Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

11-9-2021

Modeling Environmental Factors Related to Drought-Induced Tree Mortality Based on Lidar and Hyperspectral Imagery

Lauren Nicole Sharwood
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Geography Commons](#), and the [Other Forestry and Forest Sciences Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Sharwood, Lauren Nicole, "Modeling Environmental Factors Related to Drought-Induced Tree Mortality Based on Lidar and Hyperspectral Imagery" (2021). *Dissertations and Theses*. Paper 5874.
<https://doi.org/10.15760/etd.7745>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Modeling Environmental Factors Related to Drought-Induced Tree Mortality

Based on Lidar and Hyperspectral Imagery

by

Lauren Nicole Sharwood

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Geography

Thesis Committee:
Geoffrey Duh, Chair
Martin Lafrenz
Christopher Grant
Melissa Lucash

Portland State University
2021

Abstract

Climate change is projected to bring more frequent and prolonged droughts, causing widespread forest die-off. Identifying tree mortality over large spatial extents in response to the most recent California drought will help forest managers and conservationists understand where there may be a greater likelihood of future die-offs. In order to find more at-risk areas, this study evaluated how interacting site-specific topographic, climate, substrate, and stand characteristics mediated tree mortality in the Central Sierra Nevada during the 2012-2016 drought. The author used lidar and hyperspectral imagery provided by the National Ecological Observatory Network to identify individual dead trees using the Random Forest classification method and created a Random Forest Regression model to assess site-specific environmental variables that had a greater influence on tree mortality. The results show that the most influential variables were tree height, density, and elevation. Results also found higher mortality rates in pines and oaks, meaning further widespread die-off of these trees could reduce forest productivity, increase fire hazard risk, and drive a shift in community composition over the long-term. This study provides a finer resolution mapping of tree mortality over the research area than was reported by the USFS Aerial Detection Survey. Due to the confounding evidence regarding the relative influence of environmental factors on tree mortality during droughts, these results provide robust information to help maintain these changing forests in a climate-informed manner. Because this study is site-specific, more research is needed to assess how environmental factors mediate drought-induced

mortality in other regions also projected to have more intense droughts as a result of climate change.

Acknowledgments

Thank you to Geoffrey Duh, Melissa Lucash, Chris Grant, and Martin Lafrenz for helping with the paper, to the Geography Department for financial support, to Weinstein et al. 2021 for the survey points, and to the National Ecological Observatory Network (NEON) for the remote sensing data.

Table of Contents

Abstract.....	i
Acknowledgements.....	iii
List of Tables.....	v
List of Figures.....	vi
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: LITERATURE REVIEW.....	5
2.1 Remote Sensing.....	5
2.2 Digital Image Classification.....	9
2.3 Tree Mortality Modeling.....	13
2.4 Influential Environmental Predictors	15
2.5 Hypothesis.....	19
CHAPTER 3: METHODS.....	20
3.1 Study Area.....	21
3.2 Remote Sensing Data.....	23
3.3 Image Classification.....	28
3.4 Tree Mortality Modeling.....	36
CHAPTER 4: RESULTS.....	39
4.1 Live / Dead Tree Identification.....	39
4.2 Tree Mortality Modeling.....	42
CHAPTER 5: DISCUSSION.....	48
5.1 Live / Dead Tree Identification.....	48
5.2 Tree Mortality Modeling.....	50
5.3 Practical Implications.....	53
5.4 Limitations.....	55
5.5 Future Work.....	56
CHAPTER 6: CONCLUSION.....	59
References.....	60

List of Tables

Table 3.1. Sensor specifications of the digital camera, imaging spectrometer, and lidar sensor carried on the NEON AOP.....	24
Table 3.2. Image interpretation key for the live/dead and genus classifiers. True-color HSI (Left), camera imagery (Center), and CHM (Right) were the remote sensing reference sources used to extract class pixels for the live/dead (purple) and genus (yellow) classifiers. Genus samples were located using Weinstein et al. (2021) in-situ survey points.....	30
Table 3.3. Environmental factors considered for analysis for association with tree mortality..	36
Table 4.1. Confusion matrix containing class predictions for the withheld validation sample pixels, OA, and κ for the live/dead classification raster.....	41
Table 4.2. Confusion matrix containing class predictions for the withheld validation sample pixels, OA, and κ for the genus classification raster.....	42

List of Figures

Figure 1.1. The U.S. Drought Monitor shows California from 2012 - 2017 and much of the Sierra Nevada under Exceptional Drought from 2014 to 2016 (https://droughtmonitor.unl.edu/Maps/MapArchive.aspx).....	1
Figure 1.2. Trees in the red and gray stages of mortality (https://www.montanabusinessquarterly.com/recovering-from-the-mountain-pine-beetle/).....	2
Figure 2.1. Spectral signature showing reflectance and absorption of vegetation across the EMS. Photosynthetic leaf pigments like chlorophyll and carotenes strongly absorb visible light. The red-edge region of the EMS is illustrated by the steep slope around 0.7 μ m, where the absorption of red light by chlorophyll transitions to high reflectance from the scattering of NIR light within the spongy mesophyll of a leaf's internal structure (Jensen 2005).....	6
Figure 2.2. Example of a decision tree landcover classification (Horning 2010).....	11
Figure 2.3. Developed based on existing literature, the conceptual framework for this study shows the common pathways for drought-induced tree mortality in the Sierra Nevada.....	15
Figure 3.1. Workflow diagram (with Key Terms list) used for this study.....	20
Figure 3.2. The study area in relation to Fresno, California (Google Earth Pro).....	21
Figure 3.3. Soaproot Saddle field site and the dominant vegetation and landcover types in the study area (Soaproot Saddle 2019).....	22
Figure 3.4. Lidar-segmented tree crowns displayed on CHM (Left) and camera imagery (Right).....	26
Figure 3.5. PC Components one through ten, containing nearly all the total spectral variance of the original 341-band HSI dataset.....	27
Figure 3.6. PC composite PCs 4, 6, and 8 highlights differences in vegetation types (Left) not apparent in the HSI true color image (Center) and higher order PC 22 highlights sensor noise (Right).....	34
Figure 3.7. Lidar-derived tree crowns and classification on true color HSI (Top). Tree crowns with live, red-stage, or gray-stage majority attribute (Bottom).....	35
Figure 4.1. Lidar-derived tree crowns with a minimum height of 2m in the study area.....	40
Figure 4.2. Correlation matrix between tree mortality and its environmental predictors.....	43
Figure 4.3. RF variable importance plot.....	44

Figure 4.4. PDPs displaying mortality probability by each individual environmental predictor: Tree height, forest cover, elevation, density, distance to streams, slope, aspect, rock cover, soil available water supply, and soil suborder.....	45
Figure 4.5. Tukey's HSD test displaying an analysis of variance of mortality probability by genus.....	47
Figure 5.1. Mean class reflectance spectra for ground-truth pixels used in RF classification training and validation. Red-stage and gray-stage trees show visible differences in the VIS and NIR parts of the spectrum. Gray-stage trees have a similar spectra to shrubs, red-stage trees have a similar spectra to soil, and live trees have a similar spectra to groundcover.....	49

CHAPTER 1: INTRODUCTION

Climate change is projected to bring more frequent and prolonged droughts, causing widespread forest die-off (Williams et al. 2013). This study evaluated how interacting site-specific environmental characteristics mediated tree mortality in the central Sierra Nevada during the 2012-2016 drought based on a tree mortality map created using lidar (light detection and ranging) and HSI (hyperspectral imagery) acquired in 2017. The results could provide an understanding of the environmental factors that mediate drought-induced tree mortality, which is projected to be more severe as a result of climate change.

Over the last decade, climate change has caused record high temperatures and severity of droughts in western North America. During the five-year period from 2012 to 2016, California experienced one of its most severe droughts in the last several centuries (Lund et al. 2018). From 2012-2015 the state recorded its lowest four-year precipitation levels since the late 19th century, while 2014 and 2015 were the hottest years on record (Chang and Bonnette 2016). The U.S. Drought Monitor from 2012 to 2017 (**Figure 1.1**) shows much of the Sierra Nevada under Exceptional Drought from 2014 to 2016.

Figure 1.1. The U.S. Drought Monitor shows California from 2012 - 2017 and much of the Sierra Nevada under Exceptional Drought from 2014 to 2016 (<https://droughtmonitor.unl.edu/Maps/MapArchive.aspx>).

The lower precipitation levels resulted in diminished snowpack at higher elevations of the Sierra Nevada, which led to diminished snow melt early in the summer.

This runoff is millions of trees' primary water supply during peak growing season (Luce et al. 2016). Warmer temperatures intensified the effects of this reduced precipitation by simultaneously increasing potential evapotranspiration (Williams 2015). An estimated 129 million trees have died in California since 2010 (US Forest Service 2017a). A combination of warm temperatures and water stress have made trees more susceptible to bark beetle (*Dendroctonus spp.*) outbreaks, which together have led to one of the worst epidemics of tree mortality in the Sierra Nevada (US Forest Service 2017a).

About one year after bark beetle attack, a tree's needles begin to lose their moisture and pigment molecules break down, turning the needles red and brown, entering the red-stage of mortality. Over the next two or three years, the needles will fall to the ground. Following the red-stage, after the tree has lost all of its foliage, it enters the gray-stage where it can remain for years. Trees in the red-stage and gray-stages of mortality are visible from aerial photography in **Figure 1.2**. Over the course of the gray-stage the fine branches fall, bark flakes off the stem and the wood of the stem continues to deteriorate. Finally, the tree falls to the ground.

Figure 1.2. Trees in the red and gray stages of mortality
(<https://www.montanabusinessquarterly.com/recovering-from-the-mountain-pine-beetle/>).

Widespread tree mortality increases carbon emissions, affects forest functioning, alters eco-hydrology and nutrient cycling, amplifies disturbances such as soil erosion (Adams et al. 2012, Allen 2007, Breshears et al. 2001, Bigler et al. 2005). California relies on its forests for wildlife habitat, carbon storage and sequestration, water supply, timber products, ecotourism, and recreation. In the wildland urban interface, standing dead trees more likely to fall into utility corridors, roads, and trails pose an increasing risk to surrounding communities. For these reasons, the US Forest Service has redirected \$32 million to monitor and respond to tree mortality in California (US Forest Service 2017b).

Historically ground-based methods have been used to monitor tree mortality in forest ecosystems. However, because these methods are typically expensive and time-consuming, they alone cannot be applied frequently over large spatial extents. Within recent years large amounts of open ecological and remote sensing data have become available. Technological advances in remote sensing, computer software, and computer hardware in the 1980s led to the integration of remote sensing in forest and conservation practices to monitor the extent, spatial distribution, percent cover, and temporal variability of forest ecosystems (Pan et al. 2013, Verrelst et al. 2015). The US Forest Service Aerial Detection Survey (ADS) currently uses Google Earth imagery taken from satellites and photographs taken from aircraft to create sketch maps of areas containing tree mortality, defoliation, and other damage over the Sierra. From this survey the US Forest Service also hosts a web mapping application, the Tree Mortality Viewer, from the ADS which displays the number of dead trees per acre (Moore 2018).

