Classifying Blue Oaks through Geographic Object-Based Image Analysis for Biomass, Canopy Cover, and Evapotranspiration Estimations

Megan Y. Hur

ABSTRACT

As our environment changes, there is an increased need to quantify the effects on our ecosystems and verify the nature-based climate solutions combating them. The current indicative measures, such as biomass estimates, are largely dependent on field-based measurements. However, trees can be detected and characteristics can be measured in remote sensing images, allowing for more data to become available. This study focused on the extent to which computer vision algorithms' classification of blue oak trees (*Quercus douglasii*) in remote sensing images could help in understanding the carbon and water dynamics within oak woodlands. Geographic Object-Based Image Analysis (GeOBIA) was optimized to classify an oak savanna in the lower Sierra Nevada foothills by testing parameter combinations of segmentation functions: `slic` segments responded the best to the spectral features of the landscape, while the addition of an edge class between trees helped with breaking up clumps of trees. The development of a crown area-based biomass equation with initial values estimated from field diameter at breast height measurements was unsuccessful due to the difficulties in isolating individual oak trees. Analyzing land cover type and evapotranspiration yielded expected relationships, and the ability of generalized additive models to predict evapotranspiration demonstrates how a GeOBIA classification can be upscaled for ecosystem modeling. The classification and measurements from a GeOBIA classification has potential to fill gaps in data, but will require a multi-step approach of image processing algorithms and initial field measurements to develop accurate relationships.

KEYWORDS

oak savanna, segmentation, machine learning random forest classifier, NAIP imagery, OpenET

INTRODUCTION

As the environment changes due to climate change, there is a growing need to measure the impacts on our ecosystems. Biomass estimations can be an indicative measure to understand the health of vegetation in ecosystems. Knowing the amounts of biomass, the total weight of plant material, and whether it is growing or shrinking could help with analyzing how trees are responding to changes in carbon, water, and nutrient cycles as a result of increases in carbon dioxide and variations in weather patterns (Jenkins et al. 2003, Garcia et al. 2017). Many nature-based climate solutions aim to better manage and restore ecosystems by increasing vegetation and carbon stock, increasing the importance of biomass estimations to verify such programs. In addition to environmental implications, the reliable estimation of biomass has various applications, such as bioenergy and estimating fuel loads for fire management (Parresol 1999, Woodall et al. 2011). Agricultural planners can use biomass data to infer crop yield from orchards (Koc-San et al. 2018, Dong et al. 2020). Biomass data is fundamental to gauge the health of an ecosystem in the face of climate change, so more products are needed, but many biomass estimations are restricted to the limitations of field-based methods.

The current methods to estimate biomass, which are essential to understanding ecosystem ecology, are mostly field-based. Biomass values are measured through the time-consuming, costly, labor-intensive, and damaging task of felling and weighing tree components, which may lead to small sample sizes, and biased and unrepresentative biomass estimations (Jenkins et al. 2003, Karlik and Chojnacky 2014). Based on the tree's measurements and mass, allometric equations can be fit. The parameters for these equations are commonly measured tree characteristics, such as diameter at breast height (DBH), diameter at root collar (DRC), height, and crown size (Parresol 1999, Karlik and Chojnacky 2014, Ozdemir et al. 2019). For example, the equations developed in the first biomass study on California blue oak by Karlik and Chojnacky (2014) use DBH and DRC as inputs to estimate components and total aboveground biomass. However, measurements taken in the field limit biomass estimates spatially and temporally. Field-based methods to obtain biomass values are difficult, leaving a gap for remote, computer-based ways to estimate biomass.

Computer vision methods on remote sensing images could measure the characteristics of trees away from the field, eliminating restrictions on biomass estimations. The purpose of computer vision and image processing in the environmental sciences is often for land cover

classification (e.g., to detect land use change). Models could identify individual trees in remote sensing images of Normalized Difference Vegetation Index (NDVI). Geospatial softwares can measure the crown sizes of detected trees and allometric equations would determine biomass. Brandt et al. (2020) outlined the tree detection step in a proof of concept paper in the West African Sahara and Sahel, and the model had high agreement between the manually outlined tree crown areas and the predicted areas. When training a model, studies consider several factors, such as whether or not the desired objects should have a shadow, minimum crown size, or the handling of clumped trees (Gonzalez et al. 2010, Brandt et al. 2020). This entire workflow has been done to estimate carbon stock by Tucker et al. (2023): the team identified and measured over 9 billion tree crown areas in the African drylands, displaying how estimation can be done away from the study site while covering substantial areas over time, which is not possible with taking field measurements.

The study by Tucker (2023) opens up new possibilities to do a similar tree detection workflow on other ecosystems. In particular, the current literature overlooks biomass estimation of oak trees in savannas and woodlands. Oak species are the most common species group in California and account for 15% of California's forest basal area (Karlik and Chojnacky 2014). Additionally, oak trees are culturally significant to Indigenous people in California, as acorns were a direct and indirect food source and the tree mass provided material for wood products (Long et al. 2016). This study will focus on blue oaks (*Quercus douglasii*), which has the largest areal extent of oak woodlands in California and is the fourth most common oak species (Karlik and Chojnacky 2014, Harold Mooney and Erika Zavaleta 2016). Historically, the trees acted as an indicator for changing seasons, whereas today, the biomass of trees can be a measure of the oak woodland's wellbeing amongst disruptions, such as fire suppression policies, urban development, and agricultural conversion (Harold Mooney and Erika Zavaleta 2016, Long et al. 2016).

In this study, I asked how well the classification of blue oak trees (*Quercus douglasii*) in remote sensing images by computer vision algorithms could be used to understand carbon and water dynamics of oak woodlands. While the study by Tucker (2023) uses U-Net Architecture, a convolutional neural network for image segmentation, there are varying levels of image processing complexity that can be tested. For my first subquestion, I optimized steps of a land cover classification workflow called Geographic Object-Based Image Analysis (GeOBIA) by determining the best functions, function parameters, and the best truth data classes for

classification of an oak savanna remote sensing image. Next, I asked how well the biomass of oak trees at my study site, Tonzi Ranch, could be estimated using an allometric equation, field DBH measurements, and crown area detected and measured from the GeOBIA classification of a USDA NAIP image. Lastly, I predicted the evapotranspiration of the study area using generalized additive models (GAMs), based on the percentage of land classified as a certain cover type. I expected to reach a satisfactory classification through an iterative process, and that there would be a relationship between crown area and biomass, and percent land cover class and evapotranspiration.

