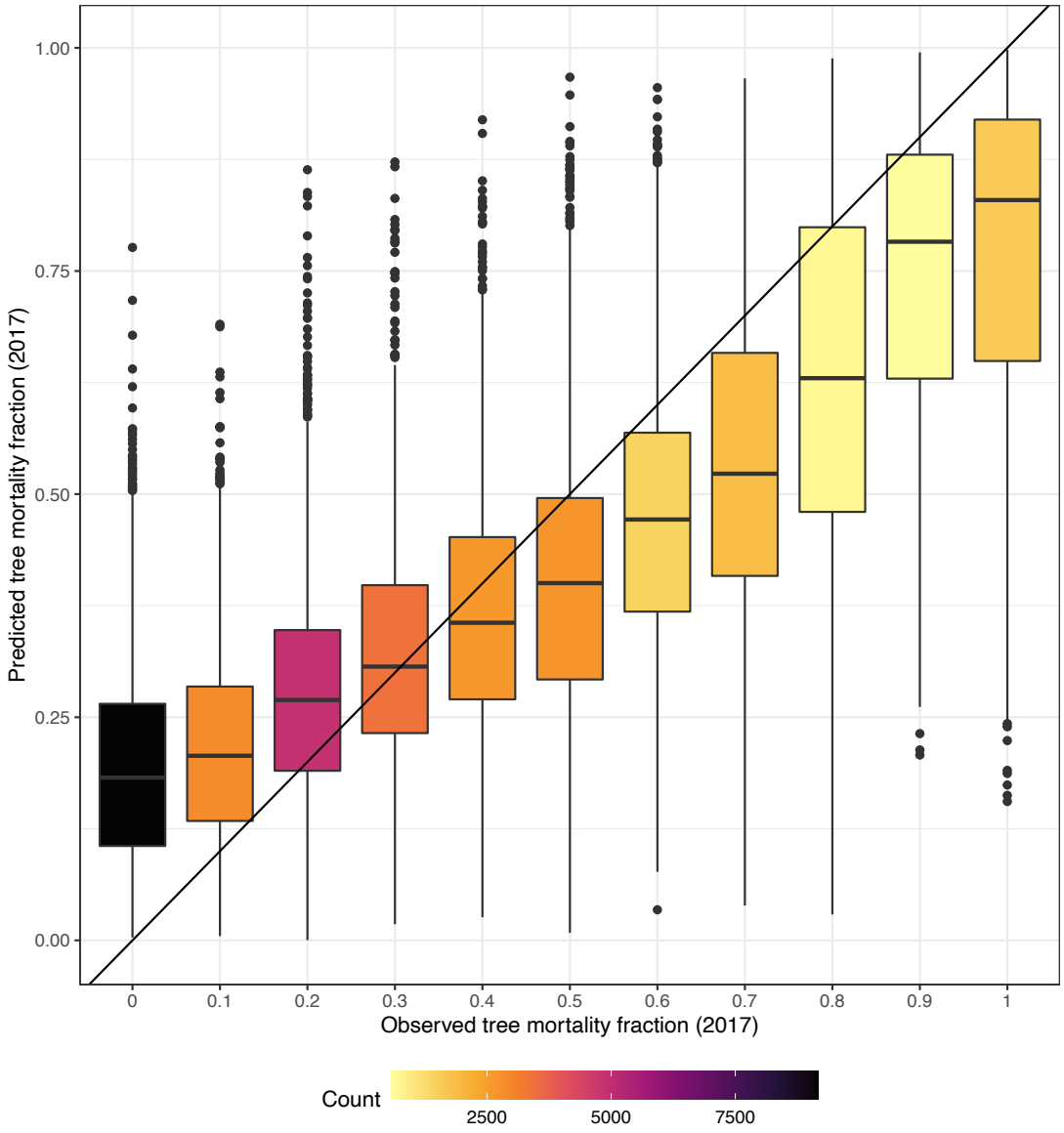

Figure 3.7. Mortality fraction in 2017 predictions vs. observations for pixels with at least four trees. We binned the observed 2017 tree mortality fraction to the nearest 10th and plotted the spread of the 2017 tree mortality fraction predictions. The line plotted on the graph is the 1:1 line, and the color of the boxes indicates the number of pixels which fall within them. Darker colors indicate more pixels.

We applied our trained CNN model to a broader domain of Landsat pixels shown in Figure 3.8 to estimate 2017 tree mortality fraction. Each grid line shown in the figure indicates 5 kilometers. The pattern of high mortality fraction observed at Soaproot Saddle is present in the gridded area to the north, east, and southeast of the study site. These areas have comparable elevation ranges as Soaproot Saddle (Figure 1.1). Similarly, estimates of tree mortality at elevation ranges comparable to Lower Teakettle to the northwest of the site show lower levels of tree mortality fraction consistent with what we observed at Lower Teakettle.

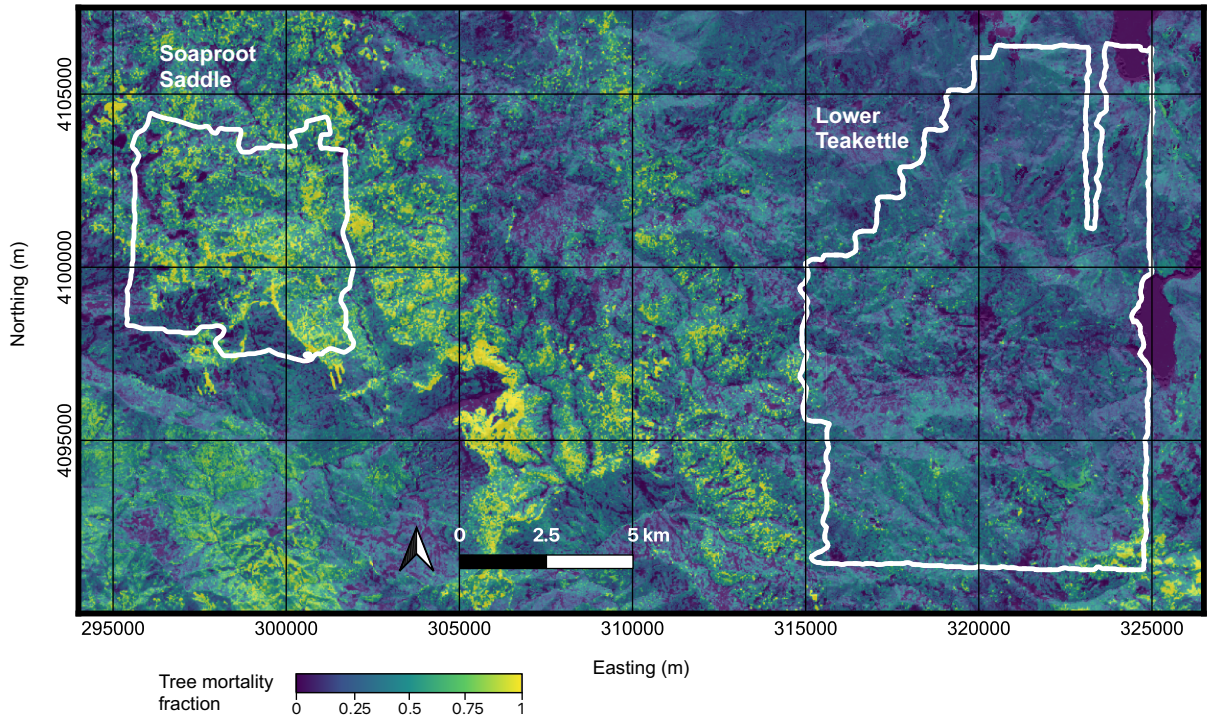

Figure 3.8 Extended tree mortality predictions for 2017. We applied our model to the time series of Landsat pixels over a broader domain. The white outlines show our study sites, and the raster data shows tree mortality fraction from 0 (dark blue) to 1 (light yellow).

3.4 Discussion

In this study, we used mortality fraction raster data from a data set of more than one million individual trees derived from airborne campaigns to model mortality fraction from Landsat time series. We found that a one-dimensional convolutional neural network trained on six bands including the relative blue, green, and red; luminosity; NDVI; and NDMI performed better than any single band. While the model overestimates low mortality and underestimates moderate to high mortality, the spatial structure of the model is well-represented.

Previous studies have shown that the change in late summer NDMI over a drought period is correlated with tree mortality (Goulden & Bales, 2019). In Chapter 1, we showed that the correlation between Landsat-derived late summer NDMI from 2013 to 2017 with the change in mortality fraction over the same time period has an R^2 value of 0.36 for pixels with at least 4 trees. In this study, we found an R^2 between our predictions and mortality fraction of 0.57 which shows that our CNN model can describe more than 50% more variation in the mortality fraction.

Our work builds upon previous studies characterizing vegetation changes after disturbance using satellite-derived time series of surface reflectance. Previous efforts focused on classification may not be able to provide the level of detail needed to characterize tree mortality from drought and insect-driven disturbance. A previous study showed that a long short term memory network can be used to estimate tree mortality fraction derived from very high resolution unmanned aerial vehicle data from the relatively high-resolution Sentinel-1 and 2 imagery (Schiefer et al., 2023). In this study, we show that

the 30-meter Landsat time series can also be used in this way with comparable correlations between the observed and predicted mortality fraction.

A key limitation of our study is that we fit a model to the 2017 tree mortality fraction which occurred after a particularly severe drought which means that we do not know how well the model would perform in estimating tree mortality at other times or for less severe disturbances. Future work might focus on assessing how well such a model can be generalized to additional time points and tree mortality scenarios. Nevertheless, the model we generated in this study can be used to expand maps of tree mortality in the Sierra Nevada to better characterize the extent of tree mortality during the 2012-2016 drought.

The 2012-2016 California drought may be indicative of future droughts as the climate warms and precipitation patterns change (Crockett & Westerling, 2018). Changes in temperature and precipitation may already exceeded the limits of forests in hotter, lower elevation forests of the Sierra Nevada (Hill et al., 2023). This means that large-scale mortality in these regions could lead to large-scale ecosystem shifts. Leveraging high-resolution tree mortality datasets to expand the domain of drought-driven disturbance maps may help us to characterize the changes that have occurred and characterize consequences for carbon cycling and ecosystem services now and in the future.

3.5 Conclusion

The 2012-2016 California drought was particularly severe with widespread consequences for California forests. We showed that a one-dimensional convolutional

neural network can be used to estimate tree mortality fraction after a severe drought from the Landsat time series with a mean absolute error of about 0.17. Our model overestimates low tree mortality and underestimates moderate-to-severe mortality fraction. However, the correlation between observed and predicted mortality fraction is 0.66 before filtering for pixels with no trees indicating that we can get a relatively good fit with little preprocessing. Our work provides a means to estimate tree mortality fraction in the Sierra Nevada after the 2012-2016 drought on a much larger spatial scale than we could achieve from currently available and public air campaign data. This model may help to characterize tree mortality across a wider domain within the Sierra Nevada to improve our understanding of the disturbance after this severe drought and the consequences for ecosystem structure and function.

CHAPTER 4

Summary, Future Directions, and Conclusions

4.1 Summary

The 2012-2016 drought in California was particularly severe and associated with widespread tree mortality in the Sierra Nevada. Throughout the investigations of this dissertation, we aimed to characterize tree mortality at the scale of individual trees, model tree mortality risk at the scale of individual trees, and extend our results to a wider domain within the Sierra Nevada by linking our individual tree mortality results to the Landsat time series.

In Chapter 1, we used NEON lidar and multispectral reflectance airborne observations to map individual tree mortality over a 160 km^2 area during and after the 2012-2016 drought for two sites in California's Sierra National Forest. We used NEON lidar to derive tree locations and crown perimeters and multispectral data to map tree mortality for more than one million trees. We found that about one quarter of the trees at Lower Teakettle were dead by 2017 and about half of the trees were dead at Soaproot Saddle. Two wildfires in 2020 and 2021 doubled tree mortality within wildfire perimeters and killed almost all trees within the 2021 Blue Fire perimeter. Finally, we compared our result to a previous study and found that the NDVI signature of trees we labeled as alive and the previous study labeled as dead after the drought was more consistent with live trees from both studies.

Consistent with previous work, we found that tree mortality risk increased as a function of tree height. Tree mortality was positively associated with distance from rivers, trees per hectare, and decreasing slope at the lower elevation site. In contrast, increasing slope was positively associated with tree mortality at the higher elevation site. To further develop our analysis of the drivers of tree mortality, we used extreme gradient boosting to estimate tree mortality risk from plant traits prior to the large tree mortality event, topography, and climate variables in Chapter 2. We found that tree mortality at the low elevation site, Soaproot Saddle, was much easier to predict. Models of tree mortality at Soaproot Saddle reached 74% accuracy on our validation dataset compared with 61% at Lower Teakettle compared with a base rate of 50% for random guessing. Consistent with our findings in Chapter 1, tree height was the most important predictor for models at both sites. At Soaproot Saddle, the second and third most important feature variables were

distance to rivers and canopy cover fraction, respectively. In contrast, Lower Teakettle predictions were more dependent on baseline NDVI and the strength of the change in dry season temperature during the drought. Our results show that tree mortality risk may be easier to predict in warmer and drier, low-elevation regions of the Sierra Nevada. A wider domain of training data may help to create more robust models of tree mortality.

In Chapter 3, we aimed to generate a wider domain of tree mortality data by linking the individual tree mortality data set from Chapter 1 to the Landsat time series. We downloaded Landsat scenes from Landsat 5 and Landsat 8 to collect a time series of surface reflectance data between 1984 and 2017 to capture longer term patterns that may contribute to tree mortality risk. We fit a one-dimensional convolutional neural network to estimate 2017 tree mortality fraction from the time series of six channels including the visible bands normalized for luminosity, luminosity, normalized difference vegetation index, and normalized difference moisture index. For predicting mortality fraction, we found a mean absolute error of 0.16 on the training data set and 0.17 on the validation and testing data sets. The model's Pearson correlation between the observations and predictions is 0.66 for pixels with any number of trees including those with no trees. For Landsat pixels containing at least four trees, the correlation is 0.75. Applying convolutional neural networks to Landsat imagery may help bridge the gap between high spatial resolution tree mortality data sets derived from air campaigns and region-wide estimates needed for characterizing drought and disturbance impacts on the carbon cycle and ecosystem function.