Though over a smaller area than the ADS, this research uses remote sensing data acquired in 2017 and in-situ tree survey measurements collected in 2015 by Weinstein et al. (2021) to create a framework for a finer-scale estimation of forest dieback in response to drought stress by identifying individual dead trees over the central Sierra Nevada during the most recent 2012-2016 drought. Accurately identifying individual dead trees can help stakeholders assess where to clear out dead fuel in high fire risk areas or understand where there may be a greater likelihood of future die-offs. With the increasing occurrence and severity of droughts, understanding how water stress is linked to tree mortality across forest types has become increasingly important for forest management, conservation, and resource policy (Morin et al. 2017, Wiens et al. 2009). While the understanding of the mechanisms and processes that underlie drought-induced mortality continues to grow, pre-mortality indicators of tree death remain poorly understood (Asner et al. 2015). Predictions where, and which trees are most vulnerable are needed to better understand forest response to future climate scenarios and inform mitigation strategies as California once again experienced record dry years in 2019 and 2020. The ability to identify vulnerable trees or stands of trees remains poor and is insufficient for prediction of timing, or spatial pattern, of tree die-off events (Dorman 2015).

The research objective is to use publicly available remote sensing data to map drought-induced tree mortality, then address the following research question: How did interacting site-specific environmental factors (relating to topography, substrate, and stand characteristics) mediate tree mortality in the central Sierra Nevada during the 2012 to 2016 drought?

CHAPTER 2: LITERATURE REVIEW

The integration of two emerging technologies, hyperspectral imagery (HSI), and light detection and ranging (lidar), has aided in vegetation monitoring by providing a more robust set of metrics on vegetation health and taxonomy. Most methods to assess health status and taxonomic groups still require highly precise, accurate, and representative ground-based field samples for classification model training. The machine learning Random Forest (RF) classification algorithm has recently gained popularity in image classification, due to its limited user-defined parameters, efficiency, and its ability to incorporate complex, hierarchical, and non-linear relationships. Numerous studies have modeled tree mortality using the RF regression built-in variable importance measure to assess the relative importance of site-specific environmental variables to predict tree mortality. In similar studies assessing drought-induced tree mortality in the Sierra, literature shows that the most influential environmental variables are linked to site water balance and tree height, yet acknowledge that bark beetle infestation is a common direct cause of tree mortality in the region.

2.1 Remote Sensing:

Vegetation monitoring has historically been performed using ground-based methods. Remote sensing, the recording of information of objects without being in direct contact with them, was introduced to help provide repeat vegetation monitoring over larger spatial extents. Passive sensors record energy reflected or emitted from surface objects along the electromagnetic spectrum (EMS), whereas active sensors, such as lidar,

emit their own energy and record their returns. Highly accurate three-dimensional surface models can be created from precise lidar sensor location and orientation measurements, timestamps that measure the time it takes each pulse to be returned to the sensor after being emitted, and the speed of light.

Reflectance measurements recorded by hundreds of narrow bands along the EMS by a hyperspectral sensor, or imaging spectrometer, are more sensitive to small variations in reflectance that a tree's structure, chemical, and biophysical properties determine (Hamlin et al. 2011). These distinctive patterns are also known as spectral signatures.

Figure 2.1 illustrates the typical spectral reflectance pattern of a live plant pixel along the EMS. Reflectance information in these regions can be used to measure levels of plant pigments which assess photosynthetic activity and assess stress (Clevers et al. 2010, Ustin et al. 2004).

Figure 2.1. Spectral signature showing reflectance of vegetation across the EMS. Photosynthetic leaf pigments like chlorophyll and carotenes strongly absorb visible light. The red-edge region of the EMS is illustrated by the steep slope around 0.7μm, where the absorption of red light by chlorophyll transitions to high reflectance from the scattering of NIR light within the spongy mesophyll of a leaf's internal structure (Jensen 2005).

The gray regions in **Figure 2.1** represent atmospheric absorption, where certain atmospheric gases such as water and carbon dioxide absorb light of certain wavelengths, diminishing the return recorded by a sensor. Multispectral satellite-based sensors like Landsat MSS, TM, ETM+, and OLI are designed to not record reflectance in regions of atmospheric absorption. These multispectral sensors also offer the benefit of repeat monitoring, but record reflectance with wider, broader spectral bands, and therefore do not capture some of these narrow spectral features (Ahmed et al. 2016, Malila 1980).

Despite the potential of hundreds of HSI bands, the amount of information in the resulting dataset poses a variety of issues along with increased processing time and storage limitations. Because wavelengths are sampled at such narrow intervals, each band contains redundant and highly correlated information between adjacent bands. For classification tasks, these highly correlated bands cause a problem known as the Hughes' Phenomenon, which states that the minimum statistical confidence of separability of classes can only be maintained if each class has enough training samples to train the classifier while still leaving enough samples for validation (Hughes 1968). This problem is often addressed through data reduction and feature extraction techniques, such as Principal Component Analysis (PCA). PCA is a linear transformation that reprojects the data to a new coordinate system where the first principal component (PC) is the axis that captures the greatest amount of variance of the original dataset, the second PC is the axis that captures the second greatest variance, and so on until the higher-level PCs represent only noise from the original dataset (Oo 2018). Past HSI classification studies have

shown reliable results using PCA (Rodarmel and Shan 2002, Yang and Everitt 2010, Makantasis et al. 2015).

While HSI can provide information about the water-related biochemical and biophysical properties of forests, lidar can provide more robust 3D information about the biophysical properties, such as tree height, fractional vegetation cover, canopy geometry, and above-ground biomass (Dees et al. 2012, Allgöwer et al. 2008). This reliable distinction of vegetation classes of different heights has been used with HSI to map urban forest structure (Alonzo et al. 2016), map biomass (Vaglio et al. 2014) assess stress in the Sierra Nevada (Asner et al. 2015), and assess fuel type to improve wildfire behavior models (Swatantran et al. 2011, Arroyo et al. 2008, Goetz et al. 2006). In a study mapping European bark beetle mortality using only HSI, Fassnacht et al. (2014) found highest misclassification rates between dead tree pixels and sparsely vegetated soil pixels, emphasizing the importance of height information to distinguish between spectrally similar classes.

Numerous studies have identified and delineated tree crowns using lidar-derived products. Stovall et al. (2019) delineated individual tree crowns using the local maxima approach, relying on a moving window to determine local high points in the canopy height model (CHM). The identified high points serve as potential crown locations and crowns are segmented using small-scale watershed delineation. Scholl et al. (2020) applied different buffer sizes around NEON woody vegetation in-situ field point data and achieved highest tree species classification accuracies with polygons half the maximum crown diameter of each tree.

2.2 Digital Image Classification:

Image classification is the process of sorting pixels in an image into classes based on their input pixel values, typically spectral or PC bands. Image classification relies on the assumption that the inputs for each class will be consistently more similar than inputs belonging to different classes. Sensor noise, differences in illumination geometry, mixed pixels containing multiple classes, varying health status of vegetation, and phenology are just some factors that affect vegetation reflectance within a forest (Thenkabail 2017, Clark and Roberts 2012). The reflectance variations caused by these factors can lead to a phenomenon which complicates image classification, where objects from the same class have different spectra, and objects from different classes have the same spectra. “Fuzzy” classification logic considers the heterogeneous and imprecise nature of the real world, where each pixel or object has a membership grade (from 0 to 1) for a class. For “hard” categorical classes, each pixel or object represents a sampled area on the ground and receives a class attribute. Image classifications generally use “hard” logic to create classification maps of categorical classes.

There are three general types of “hard” image classification: the emerging object-oriented approach, or the two pixel-based (unsupervised and supervised) approaches. Traditional pixel-based approaches decide what class each pixel belongs to on an individual basis. Object-based classifications segment the image by grouping neighboring pixels together based on how similar they are. Unlike pixel-based classifications, this allows for classification based on texture, context, and geometry. Image objects are groups of pixels that are similar to one another based on a measure of spectral properties

(i.e., color), size, shape, and texture, as well as context from a neighborhood surrounding the pixels.

Unsupervised pixel-based image classifications group pixels with similar spectral characteristics into unique clusters without any ground-truth information. Unsupervised algorithms are relatively fast and easy to perform, human error is minimized, and no prior knowledge of the image area is required. However, the predicted classes do not necessarily represent features on the ground and the classes can be time consuming to interpret (Peerbhay et al. 2015). Supervised pixel-based algorithms require users to provide samples of dominant classes in the image for model training and validation.

The results of any supervised classification task rely on the accurate and precise collection of ground-truth samples to be used for model training. Samples for model training are generally collected using high-resolution imagery or in-situ and should be representative of the dominant classes within the scene. This ground-truth information is used to train the model to predict classes for unseen pixels based on how similar its values are to each class's training pixels. A model training on inputs from a misidentified ground-truth sample could result in model predictions having low accuracy, or high confusion, between the two classes. Creating training data can be an expensive and time-consuming task that often requires external data sets and user knowledge of the task and study area (i.e., tree taxonomy, local ecology).

The Random Forest (RF) algorithm is a supervised decision tree classifier that is increasingly used in remote sensing studies (Wang et al. 2015). Decision trees are predictive models that use a set of binary rules to group pixels into different classes.

Figure 2.2 shows a simple classification tree (Left) and resulting land cover map (Right) from Horning 2010, where the tree is a set of binary decisions and terminal nodes connected by branches. At the root node, Node 1, a pixel is assigned to the Water class if Band 4 is less than or equal to 46. If Band 4 is greater than 46, another binary decision node is evaluated, and the tree continues to grow until all branches end with a terminal node. RF is an ensemble classifier such that it uses results from many (usually several hundred) different decision tree models to calculate a response. For instance, if 200 trees are grown and 150 of them predict that a given pixel is a Forest and 50 predict it is Scrub, then the predicted output for that pixel will be a Forest.

Figure 2.2. Example of a decision tree landcover classification (Horning 2010).

A study comparing various machine learning classifiers showed that Support Vector Machine (SVM) can be more accurate than neural networks and RF, especially when analyzing high-dimensional data with limited training samples typical of HSI analysis (Mountrakis et al. 2011, Melgani and Bruzzone 2004). However, RF has been widely used in image classification studies due to its advantageous processing speed and minimal user-defined parameters (Belgiu and Dra 2016, Corcoran et al. 2013, Lawrence

et al. 2006, Thenkabail 2017). Because RFs are largely insensitive to overfitting and errors stabilize before the default number of trees, $nTree$, is used, the default value of 500 is generally used. The second parameter, $mTry$, is typically set to the square root of the number of input variables (Wang et al. 2015).

Accuracy for classification tasks are typically assessed using a confusion matrix, which compares class values of withheld ground-truth reference validation pixels to the class values in the model predictions, a classified raster. Each row represents pixels of a given class from the ground-truth samples that were withheld for validation, and each column shows the class the model predicted from these pixels. Producer's accuracy (PA) quantifies how often a pixel in the classified map correctly represents that class on the ground, whereas user's accuracy (UA) quantifies how often a given class on the ground is reflected by that class on the map. Overall accuracy (OA) is calculated by dividing the total number of correctly classified validation pixels by the total number of validation pixels. Cohen's Kappa statistic (κ) is another standard measure of image classification accuracy. κ considers the possibility of agreement occurring by random chance by removing this agreement occurring by chance. κ equal to 0 represents that the agreement between the predicted and reference pixels is a result of random chance and a κ equal to 1 represents complete agreement between predicted and reference pixels. A κ value is smaller than OA because it does not account for chance agreement. As such, a κ value larger than 0.2 is considered fair (McHugh 2012).

2.3 Tree Mortality Modeling:

RF regression models are widely used in ecological studies, as they are able to incorporate complex, hierarchical, and non-linear relationships between predictor and response variables (Evans et al. 2011, Baguskas et al. 2014, Madakumbura et al. 2020, Paz-Kagan et al. 2017). Like RF classification, the number of decision trees grown is determined by *nTree* and the number of variables chosen at each split is determined by *mTry*. The RF regression model is also an ensemble of decision trees where each tree makes its own individual prediction and the predictions are averaged to produce a single result.