METHODS

Study site

Tonzi Ranch is a woody oak savanna in the lower foothills of the Sierra Nevada Mountains, located near the city of Ione, California (latitude: 38.4309°N; longitude: 120.9660°W). The site is classified as a Mediterranean and Csa Köppen climate (mild temperatures with dry, hot summer). The overstory is primarily made up of blue oak trees (40% of total vegetation) and sporadic grey pine trees (*Pinus sabiniana*), and the understory has annual C3 grasses and herbs (purple false brome, smooth cat's ear, and rose clover) that green during the winter rainy season from October to April (Figure 1) (Ma et al. 2016).

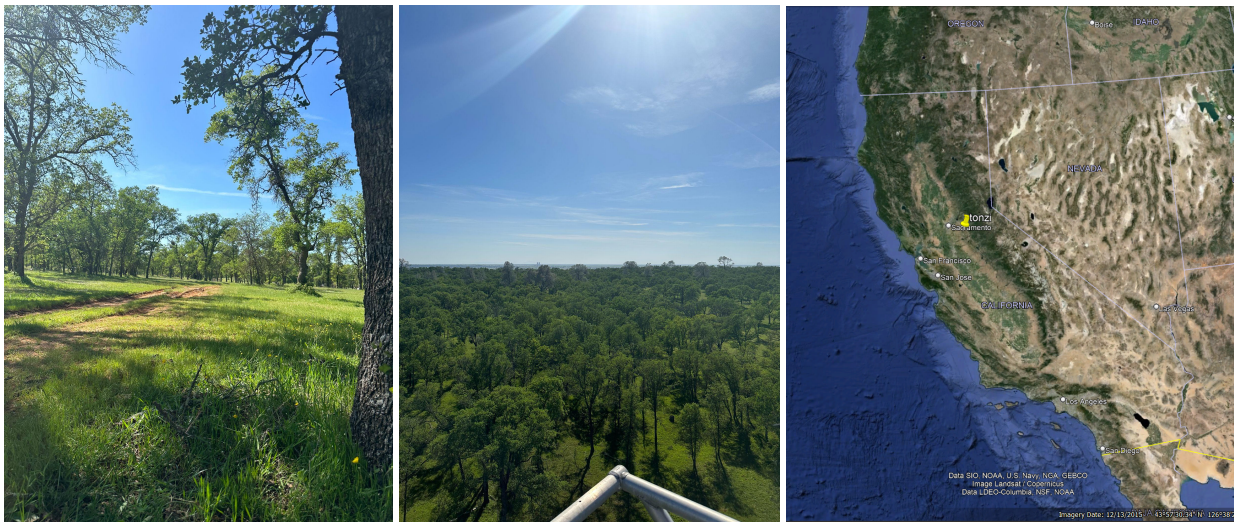

Figure 1. Photos of Tonzi Ranch and location of the study site. Oak trees along the road at Tonzi Ranch in April 2023 (left). The view of trees at Tonzi Ranch from atop the eddy covariance tower in April 2023 (middle). Tonzi Ranch relative to the state of California, viewed in Google Earth (right).

Data sources and processing

Field measurements

In order to have known biomass estimations for trees, I needed field measurements for each tree to input into allometric equations. There are tagged blue oak trees at Tonzi Ranch that have been measured in DBH surveys in 2006, 2011, 2012, and 2016. All of these trees have dendrometer bands that show incremental growth in circumference. Tonzi Ranch provided previous years' survey data: for each tree, the data includes the date, latitude and longitude, tag number, DBH, and notes.

We conducted another DBH survey in April 2023 to add another point to the data. Using a tape measure, I measured the increment first. Then, I measured the full circumference of the tree at the location of the dendrometer band, which is roughly 1.3 m (4.3 ft) above the ground. Circumference (C) was converted to diameter (d) using the following equation:

$$d = \frac{C}{\pi} \quad (1)$$

For each tree, I took photos and we recorded notes that were used to inform how to delineate tree crowns. Notes included whether or not the tree was dead and without leaves, if it had more than one trunk, and whether it was standing solo or clumped with other trees. I used a

phone camera to take a photo of the entire tree, as well as an upward facing photo of the canopy a couple feet from the trunk of each tree and at shoulder level (Figure 2).

Figure 2. Example fieldwork photos of Tree 26 taken in April 2023. Two photos were taken per tree. The first photo is of the full height of the measured tree, also showing the area surrounding the tree (left). The second photo is an upward facing photo of the canopy a couple feet from the trunk at shoulder level, in an effort to show the canopy structure (right).

Due to some irregularities in the measurements, I cleaned the data and reduced the number of trees and DBH observations available. I narrowed the DBH dataset by excluding a number of trees that had two observations per time point and trees with only one or two observations. Further, I focused on trees that were consistently growing by removing trees whose DBH shrunk or had DBH measurements that were increasing and decreasing.

NAIP imagery

To measure crown size away from the site, I needed remote sensing images, so I downloaded rasters from USDA's National Agriculture Imagery Program (NAIP) through Google

Earth Engine. The program takes aerial images in red, green, blue, and near infrared bands (R, G, B, NIR) at a high spatial resolution of 0.6 m over the continental U.S during the agricultural growing season every 1 to 2 years ("NAIP" 2018). I selected to use the NAIP image taken on 21 July 2018 for its high contrast between the overstory and understory and due to the lack of image corrections that may be needed. To test the GeOBIA methodology and develop a biomass model, I extracted a small area of Tonzi Ranch centered around where the measured trees were located, and as for the comparison with evapotranspiration, I exported a larger area that included regions with more dense canopies of blue oak trees and regions that were more sparse (Figure 3).