Through the investigations in this dissertation, we developed a large and publicly available data set of individual tree mortality to characterize tree mortality after a

particularly severe drought and subsequent wildfire. We then used this data set to analyze biophysical drivers of tree mortality risk and to create broader-scale maps of tree mortality fraction from Landsat data. These analyses, methods, and data sets may help predict the impacts of severe droughts on coniferous forests in the future.

4.2 Future Directions

One future direction for this work would be to use the expanded tree mortality map from Chapter 3 to investigate tree mortality risk factors over a wider domain. In Chapter 2, we found that our tree mortality model trained on one site did not perform well at the other. By training a tree mortality risk model on a wider domain of the feature variables, we may be able to uncover more comprehensive relationships between vegetation, topography, and climate and tree mortality risk. A second direction is to generalize the model result from Chapter 3 to additional time points to create a spatiotemporal dataset to support forest monitoring and conservation efforts and time-evolving process-based models operating at large scales.

4.3 Conclusions

We mapped more than one million trees with 92% test accuracy in the Southern Sierra Nevada, a region which experienced some of the highest rates of tree mortality after the 2012-2016 California drought. We found 50% tree mortality at the lower elevation site and found that we could model individual tree mortality risk using extreme gradient

boosting with an accuracy of 74%. Tree mortality risk increased with tree height, distance to rivers, and on shallower slopes. While tree mortality at the higher elevation site was more challenging to predict, we found that using the class likelihood from the classification predictions captured the general underlying pattern of tree mortality at both sites. Finally, we linked our individual tree mortality data set to the Landsat time series using a convolutional neural network. We found that our predictions correlated well with the observations at both sites indicating that we may be able to generalize this model to a wider domain. A wider domain of tree mortality data may help to elucidate generalizable patterns in tree mortality drivers during drought at regional scales to inform process-based models and the larger-scale Earth system models which use them. As drought may become more frequent and severe, these data sets and insights may help inform predictions of carbon fluxes from the biosphere in the Anthropocene.

REFERENCES

Adams, H. D., Guardiola-Claramonte, M., Barron-Gafford, G. A., Villegas, J. C., Breshears, D. D., Zou, C. B., et al. (2009). Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. *Proceedings of the National Academy of Sciences, 106*(17), 7063-7066.

Adams, H. D., Zeppel, M. J., Anderegg, W. R. L., Hartmann, H., Landhäusser, S. M., Tissue, D. T., et al. (2017). A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nature ecology & evolution, 1*(9), 1285-1291.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., et al. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management, 259*(4), 660-684.

Anderegg, W. R. L., Flint, A., Huang, C., Flint, L., Berry, J. A., Davis, F. W., et al. (2015a). Tree mortality predicted from drought-induced vascular damage. *Nature Geoscience, 8*(5), 367-371.

Anderegg, W. R. L., Hicke, J. A., Fisher, R. A., Allen, C. D., Aukema, J., Bentz, B., et al. (2015b). Tree mortality from drought, insects, and their interactions in a changing climate. *New Phytologist, 208*(3), 674-683.

Anderegg, W. R. L., Trugman, A. T., Badgley, G., Anderson, C. M., Bartuska, A., Ciais, P., et al. (2020). Climate-driven risks to the climate mitigation potential of forests. *Science, 368*(6497), eaaz7005.

Andrus, R. A., Chai, R. K., Harvey, B. J., Rodman, K. C., & Veblen, T. T. (2021). Increasing rates of subalpine tree mortality linked to warmer and drier summers. *Journal of Ecology, 109*(5), 2203-2218.

Asner, G. P., Brodrick, P. G., Anderson, C. B., Vaughn, N., Knapp, D. E., & Martin, R. E. (2016). Progressive forest canopy water loss during the 2012–2015 California drought. *Proceedings of the National Academy of Sciences, 113*(2), E249-E255.

Byer, S., & Jin, Y. (2017). Detecting Drought-Induced Tree Mortality in Sierra Nevada Forests with Time Series of Satellite Data. *Remote Sensing, 9*(9), 929.

Cai, W., Ng, B., Geng, T., Jia, F., Wu, L., Wang, G., et al. (2023). Anthropogenic impacts on twentieth-century ENSO variability changes. *Nature Reviews Earth & Environment, 4*(6), 407-418. https://doi.org/10.1038/s43017-023-00427-8

Carter, G. A. (1991). Primary and secondary effects of water content on the spectral reflectance of leaves. *American journal of botany, 78*(7), 916-924.

Chen, T., & Guestrin, C. (2015). *Xgboost: Reliable large-scale tree boosting system.* Paper presented at the Proceedings of the 22nd SIGKDD Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA.

Chollet, F., & Allaire, J. (2018). *Deep learning with R*. Shelter Island, NY: Manning Publications.

Coleman, T. W., Graves, A. D., Heath, Z., Flowers, R. W., Hanavan, R. P., Cluck, D. R., & Ryerson, D. (2018). Accuracy of aerial detection surveys for mapping insect and disease disturbances in the United States. *Forest Ecology and Management, 430*, 321-336. https://www.sciencedirect.com/science/article/pii/S0378112718311459

Crockett, J. L., & Westerling, A. L. (2018). Greater Temperature and Precipitation Extremes Intensify Western U.S. Droughts, Wildfire Severity, and Sierra Nevada Tree Mortality. *Journal of Climate, 31*(1), 341-354. https://journals.ametsoc.org/view/journals/clim/31/1/jcli-d-17-0254.1.xml

Dalponte, M., & Coomes, D. A. (2016). Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods in ecology and evolution, 7*(10), 1236-1245.

Das, A. J., Stephenson, N. L., & Davis, K. P. (2016). Why do trees die? Characterizing the drivers of background tree mortality. *Ecology, 97*(10), 2616-2627.

Deng, L. (2012). The MNIST database of handwritten digit images for machine learning research [best of the web]. *IEEE signal processing magazine, 29*(6), 141-142.

Fettig, C. J., Mortenson, L. A., Bulaon, B. M., & Foulk, P. B. (2019). Tree mortality following drought in the central and southern Sierra Nevada, California, US. *Forest Ecology and Management, 432*, 164-178.

Geng, T., Jia, F., Cai, W., Wu, L., Gan, B., Jing, Z., et al. (2023). Increased occurrences of consecutive La Niña events under global warming. *Nature, 619*(7971), 774-781. https://doi.org/10.1038/s41586-023-06236-9

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18-27.

Goulden, M. L., Anderson, R., Bales, R., Kelly, A., Meadows, M., & Winston, G. (2012). Evapotranspiration along an elevation gradient in California's Sierra Nevada. *Journal of Geophysical Research: Biogeosciences, 117*(G3).

Goulden, M. L., & Bales, R. C. (2019). California forest die-off linked to multi-year deep soil drying in 2012–2015 drought. *Nature Geoscience, 12*(8), 632-637.

Hemming-Schroeder, N., Gutierrez, A., Allison, S., & Randerson, J. (2023a). *Data and code from: Estimating Individual Tree Mortality in the Sierra Nevada Using Lidar and Multispectral Reflectance Data*.

Hemming-Schroeder, N., Gutierrez, A., Allison, S., & Randerson, J. (2023b). Estimating Individual Tree Mortality in the Sierra Nevada Using Lidar and Multispectral Reflectance Data. *Journal of Geophysical Research: Biogeosciences*, e2022JG007234. https://doi.org/10.1029/2022JG007234

Hill, A. P., Nolan, C. J., Hemes, K. S., Cambron, T. W., & Field, C. B. (2023). Low-elevation conifers in California's Sierra Nevada are out of equilibrium with climate. *PNAS Nexus, 2*(2). https://doi.org/10.1093/pnasnexus/pgad004

Hoerling, M., Eischeid, J., Kumar, A., Leung, R., Mariotti, A., Mo, K., et al. (2014). Causes and Predictability of the 2012 Great Plains Drought. *Bulletin of the American Meteorological Society, 95*(2), 269-282. https://journals.ametsoc.org/view/journals/bams/95/2/bams-d-13-00055.1.xml

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]*. Retrieved from https://www.ipcc.ch/report/ar5/syr/

Karpowics, B., & Kampe, T. (2022). NEON Imaging Spectrometer Radiance to Reflectance Algorithm Theoretical Basis Document (NEON.DOC.001288).

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment, 114*(12), 2897-2910. https://www.sciencedirect.com/science/article/pii/S0034425710002245

Kitzberger, T., Brown, P. M., Heyerdahl, E. K., Swetnam, T. W., & Veblen, T. T. (2007). Contingent Pacific–Atlantic Ocean influence on multicentury wildfire synchrony over western North America. *Proceedings of the National Academy of Sciences, 104*(2), 543-548. https://doi.org/10.1073/pnas.0606078104

Kolus, H. R., Huntzinger, D. N., Schwalm, C. R., Fisher, J. B., McKay, N., Fang, Y., et al. (2019). Land carbon models underestimate the severity and duration of drought's impact on plant productivity. *Scientific reports, 9*(1), 2758. https://doi.org/10.1038/s41598-019-39373-1

Koontz, M. J., Latimer, A. M., Mortenson, L. A., Fettig, C. J., & North, M. P. (2021). Cross-scale interaction of host tree size and climatic water deficit governs bark beetle-induced tree mortality. *Nature communications, 12*(1), 129.

Krause, K., & Goulden, T. (2022). NEON L0-To-L1 Discrete Return LiDAR Algorithm Theoretical Basis Document (NEON.DOC.001292).

Krauss, R. M., Courtney; Patterson, Michael; Smith, Oliver. (2018). *TOS Site Characterization Report: Domain 17*. Retrieved from https://data.neonscience.org/documents

Liang, L., Chen, Y., Hawbaker, T. J., Zhu, Z., & Gong, P. (2014). Mapping Mountain Pine Beetle Mortality through Growth Trend Analysis of Time-Series Landsat Data. *Remote Sensing, 6*(6), 5696-5716. https://www.mdpi.com/2072-4292/6/6/5696
McDowell, N. G., Williams, A., Xu, C., Pockman, W., Dickman, L., Sevanto, S., et al. (2016). Multi-scale predictions of massive conifer mortality due to chronic temperature rise. *Nature climate change, 6*(3), 295-300.
NEON. (2021). Discrete return LiDAR point cloud, RELEASE-2021 (DP1.30003.001). https://doi.org/10.48443/ca22-1n03
NEON. (2022a). Ecosystem structure, RELEASE-2022 (DP3.30015.001). https://doi.org/10.48443/hd92-rr80
NEON. (2022b). Spectrometer orthorectified surface directional reflectance - mosaic, RELEASE-2022 (DP1.30006.001). https://doi.org/10.48443/5tfk-8e74
NEON. (2023). Elevation - LiDAR, RELEASE-2023 (DP3.30024.001). https://doi.org/10.48443/kvb6-4322
Norlen, C. A., & Goulden, M. L. (2023). Recent Tree Mortality Dampens Semi-Arid Forest Die-Off During Subsequent Drought. *AGU Advances, 4*(3), e2022AV000810.
Okumura, Y. M., DiNezio, P., & Deser, C. (2017). Evolving Impacts of Multiyear La Niña Events on Atmospheric Circulation and U.S. Drought. *Geophysical Research Letters, 44*(22), 11,614-611,623. https://doi.org/10.1002/2017GL075034
Paz-Kagan, T., Brodrick, P. G., Vaughn, N. R., Das, A. J., Stephenson, N. L., Nydick, K. R., & Asner, G. P. (2017). What mediates tree mortality during drought in the southern Sierra Nevada? *Ecological Applications, 27*(8), 2443-2457.
QGIS Development Team. (2023). QGIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved from http://qgis.osgeo.org
R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
Raffa, K. F., Aukema, B. H., Bentz, B. J., Carroll, A. L., Hicke, J. A., Turner, M. G., & Romme, W. H. (2008). Cross-scale Drivers of Natural Disturbances Prone to Anthropogenic Amplification: The Dynamics of Bark Beetle Eruptions. *BioScience, 58*(6), 501-517. https://doi.org/10.1641/B580607
Restaino, C., Young, D. J., Estes, B., Gross, S., Wuenschel, A., Meyer, M., & Safford, H. (2019). Forest structure and climate mediate drought-induced tree mortality in forests of the Sierra Nevada, USA. *Ecological Applications, 29*(4), e01902.
Rippey, B. R. (2015). The U.S. drought of 2012. *Weather and Climate Extremes, 10*, 57-64. https://www.sciencedirect.com/science/article/pii/S2212094715300360
Robbins, Z. J., Xu, C., Aukema, B. H., Buotte, P. C., Chitra-Tarak, R., Fettig, C. J., et al. (2021). Warming increased bark beetle-induced tree mortality by 30% during an extreme drought in California. *Global Change Biology*.
Robbins, Z. J., Xu, C., Jonko, A., Chitra-Tarak, R., Fettig, C. J., Costanza, J., et al. (2023). Carbon stored in live ponderosa pines in the Sierra Nevada will not return to pre-drought (2012) levels during the 21st century due to bark beetle outbreaks. *Frontiers in Environmental Science, 11*, 281.
Schiefer, F., Schmidtlein, S., Frick, A., Frey, J., Klinke, R., Zielewska-Büttner, K., et al. (2023). UAV-based reference data for the prediction of fractional cover of standing deadwood from Sentinel time series. *ISPRS Open Journal of Photogrammetry and*