Because a forest consists of many deep trees, where each tree is trained on data using a random selection of features, gaining a complete picture of the decision process by examining each individual tree is not feasible. One way of gaining insight into a RF is to compute its variable importance (Zhang et al. 2018). This built-in variable importance measure assesses which predictor variables have the greatest influence over the response variable (Prasad et al. 2006). When a tree is built, the decision about which variable to consider at each node uses a calculation of Gini impurity. The Gini-based variable importance is calculated based on the reduction in sum of squared errors whenever a variable is chosen to split. For variables of different types, this method is based in favor of continuous variables and variables with more classes. Using the *permutation* feature importance measurement introduced by Breiman 2001, a feature's importance is measured by calculating the increase in the model's prediction error after randomizing, or permuting, that predictor. A predictor is “important” if shuffling its values increases the

model error, because the model relied on it when predicting. A predictor is "unimportant" if shuffling its values leaves the model error unchanged, because the model ignored it when predicting. Both methods may still overstate the importance of correlated predictors.

Partial dependence plots (PDPs) help to visualize the average partial relationship between the response and one or more predictors of a machine learning model. PDPs show the marginal effect, or average partial relationship one predictor variable has on the response variable. Plots are created by forcing all instances to have the same predictor value. Predictions are made for each instance, then averaged to yield an average prediction for a given predictor value. Because all predictor variables except for the one being examined are held constant, PDPs do not highlight the nonlinear interactions between predictor variables (Friedman 2001).

Paz Kagan et al. (2017) created a RF model where tree mortality acted as the response to site-specific environmental predictors in the southern Sierra Nevada. Authors applied this RF model such that each instance, a live or dead tree, was given a mortality probability (0 to 1). Authors combined these mortality probabilities with species information from a tree species classification map then ran a one-way analysis of variance (ANOVA) to find there was a significant difference between species composition and the probability of mortality. The post-hoc Tukey tests for group homogeneity were applied to identify the statistical differences for each of the species classes in relation to the elevation gradient and probability of mortality.

2.4 Influential Environmental Predictors:

Multiple mechanisms may influence tree mortality during a prolonged hot drought. Whether a tree lives or dies in such conditions is related to morphological defenses of an individual tree, factors related to its environment, dynamics of pests and pathogens, and the interactions between them (McDowell et al. 2008, Pureswaran et al. 2018). I created a conceptual framework (**Figure 2.3**) based on existing literature about drought-induced tree mortality that illustrates these interactions and highlights the mediating influence local environmental variables.

Figure 2.3. Developed based on existing literature, the conceptual framework for this study shows the common pathways for drought-induced tree mortality in the Sierra Nevada.

During droughts, the lack of water often results in foliage dieback as an individual defense mechanism to help conserve resources. However, this defense is not always enough to save a tree, and some succumb directly from this drought stress. Healthy trees are often able to fight off any bark beetles' entry points with a sticky resin, but under

prolonged drought conditions with little to no water and high evapotranspiration demands, trees must devote their energy to surviving, do not have the resources to produce this resin, and may succumb directly from bark beetle attack (Raffa et al. 2008).

Bark beetles become attracted to areas with mass drought-induced forest die off, causing a positive feedback loop to occur. When beetle populations become high enough, even healthy trees cannot produce enough resin to ward off a mass attack (Jenkins et al. 2008). This cause of mortality is still indirectly drought-induced, as these high temperatures and lack of water over extended periods have led to stress and initial patches of forest dieback, which attracted the large number of beetles. Certain environmental factors can either help or hinder a tree's access to resources. The mediating influence of various environmental factors is largely site-specific, and there are still conflicting findings about the relative influence of these factors within the Sierra Nevada.

Stephenson 1988 pointed out soil depth can be a good indicator of water-holding capacity and thus water availability. Generally, deeper soils may provide greater soil water availability, which leads to less water stress (Koepke et al. 2010) and lower mortality (Meddens et al. 2015, Paz-Kagan et al. 2017). However, Flake and Weisberg 2019 found higher levels of tree mortality in areas with deeper soils. Dorman et al. (2015) suggests this association may be due to trees' access to short-term water storage in surface soils as well as long-term storage in rock fissures, or less permeable soil layers, which retain water for longer during the growing season. Trees in deeper soils, which have more available water in non-drought years, may have greater leaf area, a liability for

transpiration and water loss in hot droughts. Further, they may also allocate fewer resources to developing drought-resistant characteristics such as reduced leaf area and cavitation resistance (Ogle et al. 2000, McDowell et al. 2011).

Fire suppression policies in the United States have created overcrowded forests. Forests and woodlands where excess trees are growing have fewer soil and moisture resources available per individual tree, leaving trees less healthy and less able to withstand drought, insects, and disease (Ingram and Kocher 2015). Stand density reduction treatments are often used to mitigate fire risk and restore fire health. Density reduction treatments have been effective at reducing fire intensity during following wildfires (Graham et al. 1999, Safford et al. 2009), yet its utility in improving forest health is less well understood Fettig et al. (2007). Fettig et al. (2019) found that tree mortality was positively correlated with tree density in the central and southern Sierra Nevada.

Baguskas et al. (2014) found a greater probability of mortality for shorter trees in coastal California pine forests, possibly due to a smaller root-shoot ratio than larger trees, and thus limited access to water reserves deeper in the soil. On the other hand, Paz-Kagan et al. (2017) found higher rates of mortality with increasing tree height in the Sierra Nevada. Fettig et al. 2019 supposed this to be driven by bark beetles primarily attacking larger trees, while Paz-Kagan et al. (2017) suggested the increase in water use efficiency of taller trees increased evaporative demands in the upper canopy.

Paz-Kagan et al. (2017) found a higher probability of mortality in areas further from water, at lower elevations, on southwest- and west-facing, shallow slopes in the

Sierra Nevada. Conversely, in a study quantifying the effects of forest density, climate, and environmental factors influencing forest mortality in the Lake Tahoe Basin, Gunst et al. (2016) found an increased mortality risk on north-facing aspects and steeper slopes. Fettig et al. (2019) found that tree mortality was positively correlated with a steeper slope in the central and southern Sierra Nevada. These nuanced yet confounding results within the Sierra illustrate that causes of forest mortality from environmental factors and stand characteristics are complex, site-specific, and difficult to predict.

Asner et al. (2015) found different mortality patterns within different conifer communities. Findings ranged from almost no relationship in the closed-cone pine-cypress communities to severe rates of mortality with increasing progressive water stress in Ponderosa pines (*Pinus ponderosa*), which have been prone to bark beetle attack in the region. Paz-Kagan et al. (2017) found the probability of mortality to be significantly higher, up to 20%, in areas occupied by *Pinus ponderosa* and *Pinus lambertiana*. According to the authors, over time such widespread species-specific mortality may lead to changes in species composition and associated biodiversity and ecosystem function.

2.5 Hypothesis:

Based on the aforementioned research, I expected to find topographic, substrate, and stand characteristics affecting a tree's water supply to be the most influential environmental factors related to drought-induced tree mortality in the study area. Elevation determines climatic conditions like summer temperatures and winter rainfall. Thus, I expect to find higher rates of mortality at higher elevations that typically receive greater rainfall. I also expect to find higher rates of mortality on steep slopes with higher runoff, south-west facing slopes with higher evapotranspiration rates, and trees further from streams. I expect to find trees surrounded by higher percent rock cover (a proxy for soil depth which can be used to estimate volumetric soil water content) to have shallow soils, lower volumetric water content, and lower rates of mortality. For these reasons, I expect to find trees in soils with a greater available water supply to have higher mortality probabilities. I also expect to find lower rates of mortality probabilities among trees in the xeroll soil suborder, which typically have lower water content compared to the umbrepts and orchepts suborders. Each order is divided into suborders based on properties that influence soil formation and are important to plant growth (USDA Web Soil Survey 2019). Trees experiencing more competition and thus a higher mortality probability might be surrounded by a high density of trees, and in areas with high percent forest cover. I also expect taller trees to have greater water needs and to be associated with higher rates of mortality probability. Bark beetles have been known to attack *Pinus ponderosa*, so I expected to find higher mortality probabilities among pines in the study area.

CHAPTER 3: METHODS

The workflow diagram (**Figure 3.1**) shows a brief description of the steps I took to perform this study. With the lidar-derived CHM and PC-reduced HSI mosaic, I created a RF live/dead image classification then a RF genus image classification on the live trees. I compiled and attributed environmental factors related to topography (elevation, slope, aspect, distance to streams), stand characteristics (tree height, percent forest cover per plot, and tree density in the 100,000 sq m hexagon area surrounding each tree), and substrate (percent rock cover per plot, soil available water supply up to 50 cm, and soil suborder) from USDA Web Soil Survey 2019 for identified tree crowns. These environmental variables acted as the explanatory variables in the RF regression model predicting the probability of individual tree mortality. I used the RF variable importance function to assess the most influential environmental predictors. I applied the tree mortality model to predict the mortality probability for each tree in the study area and applied a one-way ANOVA and Tukey's honestly significant difference test (Tukey's HSD) to assess the differences in mortality between oaks, pines, and incense-cedars.

Figure 3.1. Workflow diagram (with Key Terms list) used for this study.

3.1 Study Area:

The study area is in Sierra National Forest, which is located on the western slope of the central Sierra Nevada mountain range in California. **Figure 3.2** shows the study area boundary in relation to Fresno, California.

Figure 3.2. The study area in relation to Fresno, California (Google Earth Pro).

The Sierra Nevada has a Mediterranean climate with cold, wet winters and hot, dry summers. Its elevation ranges from around 150m to over 4380m and receives around 750mm to 2000mm of precipitation per year, though the actual amount of precipitation is largely determined by elevation (Ingram and Kocher 2015). This large elevation gradient, which largely determines moisture conditions, also influences the distribution of ecological communities.

The area for this study expands slightly beyond the NEON Soaproot Saddle (SOAP) field site, which sits at 975m to 1400m in elevation and has historically received

around 127cm of rain per year. Temperatures generally range from 15 °C to 27 °C but can reach up to 32 °C during the day in the summer months. South-facing slopes generally receive more insolation and consequently experience greater evapotranspiration rates. The National Land Cover Database characterizes SOAP as a mix of evergreen forest and shrub scrub vegetation types. The SOAP field site covers six sq km of mixed-conifer forests that recently experienced high levels of mortality in part directly due to pine beetles (Soaproot Saddle 2019), which can be seen in **Figure 3.3**.

Figure 3.3 Soaproot Saddle field site and the dominant vegetation and landcover types in the study area (Soaproot Saddle 2019).

Ponderosa pine (*Pinus ponderosa*) and California incense-cedar (*Calocedrus decurrens*) dominate the overstory with co-dominant Canyon live oak (*Quercus chrysolepis*) and California black oak (*Quercus kelloggii*), and an average canopy height around 32m (Soaproot Saddle 2019). Mariposa manzanita (*Arctostaphylos viscida* ssp. *Mariposa*) forms dense chaparral thickets in open areas and deerbrush (*Ceanothus*

integerrimus) and poison oak (*Toxicodendron diversilobum*) are common in the shaded understory. Mountain misery (*Chamaebatia foliolosa*) provides dense herbaceous groundcover throughout much of the site (Soaproot Saddle 2019).