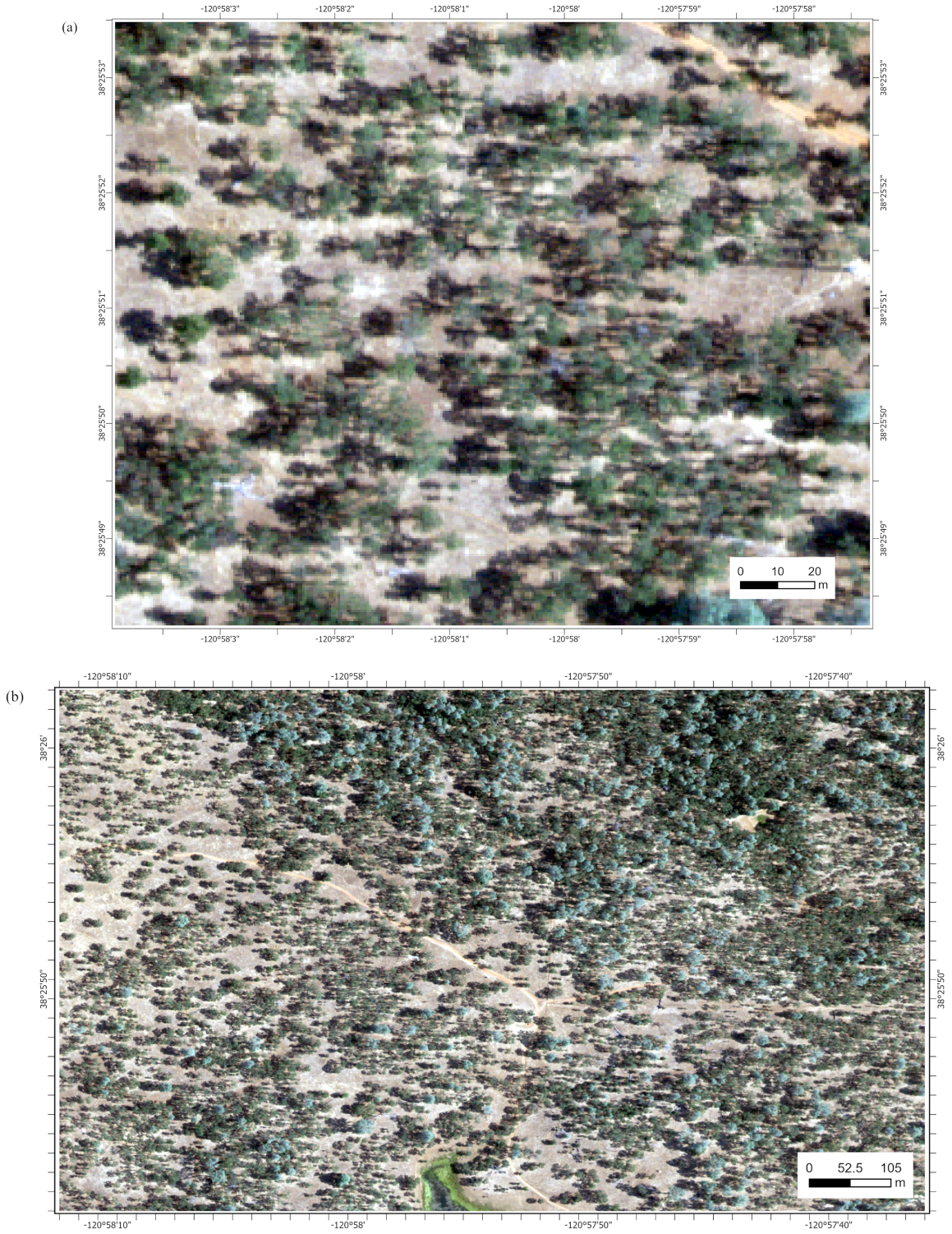

Figure 3. NAIP RGB images of Tonzi Ranch in July of 2018. (a) The small area used for GeOBIA optimization and biomass model development. (b) The large area used for predicting relationships between land cover type and evapotranspiration.

In addition to the four bands collected in the NAIP image, I calculated the Normalized Difference Vegetation Index (NDVI) using QGIS' Raster Calculator for a fifth band to inform segmentation. NDVI is known as a measure of "greenness" and is often used to quantify the relative health of vegetation where –1 indicates brown or unhealthy vegetation and +1 indicates green or healthy vegetation (GISGeography 2017). NDVI is computed from near infrared (NIR) and red bands using the following equation:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (2)$$

In the NDVI raster of Tonzi Ranch, the brighter pixels with positive NDVI are the green trees and the darker pixels with negative NDVI are areas with dried grass (Figure 4).

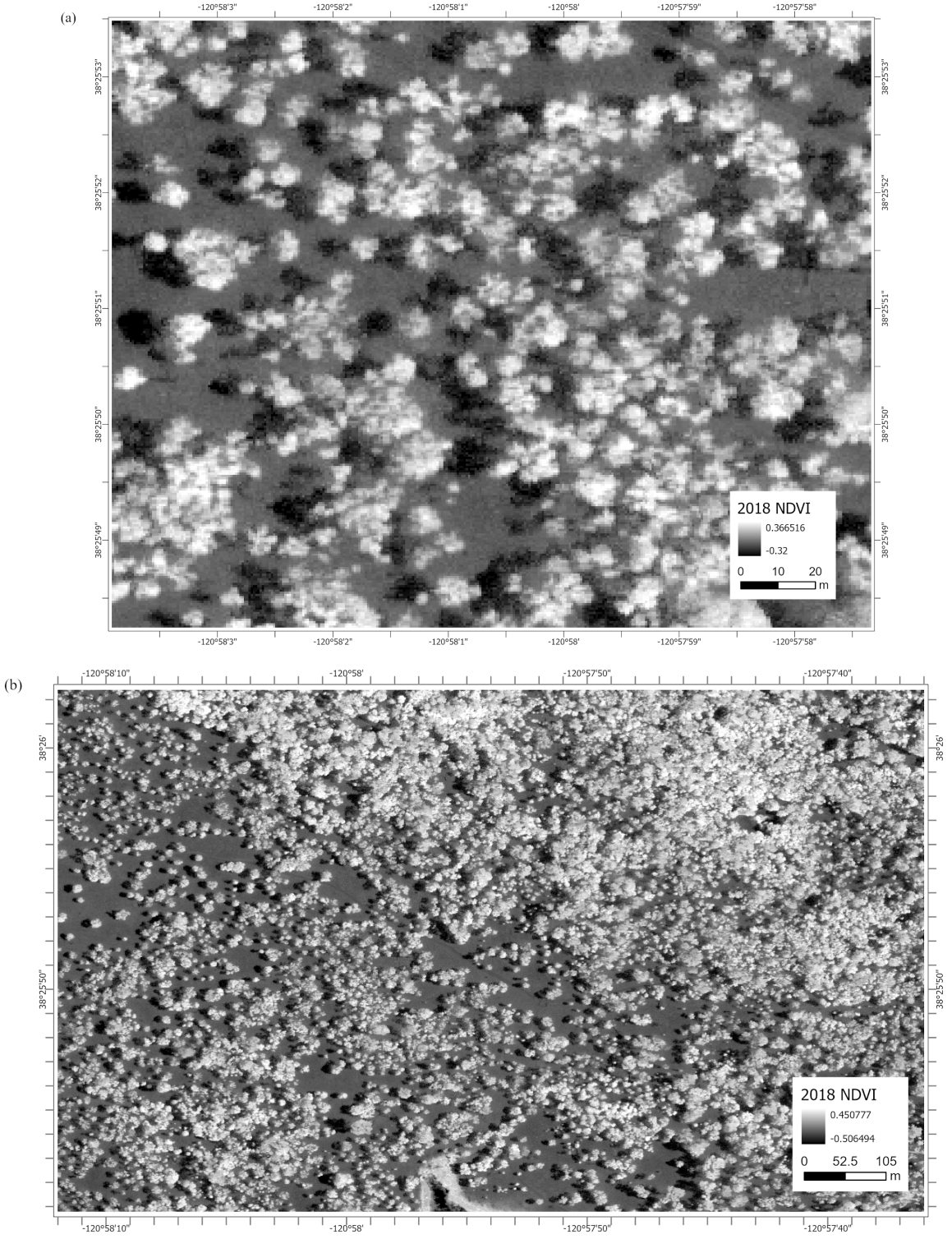

Figure 4. NDVI raster of Tonzi Ranch in July of 2018. The brighter areas with higher NDVI values are assumed to be tree crowns, while the lower NDVI values are the understory grasses. (a) The NDVI of the small area used for GeOBIA optimization and biomass model development. (b) The NDVI of the large area used for predicting relationships between land cover type and evapotranspiration.