Remote Sensing, 8, 100034. https://www.sciencedirect.com/science/article/pii/S2667393223000054
Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., & Ludwig, C. (2015). The trajectory of the Anthropocene: The Great Acceleration. *The Anthropocene Review, 2*(1), 81-98. https://doi.org/10.1177/2053019614564785
Stephenson, N. L., & Das, A. J. (2020). Height-related changes in forest composition explain increasing tree mortality with height during an extreme drought. *Nature communications, 11*(1), 1-4.
Stephenson, N. L., Das, A. J., Ampersee, N. J., Bulaon, B. M., & Yee, J. L. (2019). Which trees die during drought? The key role of insect host-tree selection. *Journal of Ecology, 107*(5), 2383-2401. https://doi.org/10.1111/1365-2745.13176
Stovall, A. E. (2019). *CA_lidar_tree_mortality (Version 2)*. Retrieved from: https://doi.org/10.6084/m9.figshare.7609193.v2
Stovall, A. E., Shugart, H., & Yang, X. (2019). Tree height explains mortality risk during an intense drought. *Nature communications, 10*(1), 1-6.
Trugman, A. T. (2021). Integrating plant physiology and community ecology across scales through trait-based models to predict drought mortality. *New Phytologist*.
Trugman, A. T., Anderegg, L. D. L., Anderegg, W. R. L., Das, A. J., & Stephenson, N. L. (2021). Why is tree drought mortality so hard to predict? *Trends in Ecology & Evolution, 36*(6), 520-532.
U.S. Geological Survey. (2019). USGS National Hydrography Dataset Plus High Resolution (NHDPlus HR) for 4-digit Hydrologic Unit - 1803.
United States Forest Service Aerial Survey Program. (2023). *Aerial Detection Monitoring*. Retrieved from: https://www.fs.usda.gov/detail/r5/forest-grasslandhealth
USDA. (2017). Record 129 Million Dead Trees in California: USDA Forest Service and CAL FIRE Working Together to Address Forest Health, Press release, December 12, 2017. https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fseprd566303.pdf
USDA Forest Service National Forest System Lands GIS and Fire personnel. (2023). *National USFS Final Fire Perimeter*. Retrieved from: https://data.fs.usda.gov/geodata/edw/datasets.php
Vilanova, E., Mortenson, L. A., Cox, L. E., Bulaon, B. M., Lydersen, J. M., Fettig, C. J., et al. (2023). Characterizing ground and surface fuels across Sierra Nevada forests shortly after the 2012–2016 drought. *Forest Ecology and Management, 537*, 120945.
Weinstein, B. G., Marconi, S., Bohlman, S. A., Zare, A., Singh, A., Graves, S. J., & White, E. P. (2021). A remote sensing derived data set of 100 million individual tree crowns for the National Ecological Observatory Network. *eLife, 10*, e62922. https://doi.org/10.7554/eLife.62922
Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., Pongratz, J., Bloom, A. A., et al. (2021). Changes in global terrestrial live biomass over the 21st century. *Science Advances, 7*(27), eabe9829.
Ye, S., Rogan, J., Zhu, Z., & Eastman, J. R. (2021). A near-real-time approach for monitoring forest disturbance using Landsat time series: stochastic continuous change detection. *Remote Sensing of Environment, 252*, 112167. https://www.sciencedirect.com/science/article/pii/S003442572030540X
Young, D. J. N., Stevens, J. T., Earles, J. M., Moore, J., Ellis, A., Jirka, A. L., & Latimer, A. M. (2017). Long-term climate and competition explain forest mortality patterns under extreme drought. *Ecology letters, 20*(1), 78-86. https://doi.org/10.1111/ele.12711

Yuan, X., Wang, Y., Ji, P., Wu, P., Sheffield, J., & Otkin, J. A. (2023). A global transition to flash droughts under climate change. *Science, 380*(6641), 187-191. https://www.science.org/doi/abs/10.1126/science.abn6301

Zald, H. S., Callahan, C. C., Hurteau, M. D., Goodwin, M. J., & North, M. P. (2022). Tree growth responses to extreme drought after mechanical thinning and prescribed fire in a Sierra Nevada mixed-conifer forest, USA. *Forest Ecology and Management, 510*, 120107.

Zellweger, F., De Frenne, P., Lenoir, J., Vangansbeke, P., Verheyen, K., Bernhardt-Römermann, M., et al. (2020). Forest microclimate dynamics drive plant responses to warming. *Science, 368*(6492), 772-775.

Zhang, H., Zimba, P. V., & Nzewi, E. U. (2019). A new pseudoinvariant near-infrared threshold method for relative radiometric correction of aerial imagery. *Remote Sensing, 11*(16), 1931.

Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment, 144*, 152-171. https://www.sciencedirect.com/science/article/pii/S0034425714000248

Zhu, Z., Zhang, J., Yang, Z., Aljaddani, A. H., Cohen, W. B., Qiu, S., & Zhou, C. (2020). Continuous monitoring of land disturbance based on Landsat time series. *Remote Sensing of Environment, 238*, 111116. https://www.sciencedirect.com/science/article/pii/S0034425719301002

APPENDIX A

Supporting Figures and Tables for Chapter 1

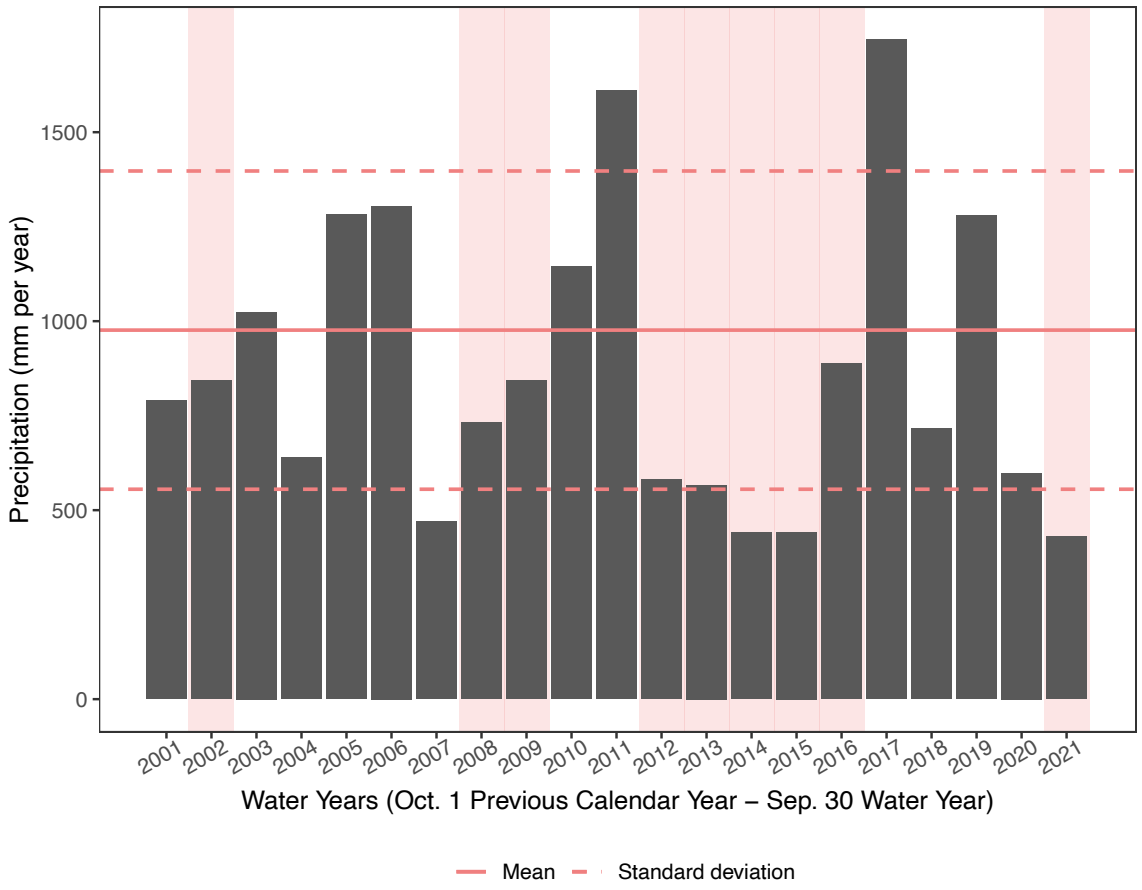

Figure A.1. The total annual water-year precipitation (mm integrated over the water-year) from PRISM with a drought index derived from data from the National Integrated Drought System. Each year is shown with a gray bar for a rectangular region equivalent to the bounding box for the Soaproot Saddle and Lower Teakettle field sites for water-years 2001-2021. The water-year for a given year goes from October 1st of the previous year through September 30th of the given year. For example, the water-year for 2020 goes from October 1, 2019, through September 30, 2020. The solid line represents the mean precipitation for the entire PRISM timeseries from 1982-2021, while the dashed line

represents one standard deviation from the mean. The shaded boxes highlight water-years of severe drought, which means at least half of Fresno County (which contains our study site shown in Figure 1.1) experienced severe drought or worse as defined by the National Integrated Drought System for at least half of the year. Because this data set begins in 2000, we only show years for which we have drought severity data for the entire water-year (2001-2021).

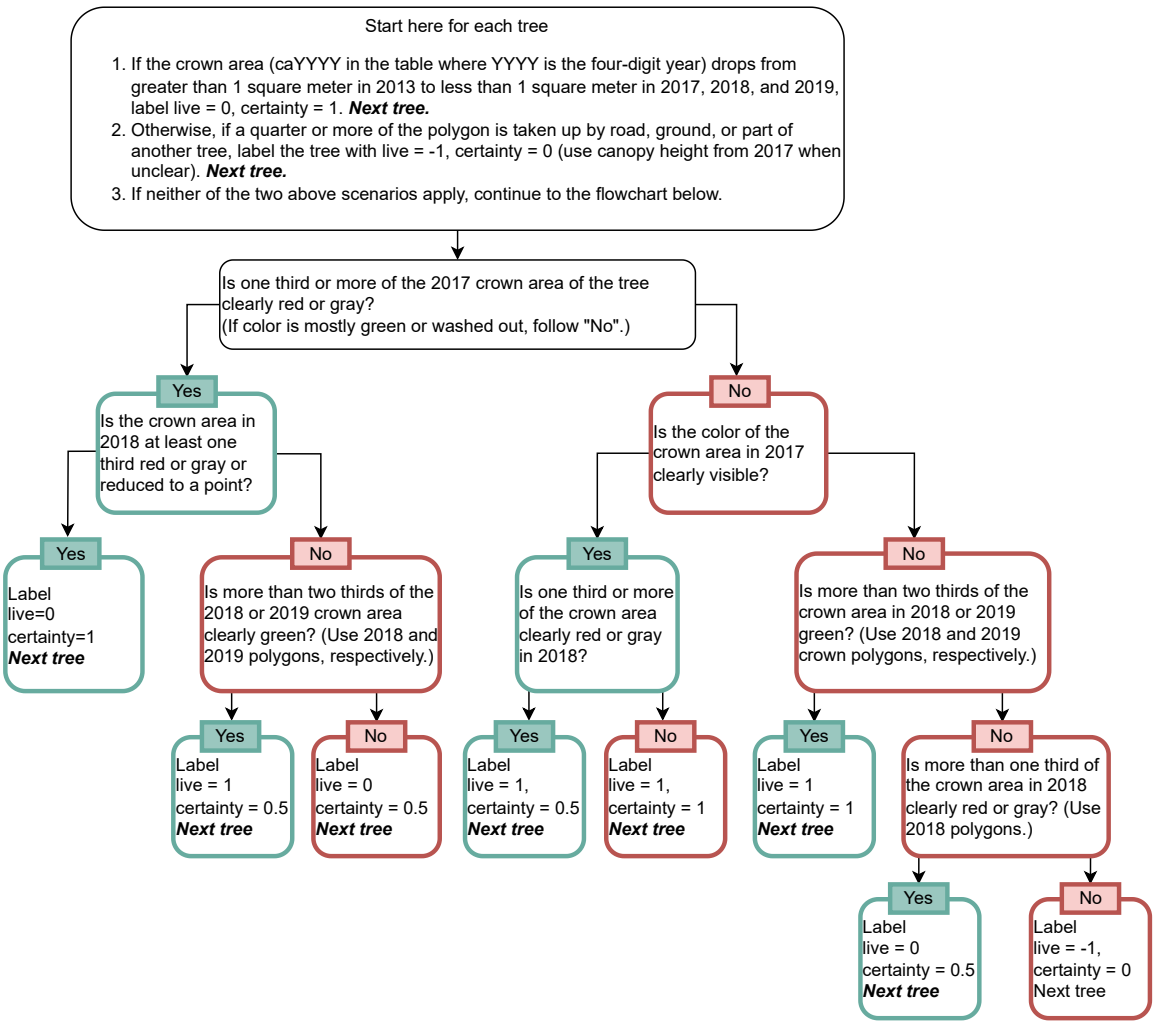

Figure A.2. Training data flowchart for manually labeling trees. We created this flowchart to label trees as dead or alive in 2017. To make the labels as consistent as possible, we created pathways that can be answered as yes (green pathway) or no (red pathway). The certainty is assigned 0.5 if the status of a tree needed to be confirmed by viewing the tree in a following year. A live label of -1 represents the lowest amount of certainty (certainty=0), where we chose to give no live or dead label to a tree. Only the live labels (0 or 1) were used in this study.