Exposed bedrock associated with shallower soils for nearby vegetation is interspersed throughout the site and intermittent streams of the watershed drain into Big Creek (Soaproot Saddle 2019). Soils are moderately deep and well-drained on mountain slopes, while shallower and excessively drained soils with poorer water retention exist on steeper slopes. Soil forming parent material includes phases of granite primarily from the Holland and Chaix series, alluvium, and glacial deposits (Ingram and Kocher 2015). Taxonomic class is defined as fine-loamy, mixed, semiactive, mesic Ultic Haploxeralfs (Soaproot Saddle 2019).

3.2 Remote Sensing Data:

The National Ecological Observatory Network (NEON) is a National Science Foundation (NSF) funded project with the purpose of discovering, understanding, and forecasting the impacts of climate change, land use change, and invasive species on continental-scale ecology. NEON provides free, publicly available ecological data available for download at <https://data.neonscience.org/data-products/explore>. NEON's Aerial Observation Platform (AOP) is flown on a fixed-wing aircraft and measures ecological change over 73 field sites over 20 distinct eco-climatic regions (<https://www.neonscience.org/data-collection/airborne-remote-sensing>). The AOP co-acquires high-resolution camera images with a Phase One digital camera, HSI with

Airborne Visible Infrared / Imaging Spectrometer (AVIRIS-NG), and lidar using Optech ALTM Gemini (**Table 3.1**).

Remote sensing data used in this study was acquired on the NEON AOP at a height above ground level of 1000m in July, 2017. During July vegetation is at least 90% of maximum greenness in the study area. A high solar zenith angle and minimal cloud coverage during this time also provides optimal acquisition conditions for HSI. AVIRIS-NG records surface reflectance along the VSWIR portion of the EMS (400nm – 2500nm), at 5nm bandwidths, resulting in 426 narrow spectral bands. Discrete return lidar point cloud data is collected from Optech ALTM Gemini, which emits pulses of light at a wavelength of 1064 nm to describe vegetation cover, height, and shape.

Table 3.1. Sensor specifications of the digital camera, imaging spectrometer, and lidar sensor carried on the NEON AOP.

	Digital Camera	Imaging Spectrometer	Lidar
Sensor	Phase One iXU-RS1000 (100 MP)	AVIRIS-NG	Optech ALTM Gemini
Data product	RGB camera imagery	Orthorectified surface directional reflectance	Discrete return lidar point cloud
Spatial Resolution	0.1m	1m	1-4 points/sq. m

I used classified lidar point cloud data to create a digital elevation model (DEM) and canopy height model (CHM). I used the lidR package in R to create a 1m digital surface model (DSM) using the grid_canopy function and first returned laser pulses from the lidar point cloud. This method consists of the interpolation of first returns with a triangulated irregular network (TIN), then rasterizes the TIN into a grid to create the DSM. Within the grid_canopy function I used the pitfree algorithm, which removes pits

in the canopy from laser pulses that penetrate deeply into the canopy before producing the first. The algorithm, developed by Khosravipour et al. (2014), replaces each point with 8 points around the original one. This densifies the point cloud to create a smoother DSM with less pits. I created the 1m DEM using the `grid_terrain` function, which interpolates classified ground points using the k-nearest neighbors approach with an inverse-distance weighting. I used the default parameter value of 10 for k (number of nearest neighbors), 2 for p (power of weighting), and 50 for rmax (maximum search radius). Next, I subtracted the DSM from the DEM to create the 1m CHM.

I used the Forest Tools package in R to perform tree crown segmentation using the lidar-derived CHM. I performed automatic treetop detection using a variable window filter (vwf) algorithm developed by Popescu and Wynne 2004. A pixel is tagged as a treetop if it is the highest within each moving window. To consider different crown sizes, the size of the window changes depending on the height of the pixel on which the window is centered. Next, tree crowns were automatically outlined from these treetop points using the Marker-Controlled Watershed Segmentation (mcws) function developed by Meyer and Beucher (1990). Mcws implements the watershed function to segment or outline crowns from the CHM and segmentation is guided by the treetop point locations. I visually inspected tree crowns against the CHM and 0.1m camera imagery throughout the study area for several functions. A subset of the results from the final function I used for tree crown attribution can be seen in **Figure 3.4**. The final tree crowns selected for analysis only include trees with a height of 2m or taller. Paz Kagan et al. (2017) and Weinstein et al. (2021) found a minimum height of 2m to be an appropriate threshold in

including established trees.

Figure 3.4. Lidar-segmented tree crowns displayed on CHM (Left) and camera imagery (Right).

The L3 HSI reflectance product is an orthorectified, atmospherically corrected reflectance mosaic using the most nadir pixels from overlapping flight paths. I downloaded 100 L3 reflectance HDF5 files from <https://www.neonscience.org/data>. Each HSI file can be represented as a “data cube” that is 100 km x 100 km x 426 HSI bands. I extracted coordinate reference information, spatial extent, reflectance values, band centers, and band widths from the HDF5 files over the study area and mosaiced them using the Georeferenced Mosaicking tool in ENVI Classic.

A sample live tree’s spectral signature can be seen back in **Figure 2.1**. This reflectance spectra can be read as the reflectance, or the fraction of reflected light (scaled by 10,000) on the y-axis, recorded continuously across the electromagnetic spectrum (from 380nm to 2500nm) on the x-axis. Bands 1322nm - 1422 nm and 1808nm - 2039nm (**Figure 2.1**) represent regions of atmospheric absorption where certain gases absorb light and stop it from being recorded by the sensor. Because of this low recorded reflectance of the surface objects of interest, I removed these bands from the mosaic. I also removed the high noise bands 381nm - 401nm and 2449nm - 2509nm, from the mosaic, leaving 341

spectral bands. In a study also using the AVIRIS sensor, Zhang et al. (2019) found these same regions to be corrupted by Gaussian noise and impulse noise.

Shadowed pixels have low reflectance information of the features of interest (i.e., live and dead trees) and are thus not suitable for image classification. Shadow effects in imagery can be extreme in mountainous regions like this study area. I masked, or removed, these pixels prior to data reduction to maximize the variance of the features of interest in HSI-derived PC bands. I created a raster mask in which pixels to be classified were set to 1 and pixels to be masked or removed were set to 0. A reflectance value of 1200, or 12 percent of reflected light at 822 nm ensured the balance between removing shadowed pixels, while ensuring no features of interest were being masked.

I used PCA as a data reduction method due to limitations in storage and computing capacity. The goal was to reduce the number of bands used as inputs for classification, while maintaining as much spectral information, or variance, of the features of interest. PCA transformed the shadow-masked, 341-band HSI dataset to 30 PC bands containing nearly all the spectral variance in the original dataset. The first PC contained around 68% of the spectral variance, the second contained 21%, the third contained 3%, and so on (**Figure 3.5**).

Figure 3.5. PC Components one through ten, containing nearly all the total spectral variance of the original 341-band HSI dataset.

3.3 Image Classification:

I chose the supervised image classification algorithm, RF, for its processing speed and limited user-defined parameters. The goal of image classification was to identify pixels that contained dead trees and live trees, then identify the genus of those live tree pixels.

The first live/dead classifier predicted whether each pixel belonged to one of the following predominant landcover classes: (1) gray-stage mortality tree (2) red-stage mortality tree (3), live tree, (4) herbaceous groundcover, (5) soil, (6) rock, or (7) shrub. While both representing a dead tree, I initially separated trees in the red- and gray-stages of mortality into distinct classes because of their unique properties and spectral signatures. Tane et al. (2018) reached 85% OA using high-resolution WorldView2 imagery to extract ground-truth distinguishing between gray-stage and red-stage mortality trees. I primarily used the high-resolution camera imagery to initially locate samples for the live/dead classification.

The second genus classifier predicted whether the live tree pixels belong to the (1) oak, (2) pine, (3) incense-cedar, or the (4) other genus. I used Weinstein et al. (2021) in-situ KML survey points to locate samples for each genus class. Each ground-truth survey point included the date of collection, XY locations of each tree point, the species, and current health status of these tree. I performed two separate classifications because of the limited number and limited spatial extent of genus ground-truth samples. The situ survey points from Weinstein et al. (2021) were collected in 2015, two years before imagery acquisition, so the number of survey trees clearly visible in the HSI and CHM were

limited and spatially clustered in the center of the study area. Creating two classifiers with separate training and validation datasets ensured a balanced, or relatively equal number of training and validation pixels per class, and ensured minimal misclassified dead tree, live oak, live pine, and live incense-cedar pixels. This would help to maximize predictive accuracy for the evaluation of any genus-specific response in mortality probability.

I created two separate training/validation sets by drawing polygons around pixels of each class to be used for the two models' training and validation in ArcGIS. An image interpretation key (**Table 3.2**) shows a sample training polygon from each class in the 1m HSI true color image, 0.1m camera imagery, and 1m CHM.

Table 3.2. Image interpretation key for the live/dead and genus classifiers. True-color HSI (Left), camera imagery (Center), and CHM (Right) were the remote sensing reference sources used to extract class pixels for the live/dead (purple) and genus (yellow) classifiers. Genus samples were located using Weinstein et al. (2021) in-situ survey points.

CLASS	1m True-color HSI	0.1m Camera Imagery	1m CHM
GRAY-STAGE
RED-STAGE
LIVE TREE

GROUND COVER
SOIL
ROCK
SHRUB

<p>OAK</p> <p>QUCH2 - <i>Quercus chrysolepis</i> Liebm.</p>
<p>PINE</p> <p>PILA - <i>Pinus lambertiana</i> Douglas</p> <p>PIPO - <i>Pinus ponderosa</i> Lawson & C. Lawson</p>
<p>INCENSE -CEDAR</p> <p>CADE27 - <i>Calocedrus decurrens</i> (Torr.) Florin</p>
<p>OTHER</p> <p>ARVIM - <i>Arctostaphylos viscida</i> Parry ssp. mariposa (Dudley) P.V. Wells</p>

I took a stratified random sampling approach to extract pixels for training and validating the live/dead supervised classifier. I collected 10 samples for each class of interest (gray-stage, red-stage, live tree) in each of the six stratified zones. For the other predominant landcover classes (groundcover, shrubs, rock, and soil), I collected at least three sample polygons in each zone. In zones where a given land cover class (i.e., rock) appeared to be sparse or nonexistent in the imagery, I collected representatively less training in that zone. I drew 309 reference polygons for final live/dead classifier training and validation and 62 polygons for the RF genus classifier. For both live/dead and genus classification models, I split the ground-truth sample sets where two-thirds of the polygons in each class were selected for model training and the remaining one-third of the polygons were set aside for a cross-validation accuracy assessment. I implemented both RF classifiers using the ‘ranger’ package in R. To determine the best combination of RF parameters for the first live/dead classification, which would also be used for the second genus classification, I went through the process to change the combination of input bands. I referenced the output classification with the high-resolution camera imagery, in-situ survey data from Weinstein et al. (2021), and output confusion matrix accuracies.

Of the 30 PCs, I displayed different PC band combinations as an RGB image, and visually assessed their ability to clearly separate different classes difficult to distinguish using camera imagery, like herbaceous groundcover versus live trees, shrubs versus gray-stage mortality trees, or soil versus red-stage mortality trees. The PC composite image displays PC band 4 loaded in red, PC band 6 in green, and PC band 8 loaded into blue

(**Figure 3.6 Left**). Shrub features stand out from dead trees and live trees more in the PC composite than in the HSI and camera imagery. The right image shows a speckled pattern from sensor noise in higher order PC bands (**Figure 3.6 Right**), as there is no longer spectral information or patterns to pull apart from the reflectance image (**Figure 3.6 Center**). A combination of 13 1m PCA bands and the 1m CHM produced the most accurate results that best represent landcover over the study area.