Biomass data

For the available DBH, I calculated the above ground biomass (AGB) in kilograms using the following allometric equation that is specific to blue oak trees, developed by Karlik and Chojnacky (2014) (Equation 3, Figure 5):

(3) $$above\ ground\ biomass\ (AGB)\ = 0.0683\ DBH^{2.5697}$$

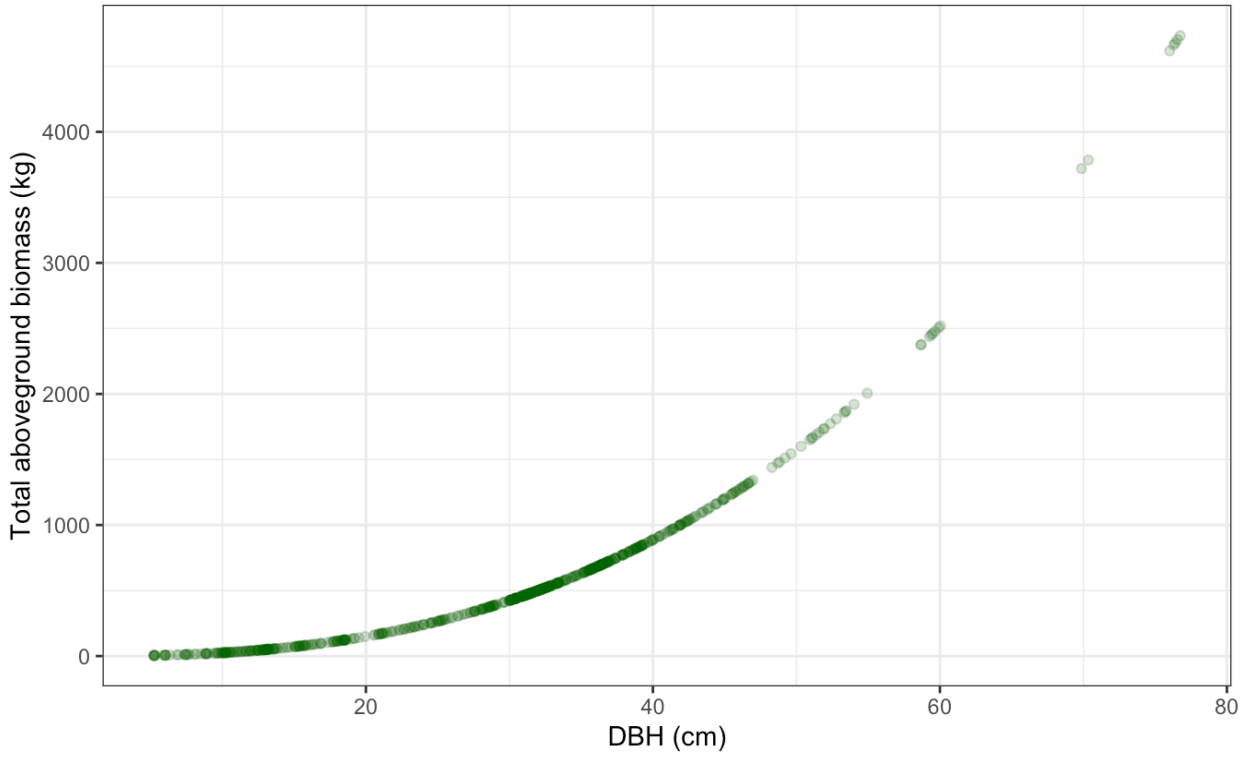

Figure 5. Aboveground biomass (AGB) of trees at Tonzi Ranch. AGB, in kilograms, was estimated using an allometric equation developed by Karlik and Chojnacky (2014) from diameter at breast height (DBH) measurements, in centimeters. AGB and DBH have an exponential relationship.

Since the dates of the DBH surveys and the dates in which NAIP imagery was collected do not match, I used the `approx` function in R to approximate the DBH and the biomass for the dates we have NAIP images. This linear interpolation step is necessary since the biomass derived from DBH values would not be representative of the estimated biomass derived from crown size measurements in the NAIP images due to the difference in years.

OpenET evapotranspiration

For the last subquestion on evapotranspiration and canopy cover patterns, I downloaded evapotranspiration data for July 2018 from OpenET in Google Earth Engine of the larger area of Tonzi Ranch (OpenET CONUS eeMETRIC Monthly Evapotranspiration v2.0) (Figure 6). The OpenET product is a monthly dataset at 30 m resolution, modeled and aggregated using a variety of sources for land surface temperature, vegetation characteristics, and soil moisture data. Evapotranspiration (ET) is the sum of all the processes that move water into the atmosphere, including evaporation from soil and transpiration from leaves.