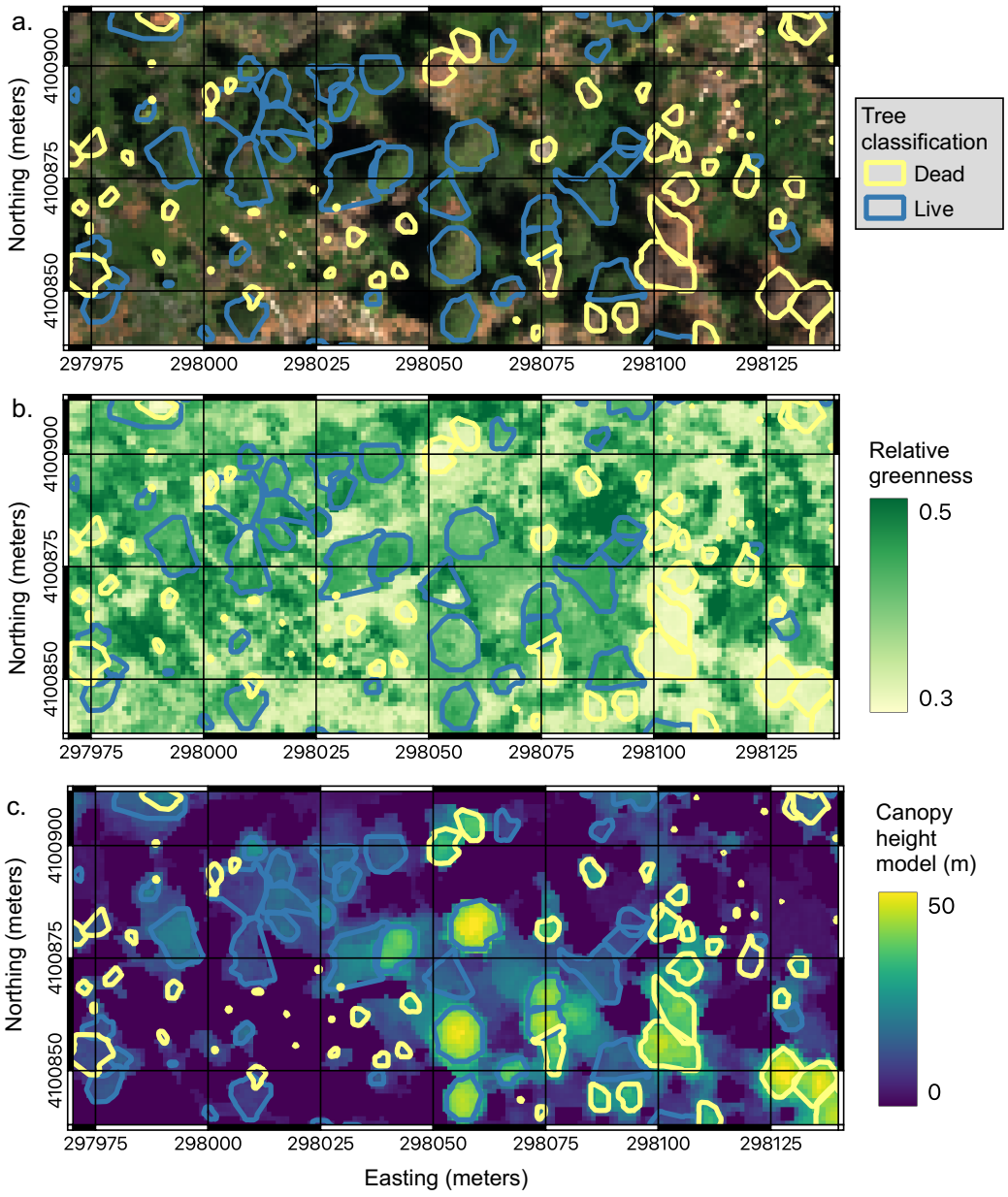

Figure A.3. Sample images with trees classified as live (medium blue) and dead (light yellow) overlaid on **a)** RGB imagery derived from NEON spectral data, **b)** relative greenness derived from NEON spectral data and **c)** the NEON canopy height model raster data from 2017.

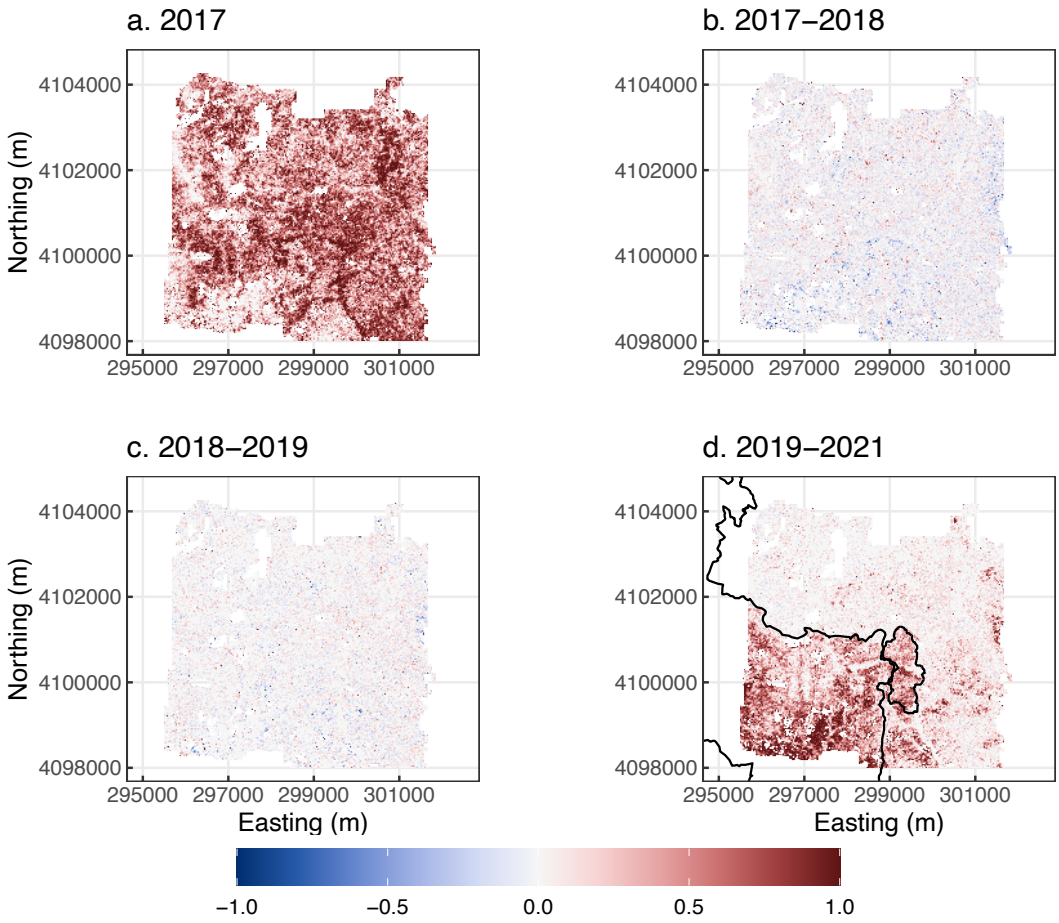

Figure A.4. The individual tree mortality for the Soaproot Saddle is aggregated to the 30-meter resolution Landsat grid. **a.** Tree mortality for 2017. **b.** The additional mortality between 2017 and 2018. **c.** The same as in panel b but for the additional mortality between 2018 and 2019. **d.** Tree mortality between 2019 and 2021 along with the National USFS Final Fire Perimeter outlines of the 2020 Creek Fire (larger fire on the west) and the 2021 Blue Fire (smaller fire to the east of the Creek Fire perimeter).

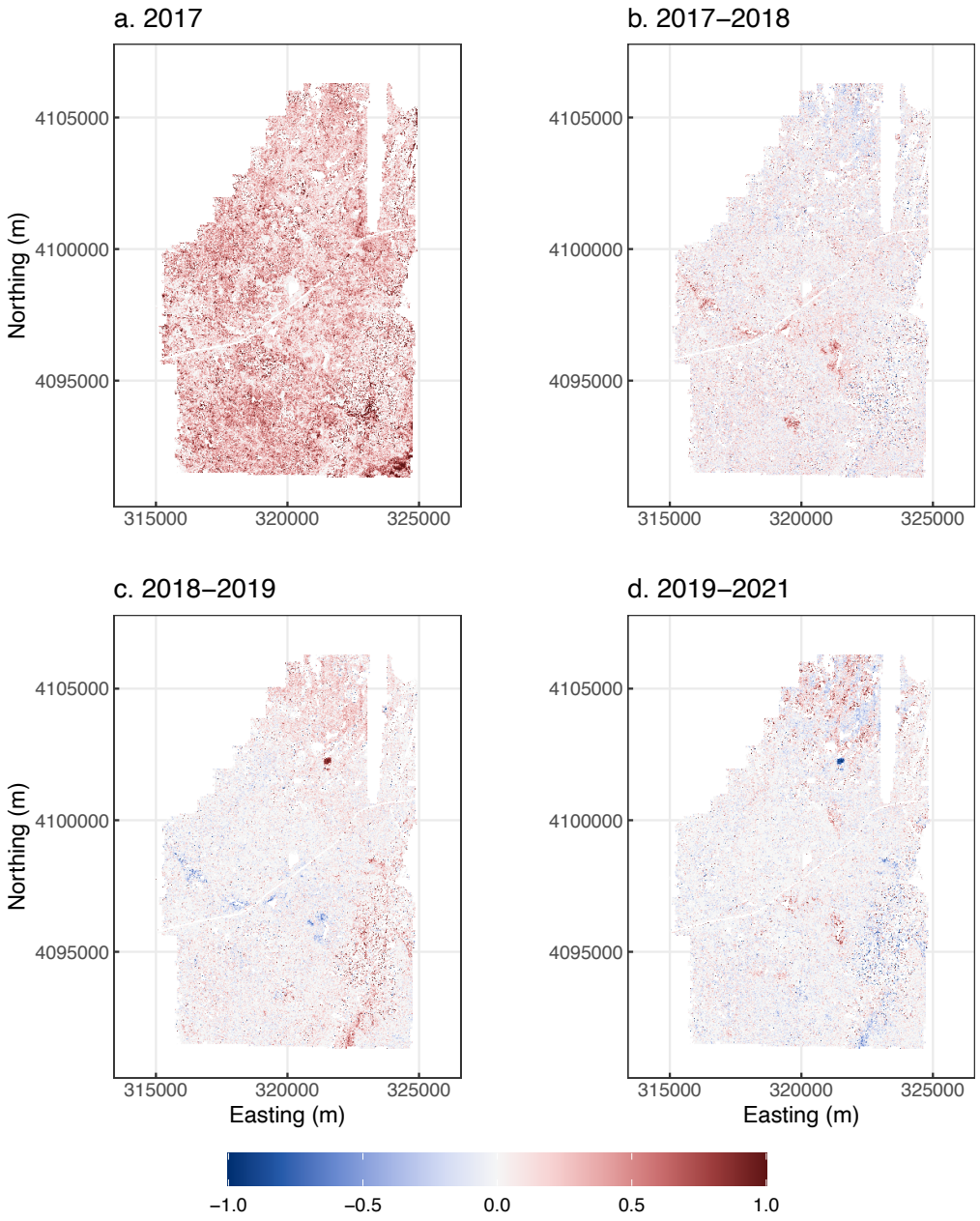

Figure A.5. The individual tree mortality for Lower Teakettle is aggregated to the 30-meter resolution Landsat grid. **a.** Tree mortality for 2017. **b.** The additional mortality between 2017 and 2018. Red pixels indicate additional mortality, whereas blue pixels indicate that trees that were classified as dead in 2017 but live in 2018. **c.** The same as in panel b but for the additional mortality between 2018 and 2019. **d.** The same as in panels b and c but for the mortality between 2019 and 2021.

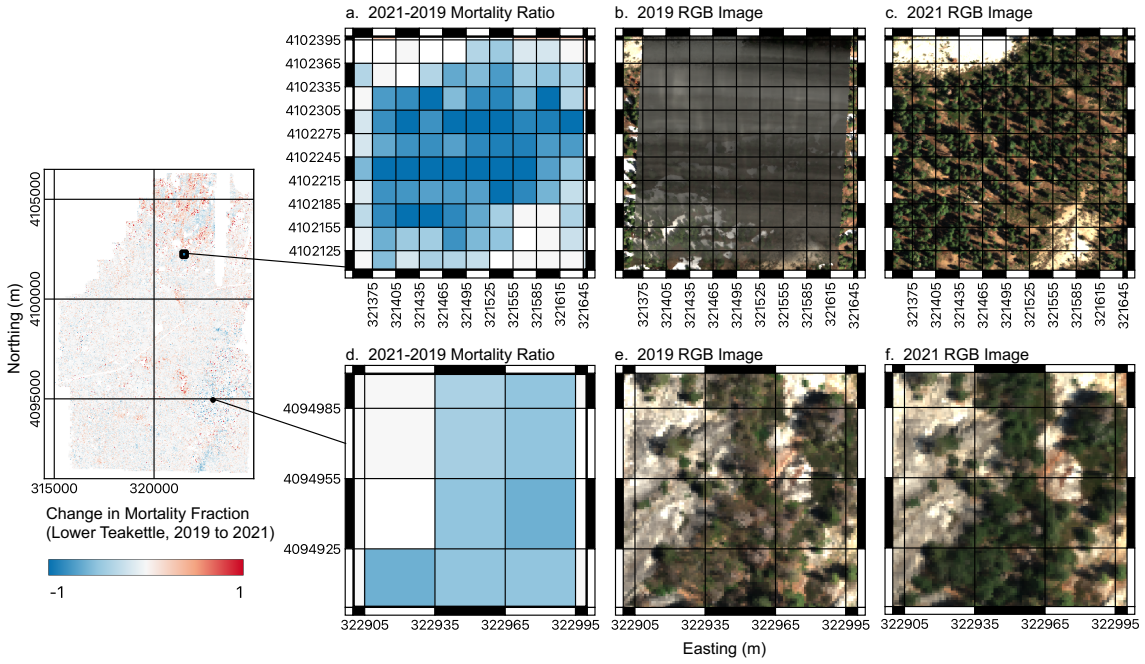

Figure A.6. Two zoomed in areas where trees were labeled dead in 2019 and live in 2021 are shown to explain possible reasons for these errors. **a.** The mortality fraction from 2019 is subtracted from the mortality fraction from 2021 for a zoomed in region in Lower Teakettle shown at right. **b.** The RGB image for the zoomed in region from panel a is shown for 2019. **c.** The RGB image for the region shown in panel b is shown for 2021. **d.** We select a second site to investigate another region of errors. The blue pixels indicate that the fractional tree mortality decreased between 2019 and 2021. **e.** This panel shows the 2019 RGB image corresponding to the region shown in panel d. **f.** The RGB image over the same region from panels d and e for 2021.