Figure 3.6. PC composite PCs 4, 6, and 8 highlights differences in vegetation types (Left) not apparent in the HSI true color image (Center) and higher order PC 22 highlights sensor noise (Right).

I also tested the RF model parameter, nTree. I chose 200 for the final model because it performed similarly to the default value of 500 during model fitting. Changing the second RF parameter, mTry, has minimal effect on classification other than increasing computational time (Gislason 2006), so I chose the commonly used default value for the final model. The mTry default value is the square root of the number of input bands, so mTry for both classification models was the square root of 13 rounded down to the smallest integer, or 3.

Based on the live/dead classification raster, I assigned each tree crown polygon a class attribute based on the most frequently occurring value within a given polygon. I created a mask where I set live tree pixel values to 1 and pixel values other than live tree to 0. I applied a second genus classification to predict genus classes on the unmasked,

live tree pixels. Based on the genus classification raster, I assigned each live tree crown polygon a genus attribute based on the most frequently occurring value within a given polygon. Most crowns were initially attributed with an oak, pine, incense-cedar, or other genus based on the raster's most frequently occurring value in a given tree crown. However, some tree crowns had shrub, groundcover, soil, or rock majorities. Shrub pixels most often occurred where lidar-derived tree crowns segmented manzanitas taller than 2m. Less often, groundcover, soil, or rock pixels were the majority. This was a result of incorrectly classified pixels and small spatial misalignment. I removed these tree crowns so that there were only dead trees, live oaks, live pines, live incense-cedars, and live others. **Figure 3.7** shows a portion of the final classified raster and the remaining live attributed tree crowns. I reclassified the gray-stage and red-stage attribute to a single dead tree attribute for tree mortality modeling.

Figure 3.7. Lidar-derived tree crowns and classification on true color HSI (Top). Tree crowns with live, red-stage, or gray-stage majority attribute (Bottom).

3.4 Tree Mortality Modeling:

I assembled a set of environmental predictors as potential variables explaining forest mortality. The assembled topographic, substrate, and stand characteristics that represent different aspects of a tree's environment. Paz Kagan et al. (2017) found all climate variables to show high levels of correlation (> 0.5) with each other and elevation, so I chose elevation to represent itself and climate variables. All environmental factors assessed in the study, along with their associated units, spatial resolutions, and data sources can be seen in **Table 3.3**.

Table 3.3. Environmental factors considered for analysis for association with tree mortality.

Environmental Factors	Units	Spatial Resolution	Source
<i>Topography</i>			
Elevation	m	1m	Lidar point cloud
Slope	°	1m	Digital Elevation Model
Aspect	°	1m	Digital Elevation Model
Distance from streams	m	1m	Digital Elevation Model
<i>Stand Characteristics</i>			
Tree height	m	1m	Canopy Height Model
Forest cover	%	100 sq km	Live/Dead Classification
Density	count per 100 sq km	100 sq km	Canopy Height Model
<i>Substrate</i>			
Rock cover	%	100 sq km	Live/Dead Classification
Soil Suborder	classes	Vector	USDA Web Soil Survey 2019
Available Water Supply	cm	Vector	USDA Web Soil Survey 2019
Organic Matter content	cm	Vector	USDA Web Soil Survey 2019
Texture	cm	Vector	USDA Web Soil Survey 2019
Water pH	cm	Vector	USDA Web Soil Survey 2019

To match the spatial resolution of the image classification products (forest and rock cover), I created 1m DEM, slope, aspect, and stream rasters. I attributed the Elevation and Slope predictors based on the DEM value and DEM-derived slope value

occurring at each tree crown centroid. I reclassified aspect values to resemble insolation levels in the Sierra during summer months. I gave trees with the least evapotranspiration demands on northwest- and north facing slopes the lowest Aspect insolation score, 1. Trees on east facing slopes were given an Aspect insolation score of 2, trees on south facing slopes were given a score of 3, and trees on southwest facing slopes were given the highest score of 4. To calculate distance to nearest streams, I calculated the Euclidean distance from each tree crown centroid to the nearest stream pixel.

I spatially joined the height attribute from the vwf-identified treetop points. I calculated density from the total number of tree crowns divided by each 100,000 sq m hexagon. The hexagon represents a hexagonal area within 196 meter radius of a tree crown centroid. I calculated fractional forest cover from the number of red-stage pixels plus gray-stage pixels plus live-tree pixels divided by total pixels, multiplied by one hundred.

Rock cover has been used as an approximation of soil depth at the landscape scale (Pellenq et al. 2003). I calculated percent rock cover from the total number of classified rock pixels divided by the total number of classified pixels in each hexagon, multiplied by one hundred. I extracted substrate data including soil suborder, soil available water supply at 25 cm, 50 cm, 100 cm, an 150 cm, soil organic matter content, texture, and water pH from USDA Web Soil Survey vectors and attributed substrate variables based on the value at each tree crown centroid.

I used the ‘ranger’ package in R to create a RF regression model predicting tree mortality based on site-specific environmental variables with the permutation feature

importance function, nTree set to 200, and mTry set to 3. I applied this RF model such that each instance, a live or dead tree, was given a mortality probability (0 to 1) based on the mean prediction from all decision trees grown. An instance where most decision trees predicted “live” would have a mortality probability closer to 0 and an instance where most decision trees predicted “dead” would have a mortality probability closer to 1. I created PDPs with the partialPlot package in R to visualize the marginal effect each environmental predictor has on the RF model’s predicted mortality.

I joined the predicted mortality probability with the classified live pine, live oak, and live incense-cedar trees based on a unique identifier created and passed down from the vwf-identified treetop points. I then ran a one-way analysis of variance (ANOVA) to find any significant difference between genus composition and the probability of mortality. Finally, I applied post-hoc Tukey tests for group homogeneity to identify any statistical differences for each of the genus classes in relation to the probability of mortality.

CHAPTER 4: RESULTS

4.1 Live / Dead Tree Identification:

After removing trees shorter than 2m, I found 835,213 lidar-segmented tree crowns in the study area. Of these crown polygons, 604 had a majority groundcover pixels, 1,890 had a majority soil pixels, 327 had a majority rock pixels, and 63,704 had a majority shrub pixels. Tree mortality modeling did not include these non-tree crown polygons. **Figure 4.1** shows the remaining trees in the study area. Dead trees appeared clustered, particularly in the higher elevation forested regions. Within clusters were mixed gray-and red-stage mortality trees. There were 471,797 live trees and 296,891 dead trees in the study area in 2017 (**Figure 4.1**). Of the live trees, 254,066 were oaks, 63,125 were pines, 95,360 were incense-cedars, and 9,246 were an other genus.

Classified Tree Crowns

Figure 4.1. Lidar-derived tree crowns with a minimum height of 2m in the study area.

The confusion matrix from the live/dead classification (**Table 4.1**) shows the actual (column total) and predicted (row total) class pixel counts within the reference sample polygons that were set aside for model validation. The live/dead classification had a κ of 0.95 and OA of 96%. Gray-stage mortality trees had the lowest UA at 86%. This means there is an 86% chance that a pixel classified as gray-stage on the map represents a

gray-stage pixel on the ground. Of the 994 classified gray-stage pixels, 856 were gray-stage, 105 were red-stage, 32 were shrub, and 1 was a live tree. Rock had the highest UA at 100%, meaning all 1572 of the pixels classified as rock on the map were rock. Of the classes of interest, live trees had the highest UA at 99%. Of the 2195 classified live tree pixels, 2179 were live trees, 3 were herbaceous groundcover, 3 were gray-stage, and 1 was a red-stage mortality tree. Red-stage mortality trees had the lowest PA, or the lowest probability of a reference pixel being correctly classified. Of the 1096 red-stage validation pixels, 990 were correctly classified, 105 were incorrectly classified as gray-stage, and 1 was classified as a live tree. Live trees had a PA of 92%, where 2179 of the 2360 live tree validation pixels were correctly classified, 169 were classified as shrub, 11 were classified as red-stage mortality, and 1 was classified as a live tree.

Table 4.1. Confusion matrix containing class predictions for the withheld validation sample sets, OA, and κ for the live/dead classification raster.

Live/Dead	Gray- Stage	Red- Stage	Ground		Soil	Rock	Shrub	Count	UA
Gray-Stage	856	105	1	0	0	0	32	994	86.1
Red-Stage	10	990	11	0	4	0	14	1029	96.2
Live	3	1	2179	3	0	0	9	2195	99.2
Groundcover	0	0	0	741	11	0	0	752	98.5
Soil	0	0	0	1	954	0	0	955	99.8
Rock	0	0	0	0	0	1572	0	1572	100
Shrub	6	0	169	18	3	3	4142	4341	95.4
Count	875	1096	2360	763	972	1575	4197	OA	κ
PA	97.8	90.3	92.3	97.1	98.1	99.8	98.6	96.48	0.95

The genus classification had an OA of 89% and a κ of 0.85 (Table 4.2). The incense-cedar class had the lowest UA 73%. Out of the 450 classified incense-cedar pixels, 329 were incense-cedar, 71 were oak, 41 were pine, and 6 were the other genus.

Of the tree genus of interest, oaks and pines had the highest UA at 89% and 91%, respectively. Out of the 699 classified oak pixels, 624 were oak, 45 were incense-cedar, 12 were pine, and 18 were the other genus. Out of the 386 classified pine pixels, 350 were pine, 29 were incense-cedar, and 7 were pine. Oaks and pines both had a PA around 87%. Out of the 713 oak validation pixels, 624 were correctly classified as oak, 74 were incorrectly classified as incense-cedar, 8 were classified as other, and 7 were classified as pine. Out of the 403 pine validation pixels, 350 were correctly classified as pine, 12 were classified as oak, and 41 were classified as incense-cedar. incense-cedar had the lowest PA at 78%, meaning 78% of the incense-cedar validation pixels were correctly classified.

Table 4.2. Confusion matrix containing class predictions for the withheld validation sample sets, OA, and κ for the genus classification raster.

Genus	Oak	Pine	Cedar	Other	Count	UA
Oak	624	12	45	18	699	89.3
Pine	7	350	29	0	386	90.7
Cedar	74	41	329	6	450	73.1
Other	8	0	15	785	808	97.2
Count	713	403	418	809	OA	κ
PA	87.5	86.8	78.7	97.0	89.12	0.85

4.2 Tree Mortality Modeling:

One of the assumptions for RF regression models is that the model variables do not exhibit multicollinearity. Soil available water supply, texture, water pH, and organic matter had correlation coefficients above 0.50, so I chose for available water supply to represent itself and the other three correlated substrate variables. Elevation, slope, aspect, distance to streams, tree height, density, forest cover, rock cover, available water supply,

and soil suborder all had correlation coefficients below 0.5 and were used in tree mortality modeling (**Figure 4.2**). VIF is also used to measure the amount of multicollinearity in a set of multiple regression variables. Each of the ten previously mentioned variables had a VIF below 2 and were statically significant in a multiple regression model.

Figure 4.2. Correlation matrix between tree mortality and its environmental predictors.

The final RF model with aforementioned 10 environmental variables had an R-squared of .4132, or explained about 41% of the variance in tree mortality. The most influential environmental variables predicting tree mortality were tree height, tree count, and distance from streams. **Figure 4.3** shows the most influential variables in decreasing importance order. The permutation feature importance function uses percent increase in mean square error (MSE) to calculate the increase in error when randomizing each predictor variable. Tree height had the largest percent increase in MSE, meaning model performance dropped most when tree height was randomized, or excluded from model

training. The next most influential predictors were forest cover followed by elevation, density, distance to streams, rock cover, soil subgroup, available water supply, slope, then aspect reclassified to match insolation levels (**Figure 4.3**).