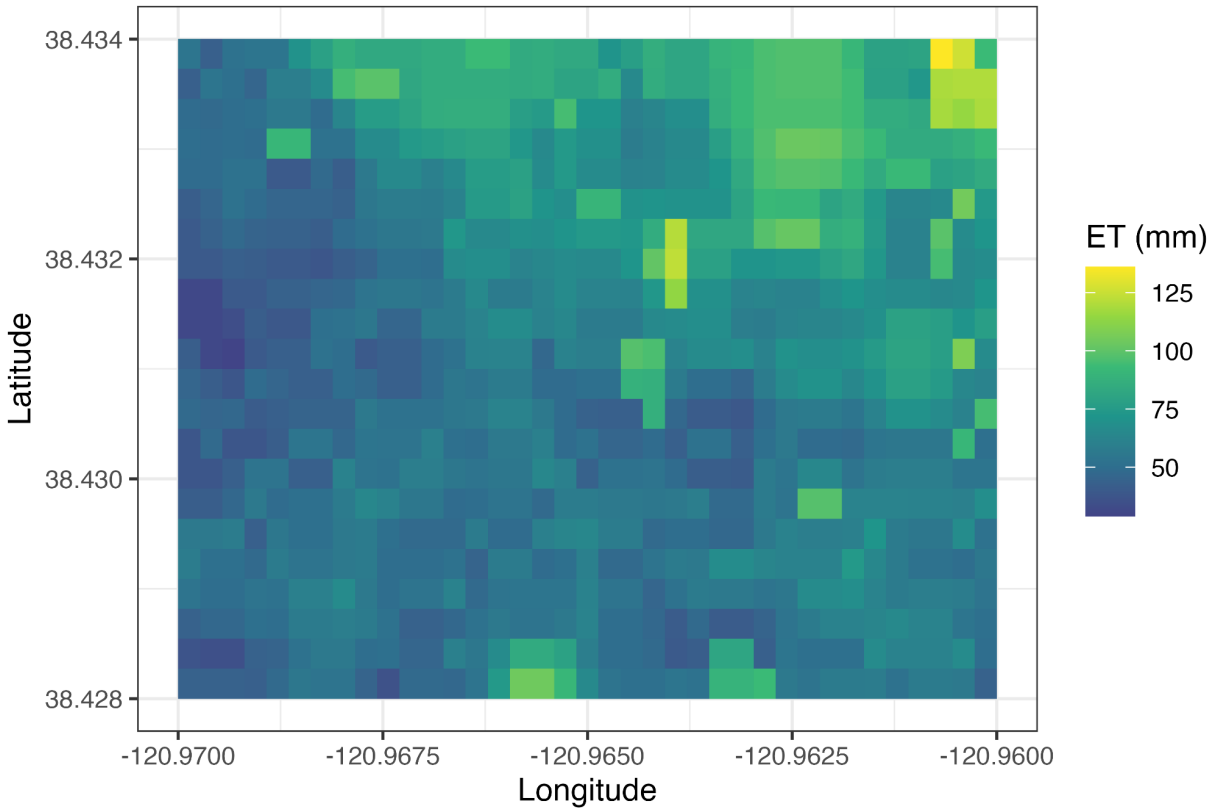

Figure 6. OpenET evapotranspiration of the larger area of Tonzi Ranch in July 2018. Higher amounts of evapotranspiration occurred in the northeast (top right) region of the expanded study area than in the southwest corner (bottom left).

Tree crown detection with GeOBIA

In order to measure tree crown measurements remotely, I delineated crowns in the NAIP image. The crowns of trees can be detected because by the time the NAIP images are taken during the summer months, the grass between the oak trees are dried up, so the understory has low NDVI values and the green leaves of the blue oak trees have higher NDVI values (see Figure 3, 4). NDVI also helps with clearly distinguishing trees apart from their shadows that are seen in the RGB image.

This is the same idea used by Brandt et al. (2020) and Tucker et al. (2023) to identify billions of trees in African drylands. These studies used manually delineated tree crowns for the training set in the deep learning network U-Net architecture. Conversely, in this study, I used Geographic Object-Based Image Analysis (GeOBIA), which is a workflow that combines computer vision and machine learning for land cover classification. There are three main parts of this methodology: image segmentation, truth data points, and classification. I followed a GeOBIA tutorial by OpenSourceOptions, which is taught using Python and QGIS, a geographic information system software (Hafen 2020a, 2020b).

Segmentation parameters

The first part in GeOBIA is the segmentation of a remote sensing image. Segmentation functions from `skimage.segmentation`, such as `slic` (simple linear iterative clustering, SLIC) and `quickshift` (quickshift mode-seeking algorithm), use algorithms to meaningfully group neighboring pixels based on similarities in the band information given. In our case, the best segmentation result would have a greater number of segments, smaller segment sizes, and more clumpiness (i.e., rather than rectangular segments). Oversegmentation is ideal to prevent a segment from having pixels that are representative of more than one land cover type, as this would confuse the statistics used for classification.

I changed the parameters of the segmentation functions to manipulate the number of segments, the size of the segments, and the clumpiness, to achieve the desired segments. The parameter in `quickshift` that was changed is `max_dist` (the maximum distance at which to stop including pixel data in a segment), while `n_segments` (the estimated number of segments wanted) and `compactness` (the weight given to balance color proximity and space proximity, with higher values making segments more rectangular) were the parameters changed for `slic`. The exact

number of segments produced and notes on the segment shapes were recorded for each segmentation parameter combination, which were used to determine the best segments.

Classes for supervision

The other major part in GeOBIA that informs classification is truth data points. In QGIS, I manually added vector point features that labeled coordinates with its land cover type. The landscape of Tonzi Ranch is moderately binary, so my points were either labeled as grass or tree. However, due to regions of the remote sensing image with a high density of blue oak trees, I included a third label and placed points in between trees, calling it the *edge* class. I hypothesized that the spectral characteristics at the edges of tree canopies may be different than at the center of the tree crown, due to factors such as leaf area index. I used visual evaluation to determine how well the inclusion of edges added to the accuracy of tree crown delineation and the disaggregation of tree clumps.

Classification and accuracy

I used Python's `geopandas` to split the truth data points randomly into training and test datasets, 70% and 30% respectively. Then, I rasterized the training vector points with `gdal` for classification where the labeled points were matched with the corresponding segment. In the last part of GeOBIA, the random forest classifier (`sklearn.ensemble.RandomForestClassifier`) was trained on the band statistics of the segments that have an associated truth label. With the understanding of the spectral distributions of grasses compared to trees, the classifier predicted and classified the remaining segments that did not have a truth label.

Pixel accuracy evaluated the classification output produced by the GeOBIA methodology. Scikit-learn's confusion matrix (`sklearn.metrics.confusion_matrix`) used the previously set aside labeled test data points and the classes of the predicted segments to assess the accuracy of the classification.

Biomass model development

One objective of this study was to create a model and data product quantifying the amount of biomass in the blue oak trees at Tonzi Ranch. I used biomass estimates known from allometric relationships with DBH to create a relationship between remote-sensing derived crown measurements and biomass that could be used for trees without DBH measurements. Since we knew the biomass for trees with DBH measurements, we hypothesized that we could fit a model between biomass and the detected crown areas of those trees. This new relationship could then be used to find the biomass of trees with only crown size data from remote sensing images and no field measurements.

I started by vectorizing the classified raster with QGIS' Vectorize (Raster to Vector) tool. Then, in the attribute table, I added a new column and calculated the area of each vector polygon using the field calculator (Geometry > $area). With the GPS locations of surveyed trees overlaid, I viewed each vector polygon that was classified as tree, and if the tree crown was near a GPS point and depending on the availability of interpolated DBH data, I added the tree's survey number and DBH to the attribute table. I plotted the DBH, crown area, and biomass points to develop crown area-biomass relationships.