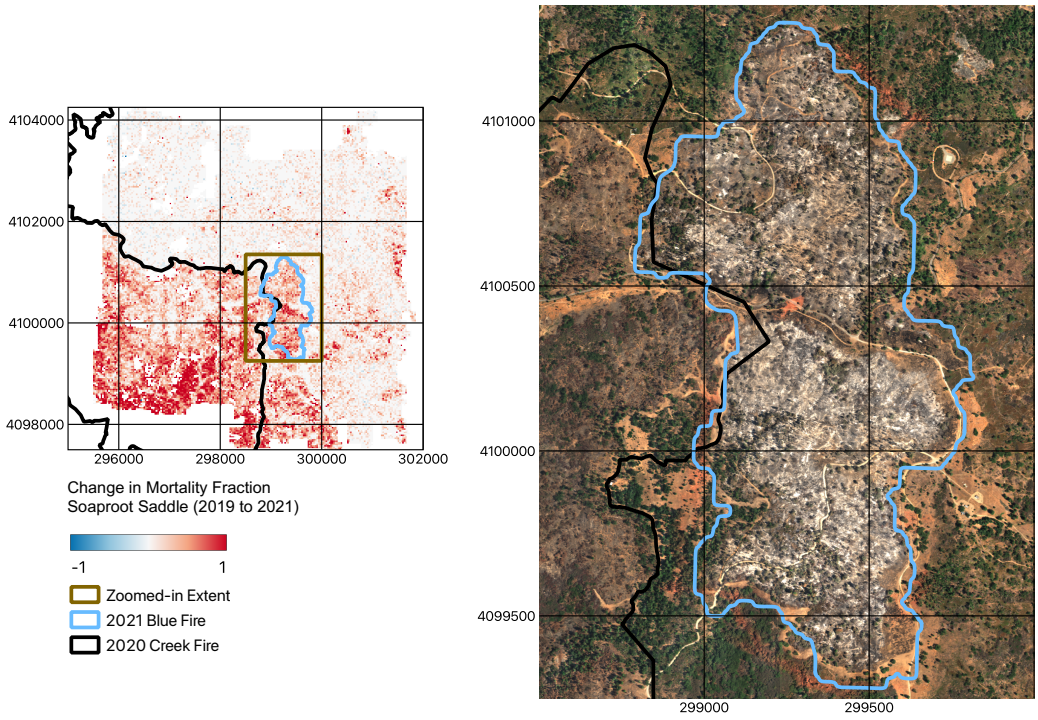

Figure A.7. Burned area from the 2021 Blue Fire. The Soaproot Saddle region shown in Figure 1.4d is shown on the left with a zoomed in region (yellow box) at the right with summer 2021 RGB spectral data derived from NEON hyperspectral datasets in the background. The black polygon shows the perimeter for the 2020 Creek Fire, while the light blue polygon shows the perimeter of the 2021 Blue Fire. Both fire perimeters are from the National USFS Final Perimeter dataset. Orange fire retardant used to contain the fire is still visible in the imagery, particularly along the northeastern and southern portions of the fire perimeter.

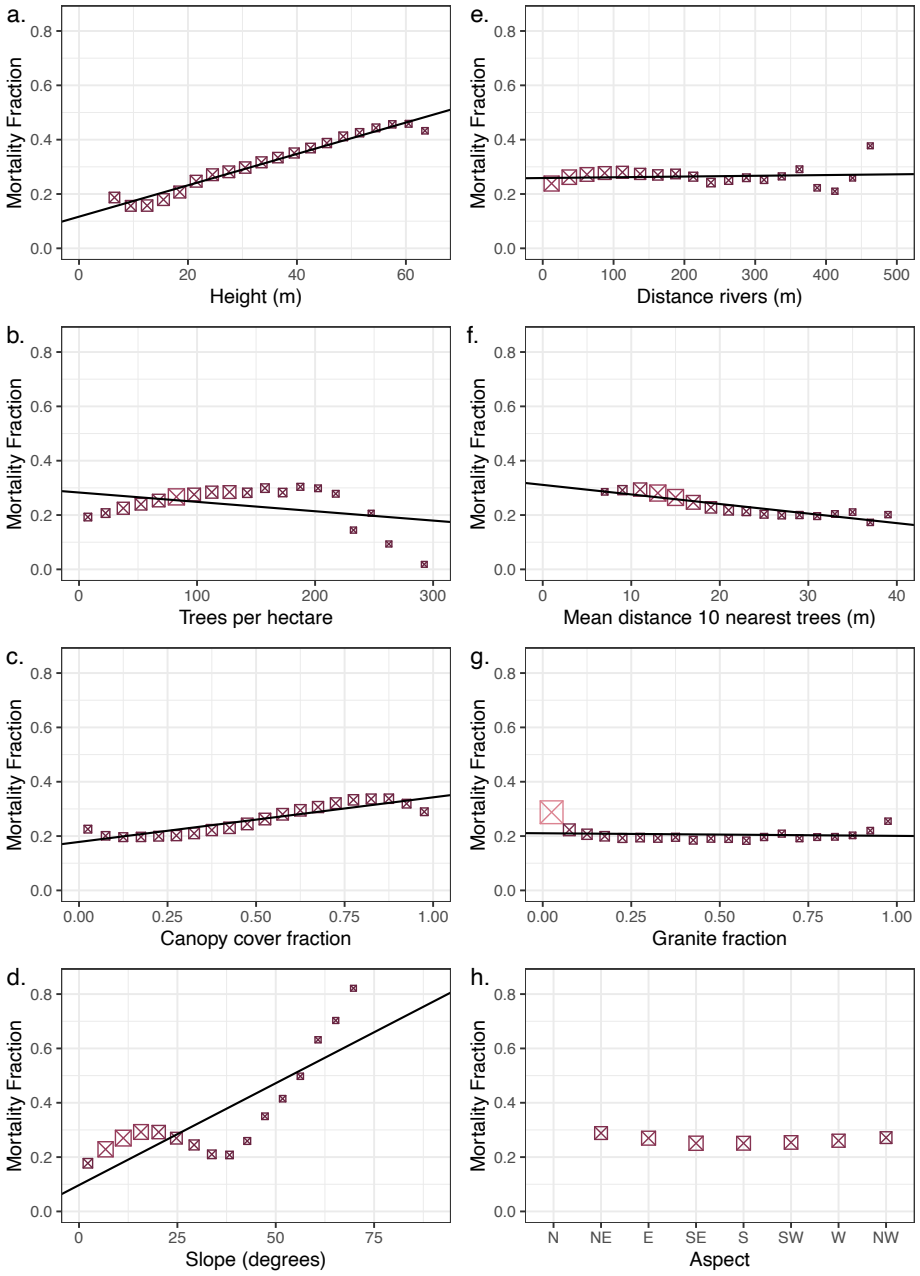

Figure A.8. Potential environmental drivers of 2017 tree mortality across the full study domain (including Soaproot Saddle and Lower Teakettle). We restricted our analyses to the domains shown along the x-axis for each feature variable and filtered out any histogram bins with fewer than 50 individual trees used to compute mortality. The feature variables include **a.** tree height from 2013 in meters, **b.** distance from the nearest large granite slab

(meters), **c.** distance from the nearest river (meters), **d.** the fraction of ground returns within 20 meters that we classified as granite, **e.** trees per hectare, **f.** mean distance of the 10 nearest trees (meters), **g.** the slope of the terrain (degrees), and **h.** the aspect (categories).

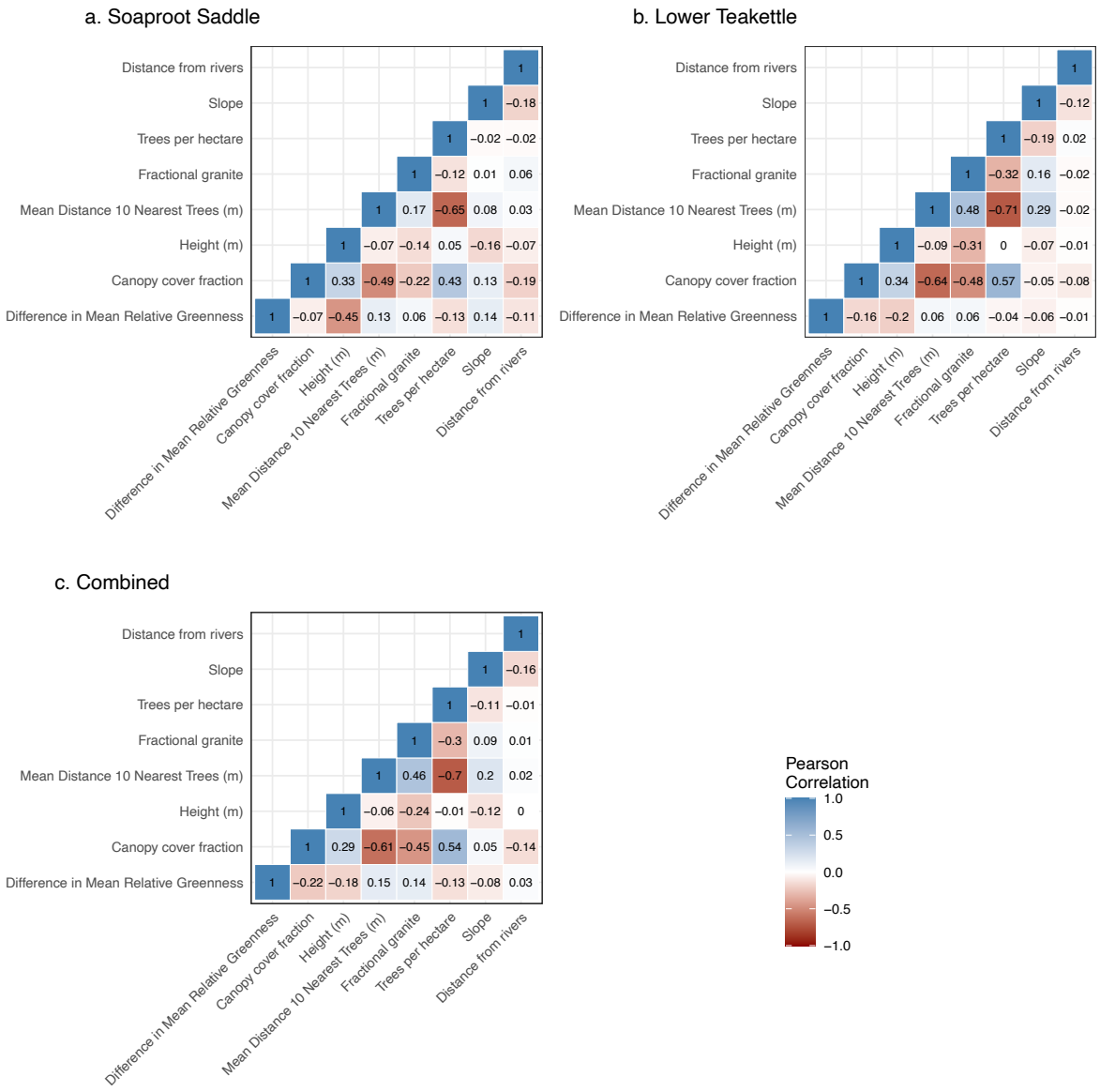

Figure A.9. Correlations among feature variables and a continuous proxy for our target variable (change in mean relative greenness between 2013 and 2017 instead of the binary labels of live or dead) for **a.** Soaproot Saddle, **b.** Lower Teakettle, and **c.** both sites combined. Each variable is computed at the individual tree level. For height, canopy cover fraction, distance from rivers, trees per hectare, and mean distance of 10 nearest trees, we

used the tree top locations and tree canopy (for tree canopy fraction) from 2013. Because of missing data in 2013, this figure is based on 686,901 trees.