Figure 4.3. RF variable importance plot.

To illustrate the nature of these relationships, PDPs were created for each environmental predictor. All other conditions are controlled in PDPs, so each plot shows the causal-effect relationship each environmental variable has on mortality probability. In each PDP, the change in mortality probability is displayed over the range of one predictor. **Figure 4.4** shows this relationship for the assessed environmental variables.

Figure 4.4. PDPs displaying mortality probability by each individual environmental predictor: Tree height, forest cover, elevation, density, distance to streams, slope, aspect, rock cover, soil available water supply, and soil suborder.

These results suggest low mortality probability in trees up to around 10m, then tree mortality increases in trees 20m to 40m tall, where it is highest and plateaus for the trees taller than 40m. These results suggest a forest cover has an inverse relationship with mortality probability. When percent forest cover is less than 40 percent, tree mortality is the highest. When forest cover increases beyond 40 percent, tree mortality decreases at nearly a linear rate. Mortality probability tends to slightly increase with elevation, though there is a small dip around 1400m to 1800m. These results suggest that tree mortality has a positive linear relationship with tree density. As the density, or number of trees per area increases, so does the probability of mortality of a given tree. These findings also show a positive linear relationship between tree mortality and distance to streams. The model predicted trees closer to streams to have lower mortality probabilities and trees further from streams to have higher mortality probabilities. Mortality probability tended to decrease in relation to slope, then reach lowest rates at the steepest slopes greater than 40 percent. This model found higher rates of mortality on southwest facing slopes than on south facing slopes, but the highest mortality rates on north facing slopes. Percent rock cover acts as a proxy for soil depth and water availability. These results suggest a high

mortality probability in areas with very low rock cover, under 5%, yet predict it rapidly decreases and plateaus by 20% rock cover. Generally, this model predicted higher mortality rates in soils with a greater soil available water supply. The tree mortality found higher rates of mortality among trees in xerolis soil suborder.

I calculated the average prediction value, or mortality probability, by genus class for all trees in the study area. The mean mortality probability for the live oaks was 0.39, 0.43 for pines, and 0.26 for incense-cedars. **Figure 4.5** shows the Tukey HSD results between mortality probability and the three genus classes. The CI for the oak-pine pair group mean difference overlaps with the vertical zero line, so I can conclude that the difference between the oak and pine genus pair group means is not significantly different but the mortality probability between oaks and incense-cedars as well as pines and incense-cedars is significantly different.

Figure 4.5. Tukey's HSD test displaying an analysis of variance of mortality probability by genus.

CHAPTER 5: DISCUSSION

5.1 Live / Dead Tree Identification:

The ability to successfully fuse this information relies on minimizing their spatial misalignment. Low spatial offsets between remote sensing data products was achieved through co-mounting and co-acquiring data with the NEON AOP. This can be seen by overlaying the lidar-derived CHM with the HSI-derived true color image, but can also be illustrated by the small percentage of tree crowns that were over a majority non-tree (soil, rock, or herbaceous groundcover) class pixel in the live/dead classification. It should be noted that any spatial misalignment likely affected tree crowns with smaller areas, thus undercounting the number of short, young trees or gray-stage mortality trees with generally smaller canopy areas. Based on visual inspection, the lidar-segmented tree crowns performed particularly well for stand-alone conical-shaped crowns. However, using remote sensing data products alone it is difficult to visually assess how the segmentation algorithm performed in homogeneous, dense, closed canopies.

I created two image classifications: The first classified live and dead tree pixels and the second used the location of these live tree pixels and classified their genus. Based on the confusion matrix, gray-stage and red-stage mortality pixels were being misclassified, or confused for one another, but those pixels were remapped to a single dead-tree class before applying a majority class attribution. Other confusion occurred between the live tree and shrub classes, but the model still predicted the live tree class with 90% Producer's Accuracy. κ was expected to be lower for the genus classification than the live/dead classification such that with fewer classes it was likely that more of the correctly classified pixels were due to random chance. κ was more in agreement with the

OA in the live/dead classification, illustrating this random chance correct classification was less common with additional classes.

The largest percentage of misidentified tree crowns over pixels of the shrub class can be explained by the manzanita shrubs in the area that grow over 2m. Trees less than 2m were removed, as Paz Kagan et al. (2017) and Weinstein et al. (2021) found this height to be appropriate in including established trees and removing tall shrubs. However, Manzanitas, a dominant cover type in the study area, can grow up to 3m. Thus, it is likely some tree crowns identified without spectral information would include these shrubs taller than 2m. More of the shrubs that may have mistakenly been identified as live tree were likely removed in the genus-specific classification when their PCs resembled the spectral information for shrubs more closely than oaks, pines, or incense-cedars. **Figure 5.1** shows how red-stage mortality and gray-stage mortality trees look different, with lower reflectance of gray-stage mortality trees in the red (625–700 nm) and NIR (800–2,500 nm) region of the spectrum.

Figure 5.1. Average class spectra for ground-truth pixels used in RF classification training and testing. Red-stage and gray-stage mortality trees show visible differences in the VIS and NIR parts of the spectrum. Gray-stage mortality trees have a similar spectral signature to shrubs, red-stage mortality trees have a similar signature to soil, and live trees have a similar signature to groundcover.

Asner et al. 2012 demonstrated that precise data fusion can greatly increase the dimensionality of ecological information derived from remote sensing data products. Similarly, the three remote sensing data products in this study provided rich information for vegetation monitoring. Each sensor's strength played to another's weakness, as the high-resolution camera imagery made it possible to locate and extract accurate, precise ground-truth samples to be used for image classification training and validation. Additionally, features that were indistinguishable in the resolution camera imagery become apparent with more spectral information, seen from the PC composite in **Figure 3.6**. The spectral resolution of the HSI dataset also made it possible to distinguish between dominant genres in the region (oaks, pines, incense-cedars, and manzanitas). Meanwhile, the addition of lidar height information likely helped to separate any spectrally similar classes (**Figure 5.1**) such as gray-stage mortality trees between shrubs, red-stage mortality trees between soil, and live trees between herbaceous groundcover.

5.2 Tree Mortality Modeling:

The highest adjusted R-squared achieved by the tree mortality model was 0.41. According to these results, the environmental factors account for around 41% of the variance of tree mortality. Because the study area is known to have had bark beetle outbreaks, I would attribute a large portion of the unexplained variance in mortality probability to areas of direct, mass bark beetle attack. I would attribute a smaller portion of the unexplained variance to data uncertainty because spatial offsets and misclassifications were shown to be small.

Ecological data is highly complex and has nonlinear interactions across different spatial and temporal scales (Fortin et al. 2012). Using the same environmental predictors I created a RF regression model predicting the mortality fraction, or the number of dead trees divided by the total number of trees per 150m by 150m plot. That model had an adjusted R-squared of 0.86, or explained 86% of the variance in the mortality fraction. The environmental predictors used in this study and their spatial resolution more accurately predicted the fraction of dead trees per area. Thus, the factors at play are likely spatially heterogeneous and scale-dependent. Over the wide elevational and climatic gradients of the Sierra, it is likely that the same variables operate differently depending on location (Fettig et al. 2019).

The low model fit may have an impact on the variable importance and PDP results so such results should be interpreted with reservations. Tree height, forest cover, elevation, and density had the strongest influence when predicting tree mortality. These results agreed with previous research by Stovall et al. (2019), in which the author found tree height to be the main predictor of drought-induced tree mortality. The higher probability of mortality with increasing tree height may be driven largely by *Pinus lambertiana* and *Pinus ponderosa*, for which the primary bark beetle attackers tend to select larger trees (Paz-Kagan et al. 2017). Tree height as an influential predictor in extreme drought conditions supports the framework supposed by McDowell et al. 2008, where hydraulic failure can be the primary driver of tree mortality during extreme drought conditions.

In the tree height vs. mortality probability plot (**Figure 4.4**), the small spike in the shorter trees may be a result of hydraulic failure from the constrained hydraulic

conductance in younger trees with small rooting volumes (McDowell et al. 2008). The greatest mortality probabilities in the tallest trees may be also be a result of hydraulic failure in tall trees with a long hydraulic path length and higher evapotranspiration demands higher up in the canopy (McDowell et al. 2008).

Elevation determines climatic conditions like maximum summer temperatures and winter rainfall. Higher mortality probabilities in the higher elevations around 1400m and 1800m may be attributed to these ranges typically receiving greater rainfall in non-drought years, so they didn't have the adaptations for sources of water or for conserving water such as deep roots or stomatal closure.

The stand density variable indicated a higher mortality probability in denser stands. In portions of western North America, over a century of fire suppression has fostered the buildup of unusually high tree densities. These unnaturally dense forests may have decreased vigor, which can increase their vulnerability to drought-induced tree mortality (Allen et al 2010). Trees in dense stands may have also attracted bark beetles and succumb to a mass bark beetle attack.

The distance to stream variable indicated a higher probability of mortality as distance to streams increased. This may indicate a reliance on the remaining soil moisture and groundwater levels to be adequate for trees closest to streams. Trees further from streams may not have had access to this last remaining soil moisture.

This model predicted higher mortality probabilities in areas with low rock cover, or in areas with deeper soils. Trees in deeper soils may allocate fewer resources to developing drought-resistant characteristics such as reduced leaf area and cavitation resistance. Lower mortality probabilities were predicted on areas with high rock cover, or

in areas with shallow soils. This may be due to trees' access to short-term water storage in surface soils as well as long-term storage in rock fissures, or less permeable soil layers, which retain water for longer during the growing season.

Higher percent forest cover indicated a lower mortality probability, possibly because areas with high percent forest cover must have a lower percent rock cover and inversely areas with a high percent rock cover must have a low percent forest cover. This may why trees in areas with a high percent forest cover had low mortality probabilities and low percent forest cover had high mortality probabilities.

Slope, Aspect (reclassified), Soil subgroup and Soil AWS variables were largely unproductive. Soil subgroup and Soil AWS were expected to be uninfluential due to the mismatch in spatial resolution between the large USDA Web Soil Survey vectors and high-resolution remote-sensing data.

These results found highest rates of mortality among pines and oaks. These results are similar, but not exactly to Paz Kagan et al. (2017), who found significantly highest rates of mortality among pines in the Sierra Nevada. It is possible in the central Sierra Nevada oak may become the next most threatened tree in this region. Linking the spatial patterns of forest dieback to species composition is essential for understanding long-term shifts among species and their associated communities (Adams et al. 2012).

5.3 Practical Implications:

This study created a framework for future studies to use open source lidar and HSI to automatically detect live and dead trees. In doing so, this research has provided a finer scale mapping of tree mortality over SOAP than is provided by the US Forest

Service ADS. This stresses the utility of open-source HSI and lidar, and the need for funding for national research organizations that help provide this. The live/dead classification and corresponding structure information provides a more accurate model of fuel type over the area, which can be used to better model fire risk. The classified locations of live trees and dead trees can provide immediate information regarding the largest threat to powerlines during wind events and fire season. More information regarding how much and what type of fuel, (i.e., dead tree, shrub, conifer, broadleaf, sparsely vegetated soil, and herbaceous groundcover) along with its corresponding structure information can be used to help more accurately assess wildfire risk. This information can be used to help assess wildfire risk and modeling in surrounding communities like Shaver Lake, which was affected by the 2020 Creek Fire.