Patterns in canopy cover and evapotranspiration

For the analysis on canopy cover and evapotranspiration, I performed geographic object-based image analysis on a larger USDA NAIP image. I segmented the expanded area using different parameters than was used with the smaller area, because the best segmentation function and parameters that were determined do not scale with the sizes of images, so a new model classified the image.

I created a land cover raster of Tonzi Ranch in 2018 by aggregating the number of pixels per class in the classified raster while resampling. In R, I made individual rasters for each class (grass, tree, and edge) and then upscaled the classified raster with `terra::resample`, because OpenET evapotranspiration data is at a coarser spatial resolution. During the resampling of each class' raster, I calculated summary statistics: for each larger pixel, I counted the number of smaller pixels in the classified raster and divided by the total number of smaller pixels to get the percentage of pixels classified as a certain cover type. In an effort to understand the amount of clumping that

was classified, I calculated the number of unique vector polygons present in the coarser pixel using `zonal_stats` in the `rasterstats` Python package, and the average tree polygon size.

After constructing a table with data on tree cover percentage, grass cover percentage, edge cover percentage, evapotranspiration, number of polygons, and mean polygon size for each pixel in the raster of Tonzi Ranch, I plotted the relationship between land cover percentage and ET. I developed generalized additive models (GAMs) (`mgcv::gam` function in R) using one or more variables in the data table to understand how well land cover type could predict evapotranspiration. I evaluated the GAMs by assessing the residuals, finding Root Mean Squared Error (RMSE), and conducting t-tests.

RESULTS

Tree crown detection with GeOBIA

Segmentation parameters

The first part of the GeOBIA workflow, image segmentation, is an iterative process that has importance, as it creates meaningful objects from which segment band statistics are collected to determine land cover class. To produce objects that best represent the spectral characteristics of the pixels, I tested different segmentation functions and parameters on the smaller 2018 NAIP image of the study site.

As noted in the `quickshift` documentation, I observed that as I increased the value of the `max_dist` parameter, the number of segments created decreased. The relationship between the `max_dist` parameter and the number of segments was not linear; instead, the number of segments decreased exponentially (Figure 7). Notably, there was a big difference in the output between `max_dist` of 1 and 1.05: each segment only had 1 to 2 pixels when using a `max_dist` of 1, so the shape of the segments were all rectangular and 71,926 objects were produced (Table 1). Other `max_dist` values created segment shapes that were very clumpy, but `quickshift` created large segments on the perimeter of the image that were present regardless of the parameters and the spectral characteristics (Figure 8).

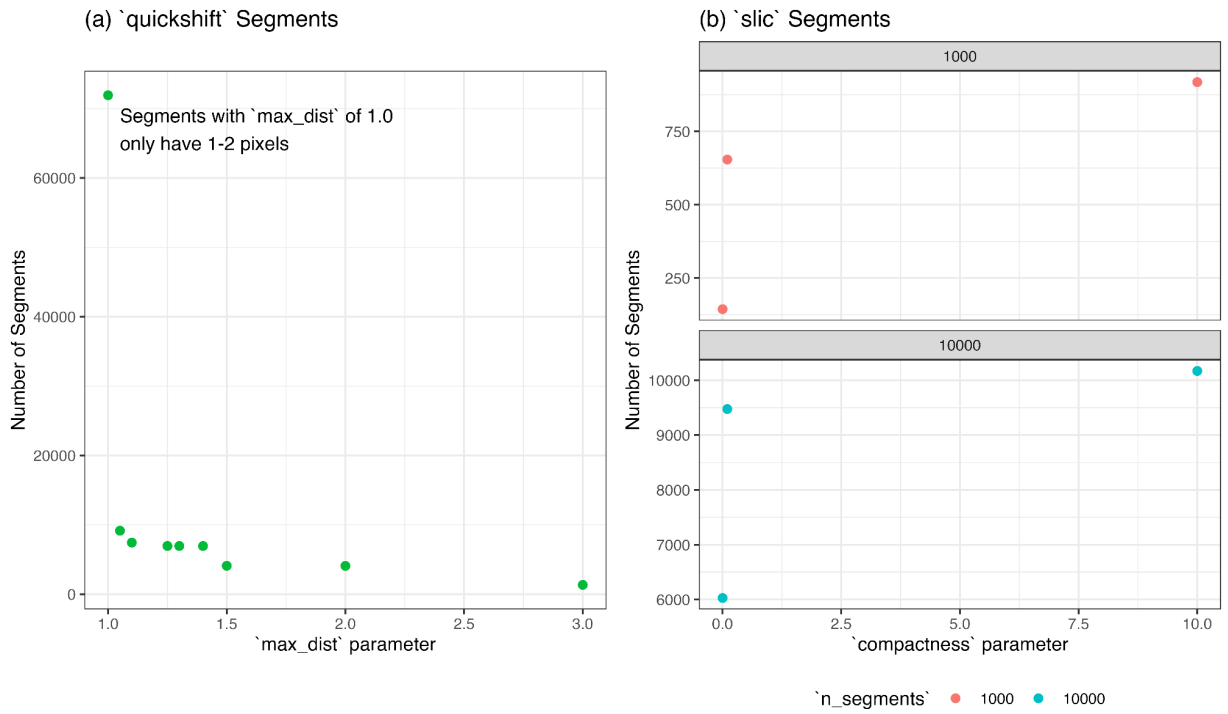

Figure 7. Number of segments created by segmentation functions. (a) `quickshift` segments decreased with increasing `max_dist` parameter. (b) `slic` segments increased with increasing `compactness` parameter.