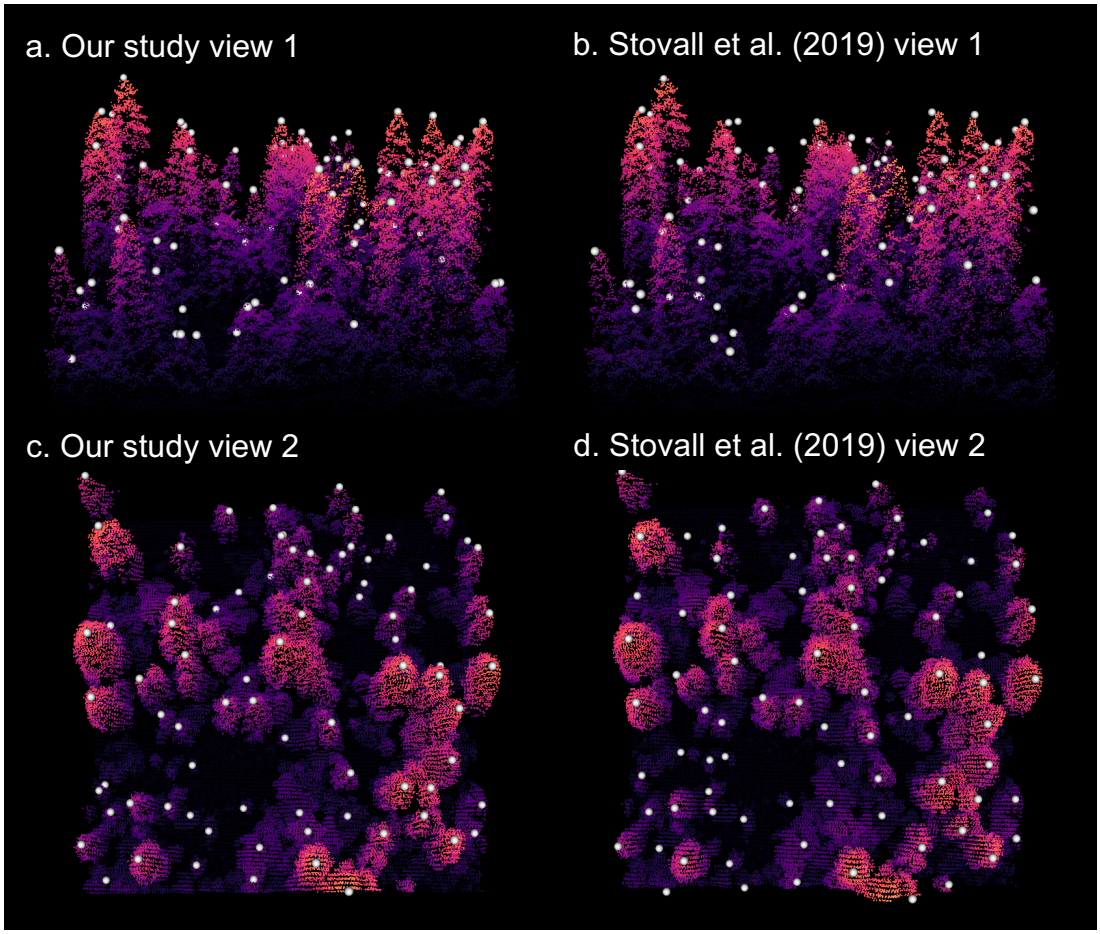

Figure A.10. The lidar point cloud from 2013 for bounding box 298050 to 298150 meters Easting and 4100800 to 4100900 meters Northing with tree top estimates. **a.** The tree top estimates (shown in white) from our method from a wolf's-eye view. **b.** The same lidar point cloud with tree top estimates from Stovall et al. (2019). **c.-d.** The same as a.-b. from a bird's-eye view.

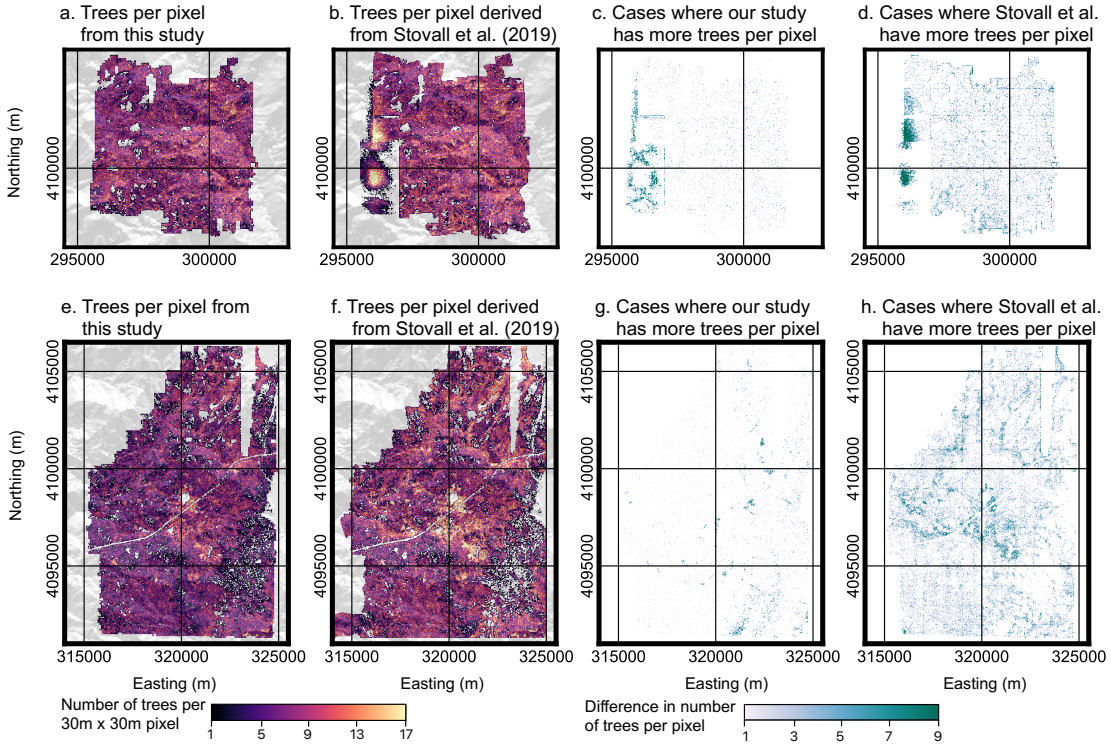

Figure A.11. The total trees per 30-meter Landsat pixel for our study and the Stovall et al. (2019) study. **a.** Trees per pixel at Soaproot Saddle for our study. **b.** Trees per pixel at Soaproot Saddle for the Stovall et al. study. **c.** Number of additional trees at our study per pixel compared to Stovall study. **d.** Number of additional trees per pixel in Stovall study compared to our study. **e.-h.** The same as a.-d. for the Lower Teakettle site.

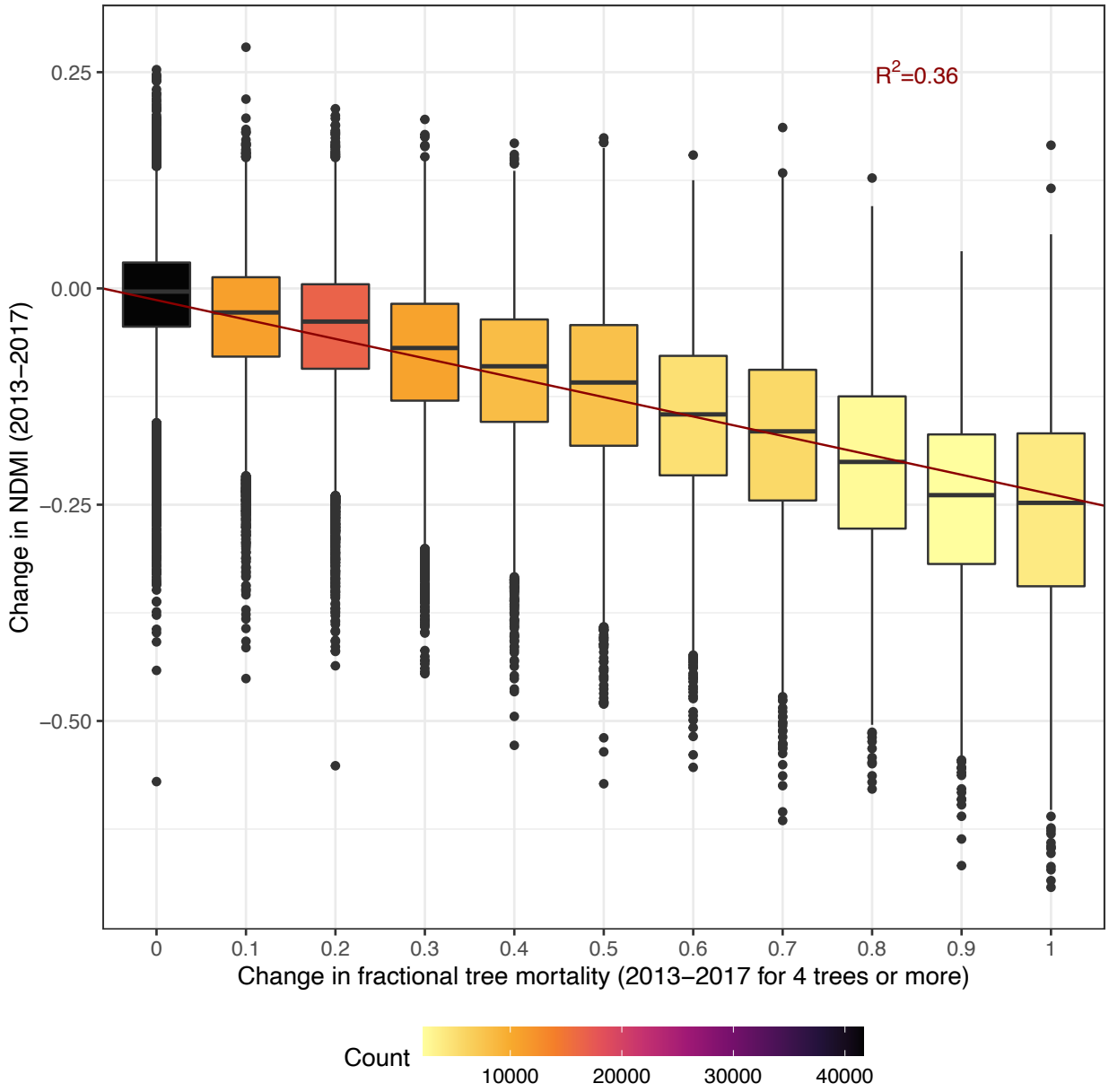

Figure A.12. Comparison of the change in fractional tree mortality to the corresponding change in the normalized difference moisture index (NDMI) between 2013 and 2017. The change in fractional tree mortality between 2013 and 2017 is shown on the x-axis for pixels with at least four trees. The y-axis shows the change in NDMI between September 2013 (mid-drought and first year of NEON data) and September 2017 (after the drought) derived from Landsat. To derive late season NDMI from Landsat imagery, we took the median pixel

values of the four Landsat images temporally nearest September 15th of 2013 and 2017. Due to missing NEON hyperspectral data from 2013 and the required trees per pixel, this analysis includes 690,072 of the trees in this study. We binned the fractional change in mortality to the nearest tenth. The total number of pixels in each bin is shown by the fill color of the box plots where light yellow represents low numbers of pixels and black represents high numbers of pixels. The dark red line shows the trend of the data before binning and is described by $f(x) = -0.22x - 0.01$ where x is the change in fractional tree mortality between 2013 and 2017 found in this study and $f(x)$ is the corresponding signal in the Landsat data with respect to the change in NDMI. The R^2 value for this relationship is 0.36.

Spectral index	Formula
Relative Greenness	$\frac{\rho_{green}}{\rho_{blue} + \rho_{green} + \rho_{red}}$
Normalized difference vegetation index (NDVI)	$\frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}}$
Normalized differenced moisture index 1 (NDMI1)	$\frac{\rho_{NIR} - \rho_{w1}}{\rho_{NIR} + \rho_{w1}}$
Normalized difference moisture index (NDMI)	$\frac{\rho_{NIR} - \rho_{SWIR1}}{\rho_{NIR} + \rho_{SWIR1}}$
Normalized difference moisture index 2 (NDMI2)	$\frac{\rho_{NIR} - \rho_{w2}}{\rho_{NIR} + \rho_{w2}}$

Table A.1. Spectral indices evaluated here for estimating tree mortality in the Sierra Nevada. To obtain each reflectance $\boldsymbol{\rho_i}$, we took the mean of the corresponding wavelength range: blue (452-512 nm), green (533-590 nm), red (636-673 nm), near infrared (851-879 nm), shortwave infrared 1 (SWIR1, 1566-1651 nm), water band 1 ($\boldsymbol{w_1}$, 1440-1460 nm), and water band 2 ($\boldsymbol{w_2}$, 1935-1955 nm).

a. Soaproot Saddle			
Feature Variable Name	Mortality Range (Maximum-Minimum)	Slope	y-intercept
Tree height (m)	0.58	0.010±0.0016***	0.19±0.064**
Trees per hectare	0.47	0.0019±0.00013***	0.24±0.017***
Mean distance 10 nearest trees (m)	0.46	-0.013±0.0017***	0.63±0.039***
Distance to rivers (m)	0.33	0.00054±0.00012***	0.38±0.027***
Slope (degrees)	0.29	-0.0067±0.0011***	0.50±0.032***
Canopy cover	0.26	0.20±0.055**	0.26±0.032***
Fractional granite	0.19	-0.12±0.064	0.36±0.028***
Aspect categories	0.06	Non-linear domain	Non-linear domain
b. Lower Teakettle			
Feature Variable Name	Mortality Range	Slope	y-intercept
Slope	0.65	0.0084±0.0013***	0.048±0.054
Tree height (m)	0.32	0.0051±0.00045***	0.093±0.019***
Trees per hectare	0.23	-0.00045±0.00011**	0.24±0.019***
Distance to rivers (m)	0.19	0.00016±0.000055*	0.20±0.015***
Canopy cover	0.09	0.089±0.016***	0.18±0.0092***
Fractional granite	0.07	0.025±0.014	0.18±0.0078***
Mean distance 10 nearest trees (m)	0.05	-0.00037±0.00030	0.21±0.0075***
Aspect (categories)	0.03	Non-linear domain	Non-linear domain
c. Combined			
Feature Variable Name	Mortality Range	Slope	y-intercept
Slope (degrees)	0.64	0.0075±0.0014***	0.097±0.057
Tree height (m)	0.30	0.0058±0.00025***	0.12±0.0099***
Trees per hectare	0.29	-0.00034±0.00020	0.28±0.033***
Distance to rivers (m)	0.17	0.000027±0.000058	0.26±0.016***
Canopy cover	0.14	0.16±0.018***	0.18±0.010***
Mean distance 10 nearest trees (m)	0.12	-0.0035±0.00041***	0.31±0.010***
Fractional granite	0.10	-0.0096±0.020	0.21±0.012***
Aspect (categories)	0.04	Non-linear domain	Non-linear domain

Table A.2. Parameters for each potential environmental driver of 2017 tree mortality for Soaproot Saddle, Lower Teakettle, and the combined dataset along with the mean standard error and significance level for each parameter. The slopes and y-intercepts for each feature variable are arranged by the range in mortality that we found in the data points in Figures 1.8 and A8. We found the range by subtracting the minimum from the maximum among the plotted values which were filtered to include only mortality values calculated from at least 50 individual trees and values that fall within the domain of interest for each feature variable (e.g. within 500 meters of a river) before plotting.