Climate change is projected to bring more prolonged hot droughts to the region. This more robust information can help forest management and conservationists assess where future die-off events are more likely. Because the relative influence of environmental factors on tree mortality during drought conditions is not well understood and largely site-specific, these results will add to this knowledge base, but can be used can be used to predict areas in the Sierra Nevada that have a higher probability of mortality in the coming years, as California continues to experience record-breaking hot, dry years in 2019 and 2020.

These results suggest that conifer communities (i.e., ponderosa pines) with a greater strain on resources from competition and other topographic resource-limiting factors are more at risk to die-off events, which could reduce forest productivity, increase fire hazard risk, and drive a shift in community composition over the long-term. Oaks

were also found to have higher rates of mortality. Sudden Oak Death is a forest disease caused by a plant pathogen that is known to infect black oak and canyon live oak, the predominant oaks in the study area, thus, more practices may need to be put in place to monitor for these trees. After a disturbance event, forests may show signs of resilience and recovery or they may show signs of recovery. These findings can be included in forest stand assessments to begin monitoring shifts in ponderosa pine dominated forests, which over the long term could alter biodiversity and ecosystem function.

5.4 Limitations:

Limitations of this study primarily relate to incomplete inputs into the regression model, which likely added to the model's unexplained variance. First, due to the nature of aerial remote sensing acquisition, the identification of dead and live trees is limited to the top of the canopy. It is likely many subcanopy, shorter trees were not classified as live or dead and were therefore not included in resulting tree mortality modeling. These results also cannot speak to the influence of environmental variables on the mortality probability of any young trees less than 2m. A high-resolution camera image mosaic of the study area in 2013 was inspected to ensure the majority of dead trees in 2017 were alive in 2013. However, because there were still some dead trees visible in 2013, this means their death was likely not a result of the 2012 - 2016 drought.

Another limitation of this assessment of model performance is that few ground-truth genus samples were available over only the central portion of the image. Because field collection occurred two years prior to remote sensing data acquisition, samples were chosen for training with caution, thus only a limited number were used. This also makes

it easy to assess model accuracy, but difficult to assess model precision, because the confusion matrix does not assess the results on the edges of the image where there are no test samples.

Further, live trees in the classification are not separated by species, masking any species-specific tree mortality response. Different species respond differently to drought conditions, as some are more or less adept at conserving water, acquiring nearby water, or surviving without it. For instance, sugar pine and ponderosa pine have both been found to have a high probability of mortality Paz-Kagan et al. (2017) in nearby communities, but this study masks any possible differences between these groups.

Site-specific environmental factors, like those related to water availability, can mediate the effects of droughts to help limit forest die off. However, drought itself is only one of several potential mechanisms of mortality. In many cases, drought induces stress within trees, which then succumb to death directly from bark beetle attacks (Raffa et al. 2008, US Forest Service 2017b). The model does not directly consider this source of mortality in the region. Though, it does assess environmental factors which have an indirect effect on tree mortality, as these factors mitigate tree stress to limit bark beetle attacks and ensure the positive feedback does not occur.

5.5 Future Work:

The combined live/dead and genus classification maps from this study could provide more robust information to model fire patterns. Future work could assess the importance of the distinction between stages of mortality in wildfire modeling. For instance, gray-stage mortality trees may limit the spread of crown fires versus red-stage

mortality trees with highly flammable needles. Future work could also identify trees that died before the 2012-2016 drought. This study's methods could be replicated using 2012 NEON AOP data over the study area to identify trees that did not die as a result of the 2012-2016 drought and should be removed from this tree mortality model. Any dead trees below a closed canopy weren't identified by this study, so field work could assess the extent of multi-level canopy mortality that was not included in this RF tree mortality model. Field work could be done to gather more survey points to perform this species classification. A species-specific classification would allow the analysis of influential environmental predictors to control for species-specific responses, which could include vegetation indices to enhance classification model performance.

Field work could also check for the cause of mortality in a subset of trees in the study area. Tall, dead trees may be attributed to natural mortality and not necessarily from increased demands from the prolonged hot drought. Additionally, areas with bark beetle damage can be compared to areas of high residuals in the RF tree mortality model. If with high model residuals are spatially autocorrelated with areas of bark beetle damage, then one could confidently attribute that unexplained model variance to direct bark beetle attack.

Existing projections of tree mortality are based on models that lack functionally realistic mortality mechanisms (Allen et al. 2010). Spatially extensive tree-level assessments in specific forest types are needed to substantially improve future predictions of global tree mortality (Allen et al. 2010). As more satellite-based HSI sensors such as HyMap and HypIri are launched in coming years, this methodology can be used to help provide repeat coverage of forests over the world. Prediction of mortality events remains

difficult because the physiological mechanisms underlying drought survival and mortality are poorly understood. The most important future work could include an integrative approach that includes climate, biotic factors and site-specific conditions that may enable us to predict future mortality patterns in response to drought conditions (McDowell et al. 2008).

CHAPTER 6: CONCLUSION

With the increased occurrence and severity of droughts, understanding how water stress is linked to future tree mortality across forest types has become increasingly important for forest management, conservation, and resource policy. Forest responses to drought depend on local water availability, temperature conditions, and pest outbreaks in combination with environmental conditions (i.e., topographic, substrate, and stand characteristics), species-specific functional traits, and intra-species specific variation. This research used publicly available remote sensing data to map drought-induced tree mortality and then assess areas and tree genus of high risk for future events by finding that tree height, elevation, and density were strongest predictors related to mortality probability. Results also found higher mortality rates in pines and oaks, meaning further widespread die-off of these trees could reduce forest productivity, increase fire hazard risk, and drive a shift in community composition over the long-term. This study provides a finer resolution mapping of tree mortality over the research area than was reported by the USFS Aerial Detection Survey. Given the potential risks of climate induced forest die-off, forest managers need to develop adaptation strategies to improve the resistance and resilience of forests to projected increases in climate stress (Allen et al. 2010). Because this study is site-specific, more research is needed to assess how environmental factors mediate drought-induced mortality in other regions also projected to have more intense droughts as a result of climate change. Understanding how interacting environmental conditions explain the spatial patterns of mortality may improve our ability to assess forest response to drought stress.

References

- Allen, C. D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N. 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.
- Allen, C. D., Breshears, D. D., and McDowell, N.G. 2015. On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere*. 6(8):129.
- Adams, H. D., C. H. Luce, D. D. Breshears, C. D. Allen, M. Weiler, V. C. Hale, A. M. S. Smith, T. E. Huxman. 2012. Ecohydrological consequences of drought- and infestation- triggered tree die-off: insights and hypotheses. *Ecohydrology*. 5:145–159.
- Ahmed, O. S., M. A. Wulder, J. C. White, T. Hermosilla, N. C. Coops, and S. E. Franklin. 2016. Classification of annual non-stand replacing boreal forest change in Canada using Landsat time series: a case study in northern Ontario. *Remote Sensing Letters*. 8 (1):29–37.
- Allen, C. D. 2007. Interactions across spatial scales among forest dieback, fire, and erosion in Northern New Mexico landscapes. *Ecosystems*. 10:797–808.
- Allgöwer, B., van der Linden, S., Curt, T., Koetz, B., Morsdorf, F., van der Linden, S. Morsdorf, F. 2008. Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecology and Management*. 256(3): 263–271.
- Alonzo, M., McFadden, J. P., Nowak, D. J., & Roberts, D. A. 2016. Mapping urban forest structure and function using hyperspectral imagery and lidar data. *Urban Forestry and Urban Greening*. 17:135–147.
- Arroyo, L. A., Pascual, C., & Manzanera, J. A. 2008. Fire models and methods to map fuel types: The role of remote sensing. *Forest Ecology and Management*. 256(6):1239–1252.
- Asner, G. P., Knapp, D. E., Boardman, J., Green, R. O., Kennedy-Bowdoin, T., Eastwood, M. et al. 2012. Carnegie Airborne Observatory-2: Increasing science data dimensionality via 512 high-fidelity multi-sensor fusion. *Remote Sensing of Environment*. 124:454–465.
- Asner, G. P., Brodrick, P. G., Anderson, C. B., Vaughn, N., Knapp, D. E., & Martin, R. E. 2015. Progressive forest canopy water loss during the 2012 – 2015 California drought. *PNAS*. E249–E255.

- Baguskas, S. A., S. H. Peterson, B. Bookhagen, and C. J. Still. 2014. Evaluating spatial patterns of drought-induced tree mortality in a coastal California pine forest. *Forest Ecology and Management*. 315:43–53.
- Bigler, C., D. Kulakowski, and T. T. Veblen. 2005. Multiple disturbance interaction and drought influence fire severity in rocky mountain subalpine forest. *Ecology*. 86:3018–3029.
- Belgiu, M., and Dra, L. 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114:24–31.
- Breiman, L. 2001. Random forest. *Machine Learning*. 45.
- Breshears, D. D., L. Lopez-Hoffman, and L. J. Graumlich. 2011. When ecosystem services crash: preparing for big, fast, patchy climate change. *Ambio*. 40:256–263.
- Chang, H., Bonnette, M.R. 2016. Climate change and water-related ecosystem services: Impacts of drought in California, USA. *Ecosystem Health and Sustainability*. 2:e01254.
- Clark, M.L. and Roberts, D.A., 2012. Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sensing*. 4 (6):1820–1855.
- Clevers, J. G. P. W., Kooistra, L., & Schaepman, M. E. 2010. Estimating canopy water content using hyperspectral remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*. 12(2):119–125.
- Corcoran, J., J. Knight, and A. Gallant. 2013. Influence of Multi-Source and Multi-Temporal Remotely Sensed and Ancillary Data on the Accuracy of Random Forest Classification of Wetlands in Northern Minnesota. *Remote Sensing*. 5(7):3212–3238.
- Dees M., C. Straub, and B. Koch. 2012. Can biodiversity study benefit from information on the vertical structure of forests? Utility of LiDAR remote sensing. *Current Science*. 102(8):1181–1187.
- Dorman, M., T. Svoray, A. Perevolotsky, Y. Moshe, and D. Sarris. 2015. What determines tree mortality in dry environments? A multi-perspective approach. *Ecological Applications*. 25:1054–1071.
- Evans, J. S., M. A. Murphy, Z. A. Holden, and S. A. Cushman. 2011. Modeling Species Distribution and Change Using Random Forest. *Landscape Ecology*. 139–159.
- Fassnacht, F. E., Latifi, H., Ghosh, A., Joshi, P. K., & Koch, B. 2014. Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. *Remote Sensing of Environment*. 140:533–548.

- Fettig, C.J., Klepzig, K.D., Billings, R.F., Munson, A.S., Nebeker, T.E., Negrón, J.F., Nowak, J.T. 2007. The effectiveness of vegetation management practices for prevention and control of bark beetle infestations in coniferous forests of the western and southern United States. *Forest Ecology and Management*. 238:24–53.
- 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, U.S. *Forest Ecology and Management*. 432:164-178.
- Flake, S. W., and Weisberg, P. J. 2019. Fine-scale stand structure mediates drought-induced tree mortality in pinyon–juniper woodlands. *Ecological Applications*. 29(2):e01831.
- Fortin, M., James, P. M. A., Mackenzie, A., Melles, S. J., & Rayfield, B. 2012. Spatial statistics, spatial regression, and graph theory in ecology. *Spatial Statistics*. 1:100–109.
- Friedman, Jerome H. Greedy function approximation: A gradient boosting machine. 2001. *Annals of statistics*. 1189-1232.
- Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R.. 2006. Random forests for land cover classification. *Pattern Recognition Letters*. 27:294-300.
- Graham, R., Harvey, A., Jain, T., Tonn, J., 1999. Effects of thinning and similar stand treatments on fire behavior in western forests. USDA Forest Service, Pacific Northwest Research Station, General Technical Report PNW-GTR-463.
- Goetz, A. F. H., Kindel, B. C., Kaufmann, M. R., Jia, G. J., & Burke, I. C. 2006. Assessing spatial patterns of forest fuel using AVIRIS data. *Remote Sensing of Environment*. 102(3):318–327.
- Gunst, K. J., Weisberg, P. J., Yang, J., & Fan, Y. 2016. Forest Ecology and Management Do denser forests have greater risk of tree mortality: A remote sensing analysis of density-dependent forest mortality. *Forest Ecology and Management*. 359:19–32.
- Hamlin, L., Green, R. O., Mouroulis, P., Eastwood, M., Wilson, D., Dudik, M., & Paine, C. 2011. Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments. *IEEE Aerospace Conference Proceedings*. 1–7.
- Horning, N. 2010. Random Forests: An algorithm for image classification and generation of continuous fields data sets. Proceeding of International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences.
- Hughes, G. 1968. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*. 14 (1):55–63.