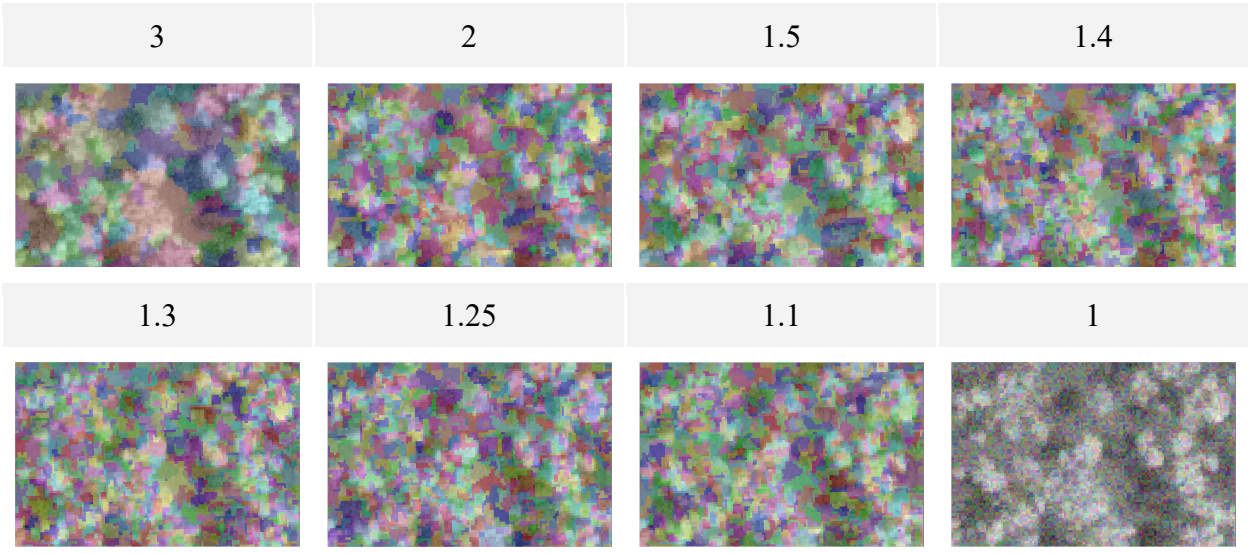

Figure 8. Segments created by the `quickshift` segmentation tests with `max_dist` parameters. Segments are overlaid on top of the NDVI image.

Table 1. Results of `quickshift` segmentation tests. The table lists the unique identifier name for the output raster of segments, the value inputted for the `max_dist` parameter, the time it took for the segmentation to complete, the number of segments created, and notes on the shape of the segments. The bolded rows were determined as the best `quickshift` segments.

Name	`max_dist`	Time (s)	# of Segments	Notes
2018_segments_test_quickshift1	3	2.36	1341	large segments, some with both classes
2018_segments_test_quickshift2	2	2.31	4081	large edge segments, very small segments
2018_segments_test_quickshift3	1	2.69	71926	all segments 1-2 pixels = too small
2018_segments_test_quickshift4	1.5	2.31	4091	large edge segments, very small segments
2018_segments_test_quickshift5	1.25	2.46	6951	very small, many segments under 10 pixels better medium edge segments
2018_segments_test_quickshift6	1.3	2.34	6947	very small, many segments under 10 pixels better medium edge segments
2018_segments_test_quickshift7	1.4	2.32	6945	very small, many segments under 10 pixels better smaller edge segments
2018_segments_test_quickshift8	**1.1**	**2.77**	**7436**	**very small, many segments under 10 pixels better smaller edge segments**
2018_segments_test_quickshift9	**1.05**	**4.33** **2.34**	**9150**	**very small, many segments under 10 pixels better smaller edge segments** **larger segments in grass, very small segments for shadows and tree crowns**

For `slic`, the number of segments created decreased with a lower `compactness` value and the same `n_segments` value (Figure 7, Table 2). A `compactness` of 10 created uniform rectangles across the entire image, a `compactness` of 0.1 segmented well around the spectral features of the trees and grass, and a `compactness` of 0.001 created longer vertically oriented segments (Figure 9).

Table 2. Results of `slic` segmentation tests. The table lists the unique identifier name for the output raster of segments, the value inputted for the `n_segments` parameter, the value inputted for the `compactness` parameter, the time it took for the segmentation to complete, the number of segments created, and notes on the shape of the segments. The bolded row was determined as the best `slic` segments.

Name	`n_segments`	`compactness`	Time (s)	# of Segments	Notes
2018_segments_test_slic1	1000	10	0.14	918	large uniform rectangles
2018_segments_test_slic2	1000	0.1	0.18	654	large clumpy segments, seems to respond to spectral features better than quickshift
2018_segments_test_slic3	1000	0.001	0.16	144	very large streaky clumps
2018_segments_test_slic4	10000	10	0.16	10170	small uniform rectangles
2018_segments_test_slic5	**10000**	**0.1**	**0.18**	**9475**	**very small segments, grass segments more rectangle, tree segments a little more clump**
2018_segments_test_slic6	10000	0.001	0.20	6028	longer vertical rectangles

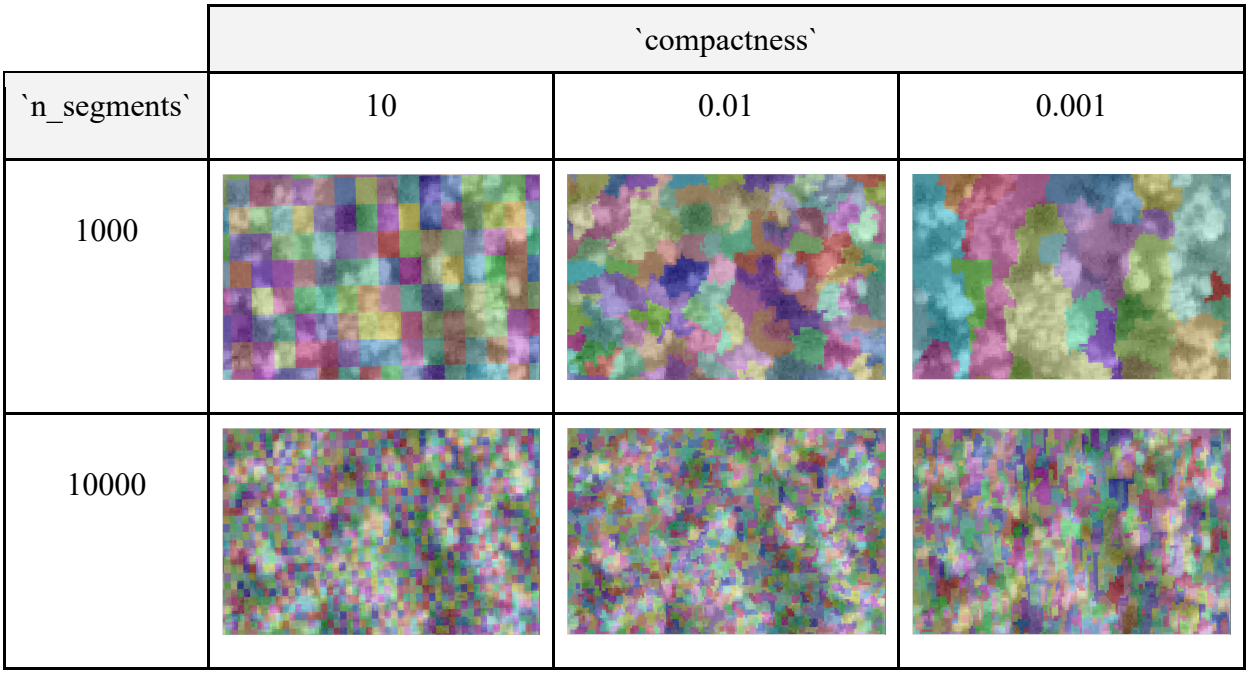

Figure 9. Segments created by the `slic` segmentation tests with `n_segments` and `compactness` parameters. Segments are overlaid on top of the NDVI image.