*The parameter has a p-value on the interval (0.01, 0.05].

**The p-value is on the interval (0.001, 0.01].

*** The p-value is less than or equal to 0.001.

Values without an asterisk have a p-value greater than 0.05.

APPENDIX B

Supporting Figures and Tables for Chapter 2

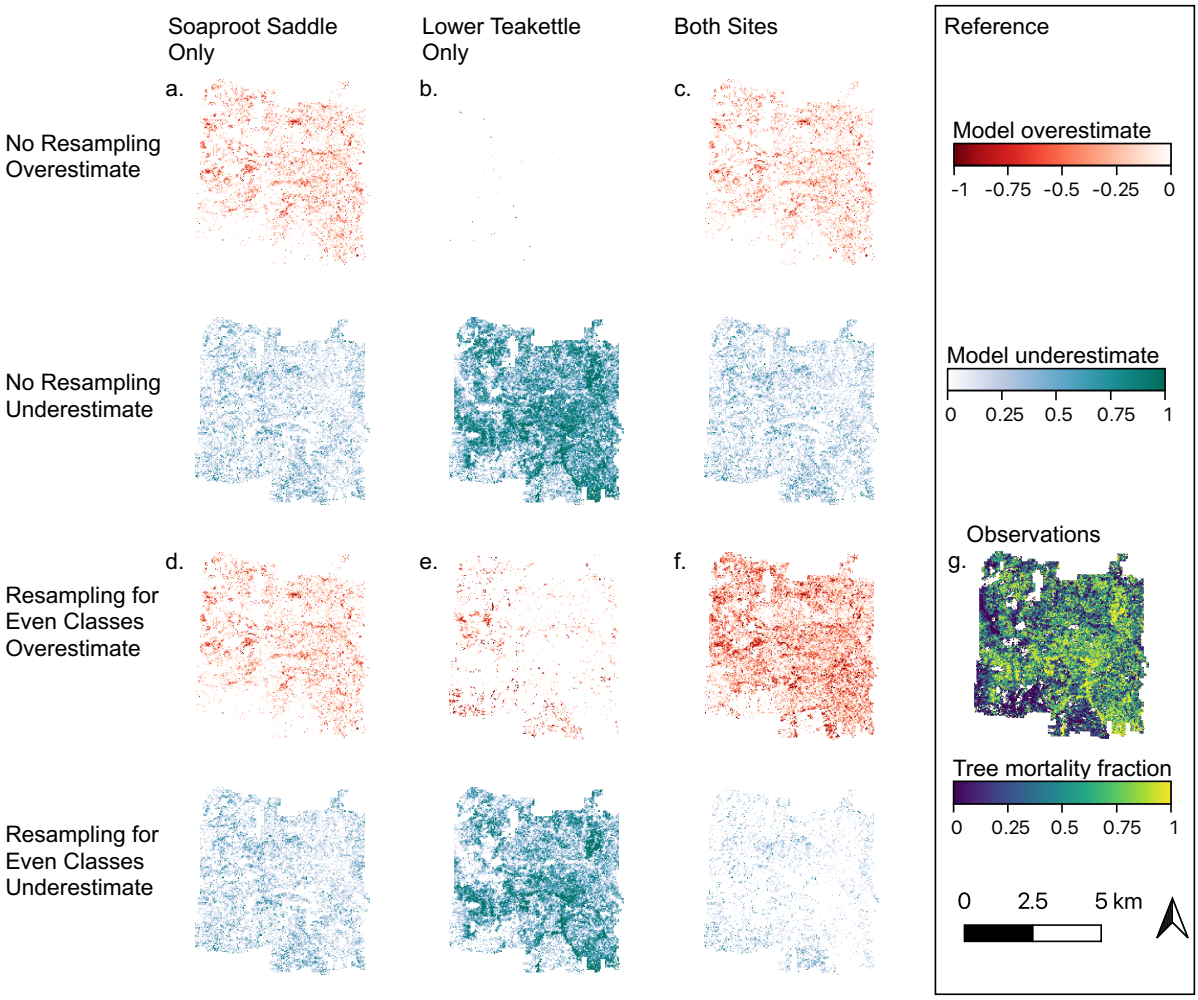

Figure B.1. Difference between maps of observed mortality fraction and modeled mortality fraction for Soaproot Saddle.

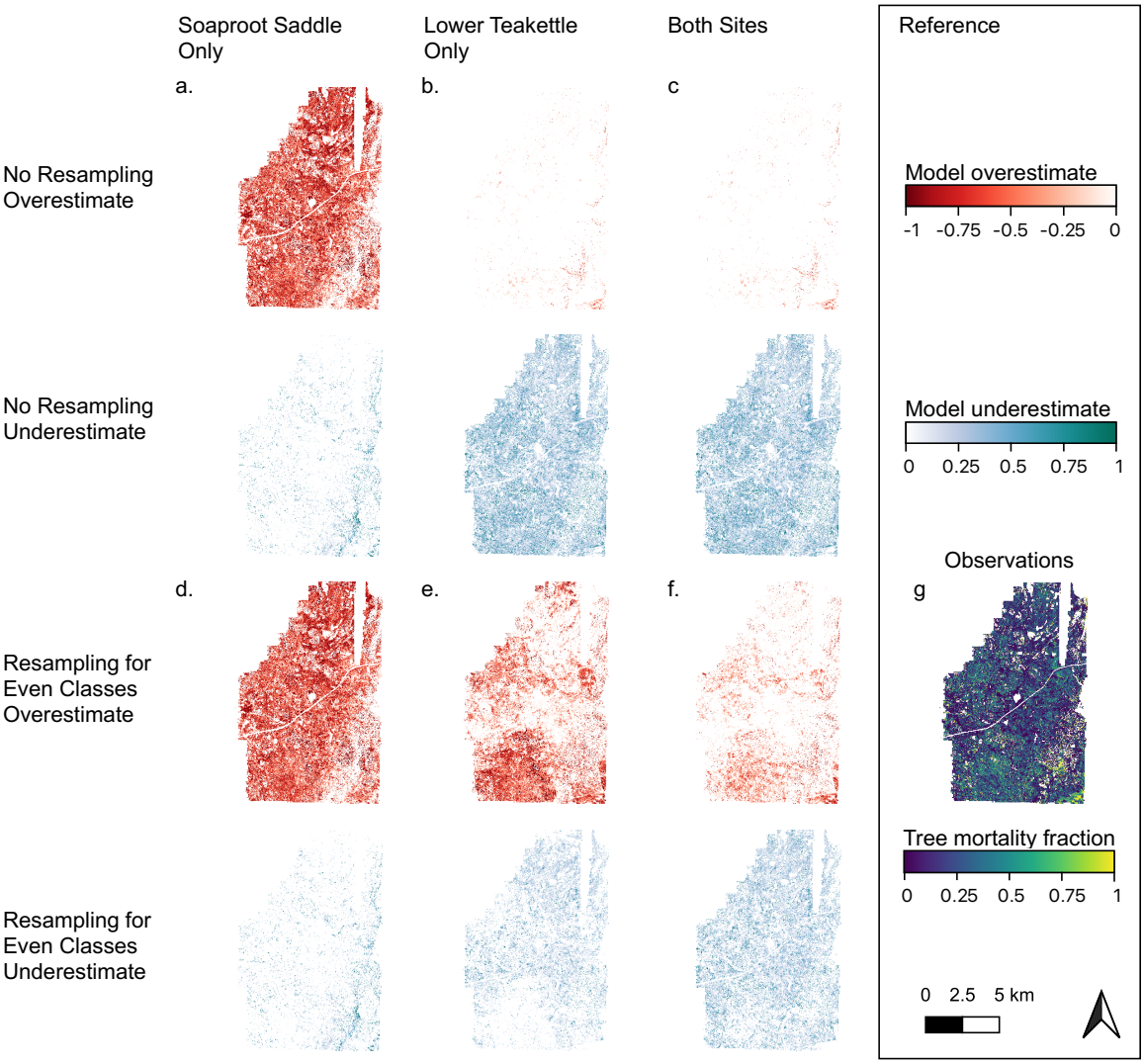

Figure B.2. Difference between maps of observations and modeled mortality fraction for best models at Lower Teakettle.

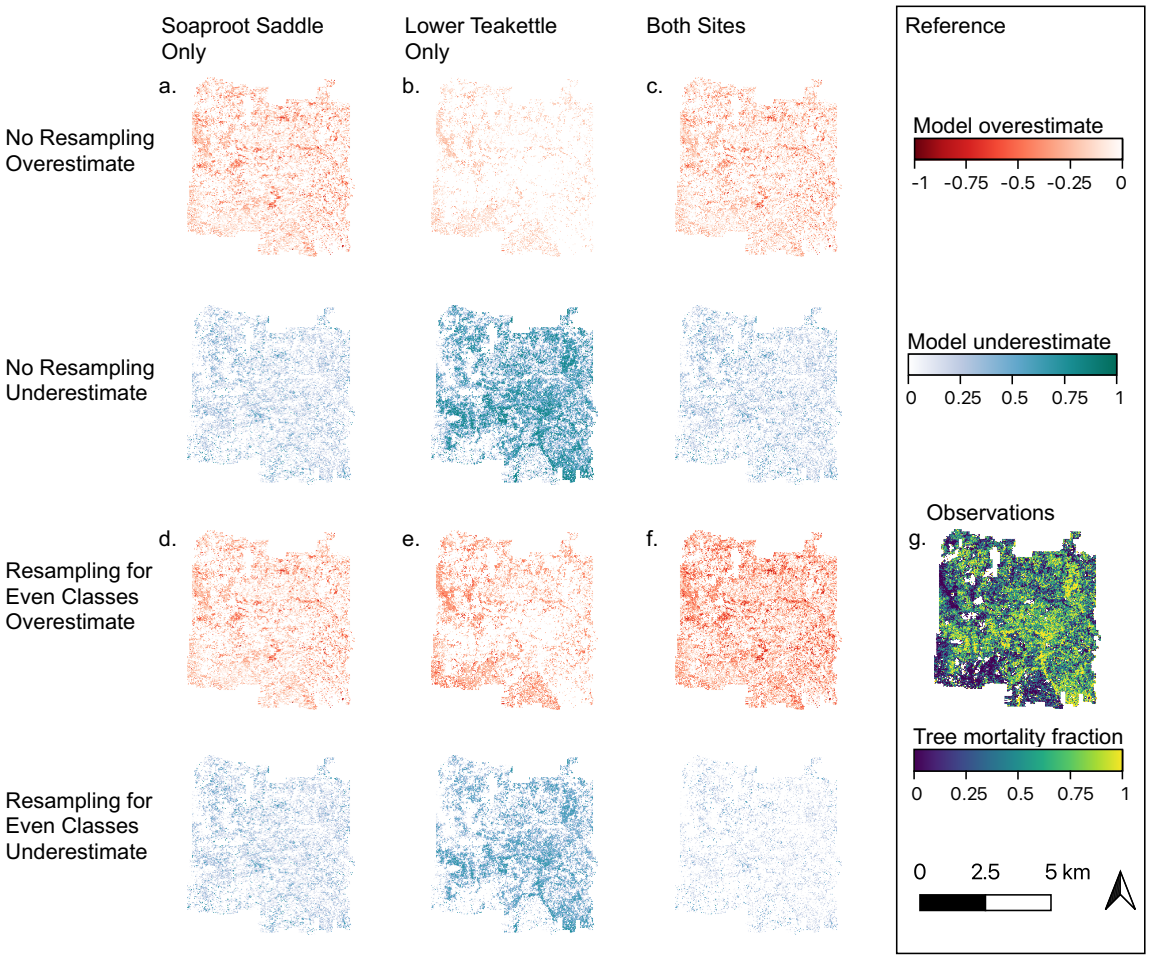

Figure B.3. Difference between observations and modeled expected mortality fraction at Soaproot Saddle.