- Ingram, K., Kocher, S. 2015. Natural History of the Sierra Nevada. *UC Agriculture and Natural Resources*. 8535:1-19.
- Jenkins MJ, Hebertson E, Page W, Jorgensen WP. 2008. Bark beetles, fuels, fires and implications for forest management in the intermountain west. *Forest Ecology and Management*. 254:16–34.
- Jensen, J.R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd Edition, Prentice Hall, Upper Saddle River.
- Khosravipour, Anahita & Skidmore, Andrew & Isenburg, Martin & Wang, Tiejun & Hussin, Yousif. 2014. Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogrammetric Engineering & Remote Sensing*. 80:863-872.
- Koepke, D. F., T. E. Kolb, and H. D. Adams. 2010. Variation in woody plant mortality and dieback from severe drought among soils, plant groups, and species within a northern Arizona ecotone. *Oecologia*. 163:1079–1090.
- Lawrence, R. L., Wood, S. D., & Sheley, R. L. 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment*. 100(3):356–362.
- Luce, C. H., Vose, J. M., Pederson, N., Campbell, J., Millar, C., Kormos, P., & Woods, R. 2016. Contributing factors for drought in United States forest ecosystems under projected future climates and their uncertainty. *Forest Ecology and Management*. 380:299–308.
- Lund, J., Medellin-Azuara, J., Durand, J., & Stone, K. 2018. Lessons from California's 2012-2016 drought. *Journal of Water Resources Planning and Management*. 144:10.
- Madakumbura, G. D., Goulden, M., Hall, A., Fu, R., Moritz, M., Koven, C., & Kueppers, L. 2020. Recent California tree mortality portends future increase in drought-driven forest die-off. *Environmental Research Letters*. 15:124040.
- Makantasis, K., K. Karantzalos, A. Doulamis, and N. Doulamis. 2015. Deep supervised learning for hyperspectral data classification through convolutional neural networks. 2015. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 4959–4962.
- Malila, W.A. 1980. Change Vector Analysis: An Approach for Detecting Forest Changes with Landsat. LARS Symposia. 385.
- McDowell, N. G., D. J. Beerling, D. D. Breshears, R. A. Fisher, K. F. Raffa, and M. Stitt. 2011. The interdependence of mechanisms underlying climate-driven vegetation mortality. *Trends in Ecology & Evolution*. 26:523–532.

- McDowell, N., Pockman, W.T., Allen, C.D., Breshears, D.D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D.G., Yepez, E.A. 2008. Mechanisms of plant survival and mortality during drought: Why do some plants survive while others succumb to drought?. *New Phytologist*. 178: 719-739.
- McHugh M.L. 2012. Interrater reliability: the kappa statistic. *Biochem Med*. 22(3):276-82.
- Meddens, A.J.H., Hicke, J.A., Macalady, A.K., Buotte, P.C., Cowles, T.R., Allen, C.D. 2015. Patterns and causes of observed pinyon pine mortality in the southwestern United States. *New Phytologist*. 206: 91–97.
- Melgani, F., and Bruzzone, L. 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*. 42(8):1778–1790.
- Meyer, F., and Beucher, S. 1990. Morphological segmentation. *Journal of visual communication and image representation*. 1(1):21-46.
- Morin, M., Powell, S., Repasky, K., Sterling, T., Lawrence, R., McCann, C. 2017. Agreement analysis and spatial sensitivity of multispectral and hyperspectral sensors in detecting vegetation stress at management scales. *Journal of Applied Remote Sensing*. 11(4).
- Moore, J. 2108. Aerial Detection Monitoring. Region 5 - Forest & Grassland Health https://www.fs.usda.gov/detail/r5/forest-grasslandhealth/?cid=fsbdev3_046696.
- Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. *ISPRS Journal of Photogrammetry and Remote Sensing*. 66:247–259.
- Ogle, K., T. G. Whitham, and N. S. Cobb. 2000. Tree ring variation in pinyon predicts likelihood of death following severe drought. *Ecology*. 81:3237–3243.
- Oo, A. N. 2018. Comparative Study of Principal Component Analysis (PCA) based On Decision Tree Algorithms. *International Journal of Advances in Scientific Research and Engineering*. 4(6): 122–126.
- Pan, Y., R. A. Birdsey, O. L. Phillips, and R. B. Jackson. 2013. The structure, distribution, and biomass of the world's forests. *Annual Review of Ecology, Evolution, and Systematics*. 44:593–622.
- Paz-Kagan, T., P.G. Brodrick, N.R. Vaughn, A.J. Das, N.L. Stephenson, K.R. Nydick, and G.P. Asner. 2017. What mediates tree mortality during drought in the southern Sierra Nevada?. *Ecological Applications*. 27(8):2443–2457.
- Peerbhay, K. Y., O. Mutanga, and R. Ismail. 2015. Random Forests Unsupervised Classification: The Detection and Mapping of Solanum mauritianum Infestations

- in Plantation Forestry Using Hyperspectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 8 (6):3107–3122.
- Pellenq, J., J. Kalma, G. Boulet, G. M. Saulnier, S. Wooldridge, Y. Kerr, and A. Chehbouni. 2003. A disaggregation scheme for soil moisture based on topography and soil depth. *Journal of Hydrology*. 276:112–127.
- Popescu, S. C., and Wynne, R. H. 2004. Seeing the trees in the forest. *Photogrammetric Engineering & Remote Sensing*. 70(5):589-604.
- Prasad, A. M., Iverson, L. R., & Liaw, A. 2006. Newer Classification and Regression Tree Techniques : Bagging and Random Forests for Ecological Prediction. *Ecosystems*. (9):181–199.
- Pureswaran, D.S., Roques, A. & Battisti, A. 2018. Forest Insects and Climate Change. *Current Forestry Reports*. 4:35–50.
- 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*. 58:501–17.
- Rodarmel, C., and Shan, J. 2002. Principal Component Analysis for Hyperspectral Image Classification. *Surveying and Land Information Systems*. 62(2):115-123.
- Safford, H.D., Schmidt, D.A., Carlson, C.H., 2009. Effects of fuel treatments on fire severity in an area of wildland–urban interface, Angora Fire, Lake Tahoe Basin, California. *Forest Ecology Management*. 258:773–787.
- Scholl, V.M., Cattau, M.E., Joseph, M.B., Balch, J.K. 2020. Integrating National Ecological Observatory Network (NEON) Airborne Remote Sensing and In-Situ Data for Optimal Tree Species Classification. *Remote Sensing*. 12(9):1414.
- Soaproot Saddle. 2019. NSF NEON | Open Data to Understand our Ecosystems. <https://www.neonscience.org/field-sites/field-sites-map/SOAP> (last accessed 29 April 2019).
- Swatantran, A., Dubayah, R., Roberts, D., Hofton, M., & Blair, J. B. 2011. Mapping biomass and stress in the Sierra Nevada using lidar and hyperspectral data fusion. *REMOTE SENSING OF ENVIRONMENT*. 115(11, SI).
- Stovall, A.E.L., Shugart, H. & Yang, X. 2019. Tree height explains mortality risk during an intense drought. *Nat Commun*. 10:4385.
- Tane, Z., Roberts, D., Koltunov, A., Sweeney, S., Ramirez, C. 2018. A framework for detecting conifer mortality across an ecoregion using high spatial resolution spaceborne imaging spectroscopy. *Remote Sensing of Environment*. 209:195–210.

- Thenkabail, P.S. 2017. "Hyperspectral remote sensing of vegetation: knowledge gain and knowledge gap after 50 years of research (Conference Presentation)", Proc. SPIE 10213, Hyperspectral Imaging Sensors: Innovative Applications and Sensor Standards 2017, 102130C.
- USDA Web Soil Survey. 2019. Web Soil Survey.
<https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx>
- US Forest Service(a). 2017. Tree Mortality Aerial Detection Survey Results.
<https://www.fs.usda.gov/detail/catreemortality/toolkit/?cid=FSEPRD565838>.
- US Forest Service(b). 2017. Our response. California Tree Mortality.
<https://www.fs.usda.gov/main/catreemortality/response>.
- Ustin, S. L., D. A. Roberts, J. A. Gamon, G. P. Asner, and R. O. Green. 2004. Using Imaging Spectroscopy to Study Ecosystem Processes and Properties. *BioScience*. 54(6):523–534.
- Vaglio Laurin, G., Chen, Q., Lindsell, J. A., Coomes, D. A., Frate, F. Del, Guerriero, L., Valentini, R. 2014. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing*. 89:49–58.
- Verrelst, J., Camps-Valls, G., Munoz-Mari, J., Rivera, J.P., Veroustraete, F., Clevers, Jan. G.P.W., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - a review. *Journal of Photogrammetry and Remote Sensing*. 108:273–290.
- Wang, H., Y. Zhao, R. Pu, and Z. Zhang. 2015. Mapping Robinia Pseudoacacia Forest Health Conditions by Using Combined Spectral, Spatial, and Textural Information Extracted from IKONOS Imagery and Random Forest Classifier. *Remote Sensing*. 7 (7):9020–9044.
- Weinstein, B., Marconi, S., Bohlman, S., Zare, A., Singh, A., Graves, S., and White, E. 2021. A remote sensing derived data set of 100 million individual tree crowns for the National Ecological Observatory Network. eLife. 10. 10.7554/eLife.62922
<https://elifesciences.org/articles/62922#sa1>.
- Wiens, J., Sutter, R., Anderson, M., Blanchard, J., Barnett, A., Aguilar-Amuchastegui, N., Avery, C., Laine, S., 2009. Selecting and conserving lands for biodiversity: the role of remote sensing. *Remote Sensing of the Environment*. 113:1370–1381.
- Williams, A.P. 2015. Contribution of anthropogenic warming to California drought during 2012 – 2014. *Geophysical Research Letters*. 42(16): 6819–6828.
- Williams, A.P., Macalady, A.C, Macalday, A.K. 2013. Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change*. 3:292–297.

- Yang, C. and Everitt, J. H. 2010. Mapping three invasive weeds using airborne hyperspectral imagery. *Ecological Informatics*. 5(5):429–439.
- Zhang, Q., Yuan, Q., Li, J., Liu, X., Shen, H. and Zhang, L. 2019. Hybrid Noise Removal in Hyperspectral Imagery With a Spatial–Spectral Gradient Network. *IEEE Transactions on Geoscience and Remote Sensing*. 57(10):7317-7329.
- Zhang, B., MacLean, D., Johns, R., & Eveleigh, E. 2018. Effects of Hardwood Content on Balsam Fir Defoliation during the Building Phase of a Spruce Budworm Outbreak. *Forests*. 9:530.