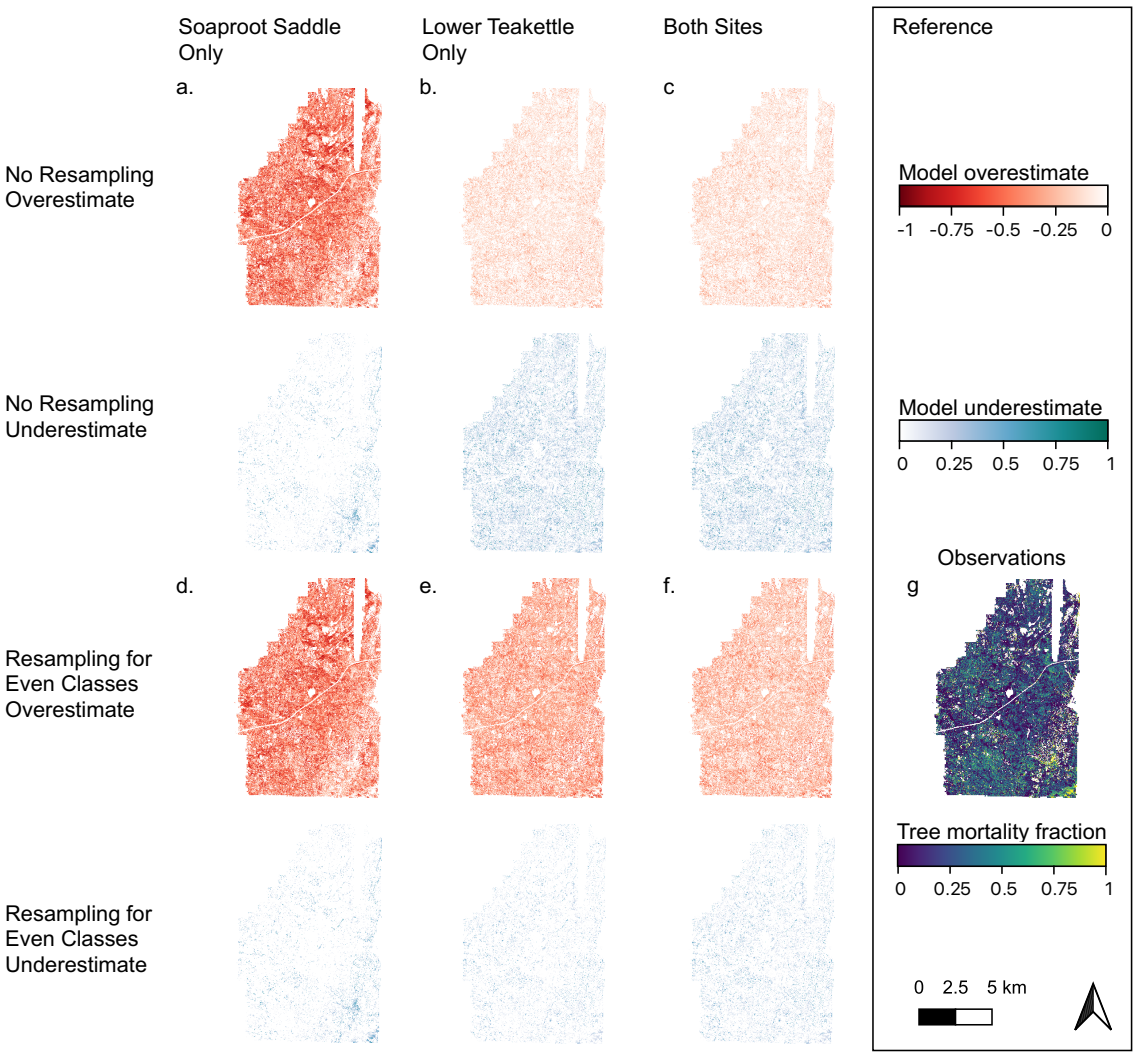

Figure B.4. Difference between observations and modeled expected tree mortality fraction at Lower Teakettle.

Feature	Soaproot Saddle Domain		Lower Teakettle Domain		Domain shown in Figures 2.2 and 2.3	
	Lower boundary	Upper boundary	Lower boundary	Upper boundary	Lower boundary	Upper boundary
Tree height in 2013 (meters)	5.4	49.2	6.3	53.6	5.9	52.8
Canopy cover fraction	0.12	0.95	0.07	0.89	0.08	0.92
Trees per hectare	33.3	166.7	22.2	166.7	22.2	166.7
NDVI before the drought (2011)	0.52	0.78	0.26	0.78	0.28	0.78
NDMI before the drought (2011)	0.04	0.47	-0.11	0.54	-0.10	0.53
Annual ET before the drought (2009-2011, mm)	441.1	822.7	226.5	776.6	239.7	799.7
Granite fraction within 20 m	0.00	0.28	0.00	0.66	0.00	0.61
Distance to rivers (m)	2.0	213.8	2.5	293.4	2.3	281.0
Slope (degrees)	5.5	34.1	3.9	35.4	4.1	34.9
North-south aspect	0.00	0.97	0.00	0.96	0.00	0.96
East-west aspect	0.00	1.00	0.00	1.00	0.00	1.00
Annual mean precipitation before the drought (2009-2011 water-years, mm)	1031	1177	1124	1365	1040	1365
Fraction of baseline annual mean precipitation during the drought (2012-2016 water-years)	0.58	0.61	0.55	0.61	0.55	0.61
Dry season (summer and fall) mean temperature before the drought (2009-	14.3	16.8	7.9	12.8	8.0	16.6

2011 water-years, °C)						
Difference in dry season (summer and fall) mean temperature during the drought (2012-2016 water-years, °C)	0.4	0.5	0.1	0.3	0.1	0.5
Winter mean monthly minimum temperature before the drought (2009-2011 water-years, °C)	-0.6	1.4	-7.1	-2.9	-7.0	1.3
Difference in winter mean monthly minimum temperature during the drought (2012-2016 water-years, °C)	1.4	1.6	1.6	2.0	1.4	2.0

Table B.1. Domain of each feature variable from Figures 2.2 and 2.3.

APPENDIX C

Supporting Figures and Tables for Chapter 3

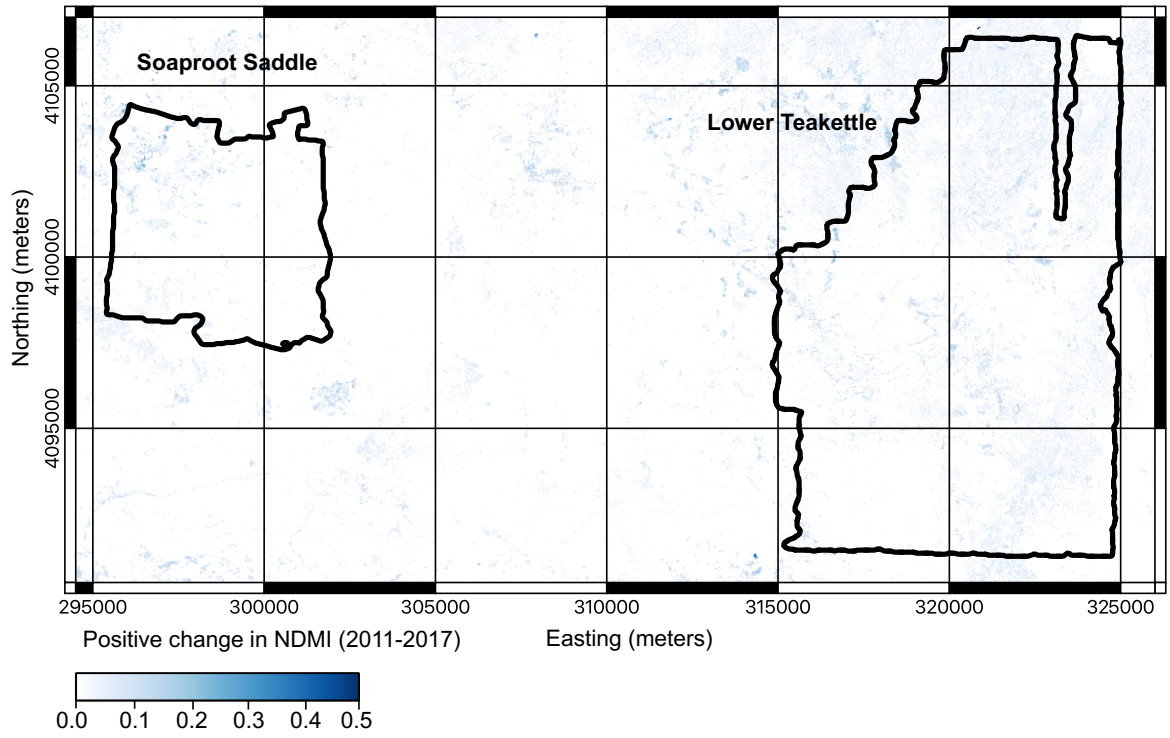

Figure C.1. The increase in NDMI between 2011 (before the drought) and 2017 (after the drought) for our study region. The outline of Soaproot Saddle is shown to the west and Lower Teakettle to the east. For visual clarity, we show the negative change (decrease in NDMI during the drought) in Figure 3.1 and the positive change here.

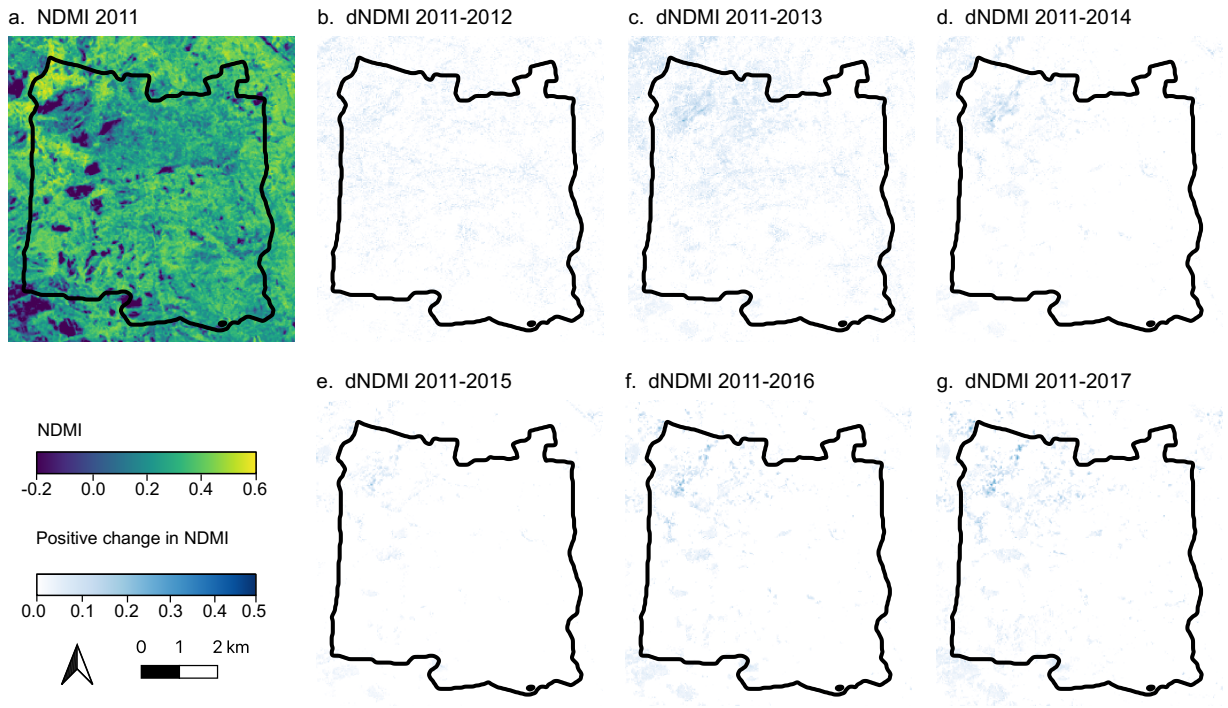

Figure C.2. Increases in NDMI during and after the 2012-2016 drought at Soaproot Saddle. a. The baseline NDMI in 2011. b. The increase in NDMI between 2011 and 2012, c. 2011 and 2013, d. 2011 and 2014, e. 2011 and 2015, f. 2011 and 2015, g. 2011 and 2016, and h. 2011 and 2017. The respective decreases in NDMI at Soaproot Saddle are shown in Figure 3.2 in the main text.

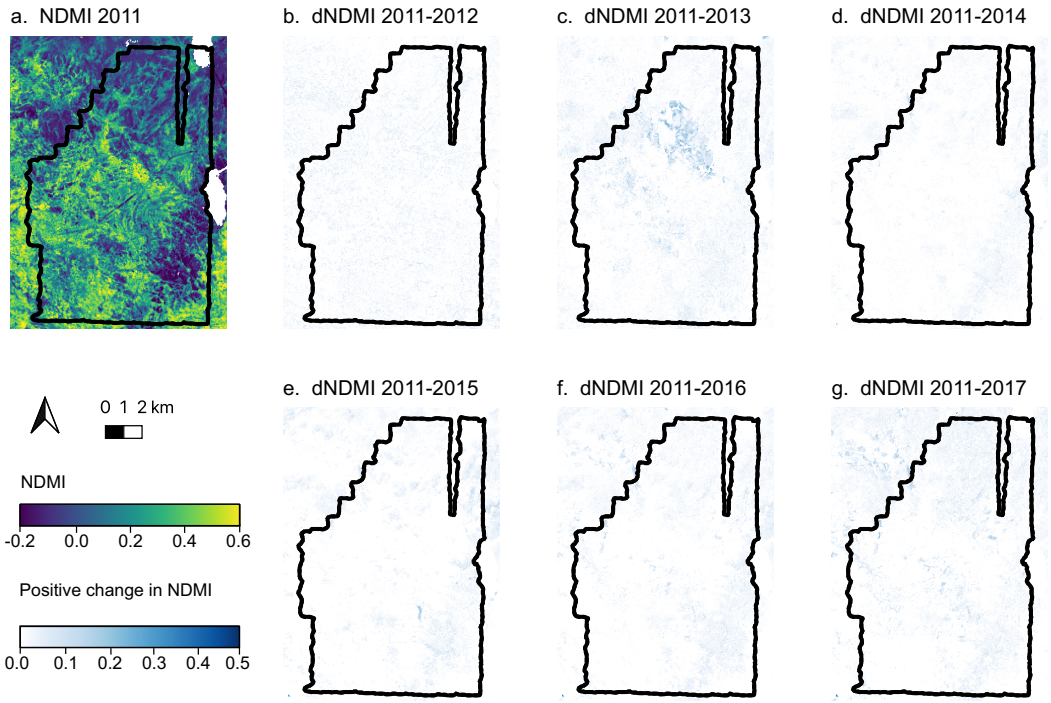

Figure C.3. Increases in NDMI during and after the 2012-2016 drought at Lower Teakettle. a. The baseline NDMI in 2011. b. The increase in NDMI between 2011 and 2012, c. 2011 and 2013, d. 2011 and 2014, e. 2011 and 2015, f. 2011 and 2015, g. 2011 and 2016, and h. 2011 and 2017. The respective decreases in NDMI at Lower Teakettle are shown in Figure 3.3 in the main text.