`Quickshift` took an average of 2.44 seconds, whereas `slic` took an average of only 0.17 seconds to complete segmenting. However, segmentations run after having already been run once using the same functions and parameters would sometimes take less time, so time may not scale with the parameters and was not a factor when deciding which segments are best to use.

Based on the number of segments created, the notes on the shapes of the segments, and viewing the output in QGIS, I decided on the following segmentation functions and parameters:

$$quickshift\_segments = quickshift(img, max\_dist = 1.1, convert2lab = False)$$

$$slic\_segments = slic(img, n\_segments = 10000, compactness = 0.1)$$

When the study site was the smaller area of Tonzi Ranch, the `slic` segments (`slic_segments`) were used for the remainder of the study, as the segments were smaller and clumpier around trees.

Classes for supervision

For the classification without the edge class, I added 123 grass points and 147 oak points (total of 270 truth data points), and for the classification with the edge class, I added 123 grass points, 146 oak points, and 53 edge points (total of 322 truth data points) (Figure 10). I aimed to add oak tree points at the center of tree crowns, placed grass points also in regions that were cast by shadows, and labeled edge points on the pixels of segments located in tree clump areas that were classified as trees in the classification with two classes.

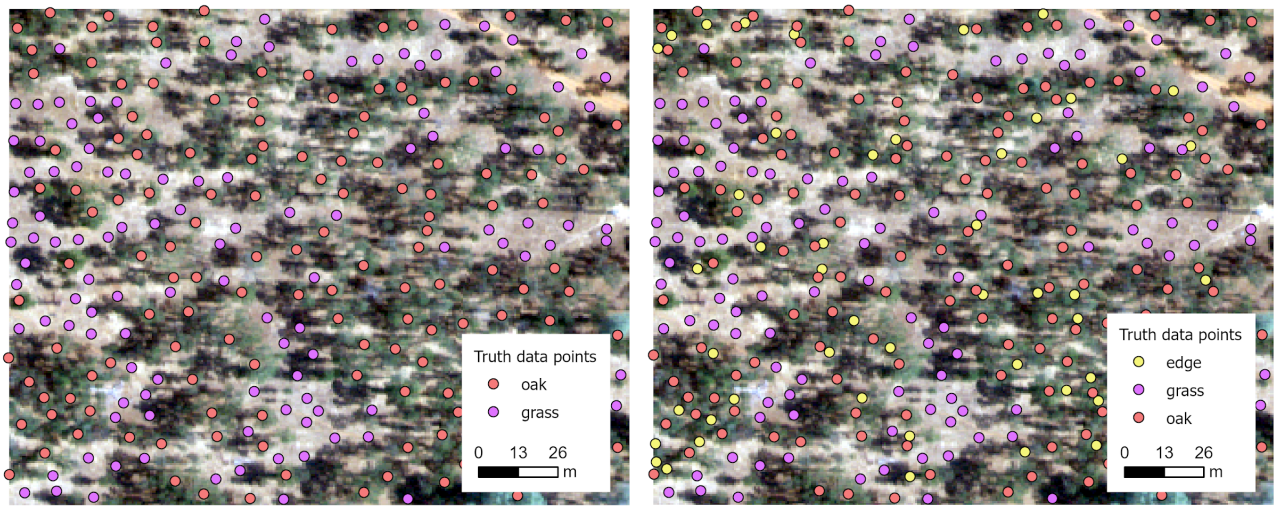

Figure 10. Truth data points over 2018 NAIP RGB image. Grass (purple) and oak (red) classes (left). Grass (purple), oak (red), and edge (yellow) classes (right).

Classification and accuracy

By appearance, the classification with the edge class was more accurate than the classification without an edge class, but both classifications had similar pixel accuracies. The amount of area that was classified as tree reduced and larger tree polygons were broken up into smaller ones (Figure 11).

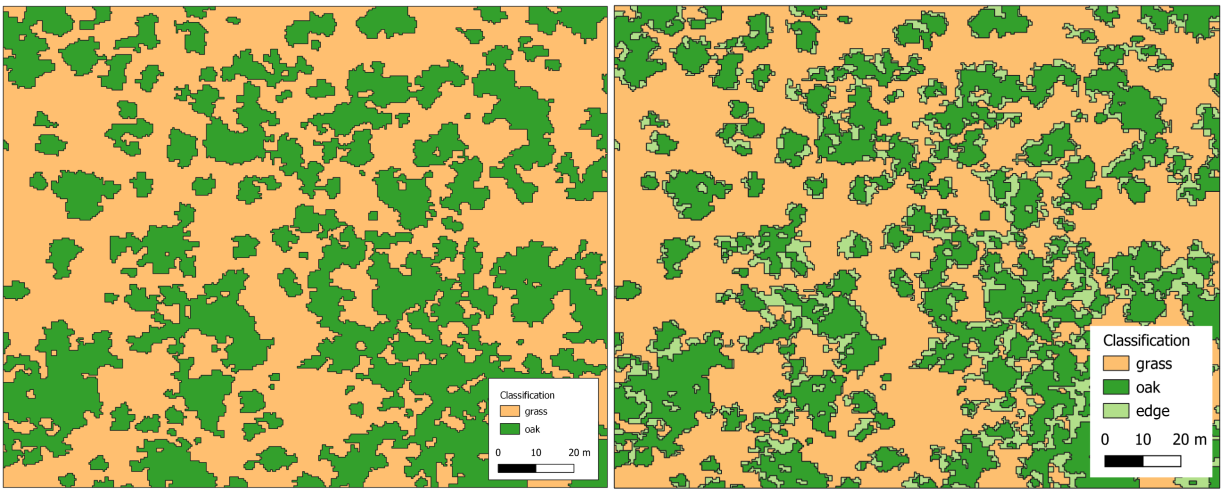

Figure 11. GeOBIA classification results of the smaller area of Tonzi Ranch in July of 2018. (a) Image was classified with two classes: grass and oak. (b) Image was classified with three classes: grass, oak, and edge.

Both classifications output by the GeOBIA methodology resulted in high pixel accuracy for all classes. For the classification with two classes (grass, oak), the random forest classifier was 97.4% accurate in predicting that a segment was grass and 100% accurate when predicting that a segment was oak tree (Table 3). For the classification with three classes (grass, oak, edge), the random forest classifier was 97.7% accurate in predicting that a segment was grass, 93.8% accurate when predicting that a segment was oak tree, and 86.4% accurate when predicting that a segment was edge (Table 3).

Table 3. GeOBIA classification pixel-based validation for the smaller area of Tonzi Ranch. Confusion matrix and accuracy for classification with two classes: grass and oak (top). Confusion matrix and accuracy for classification with three classes: grass, oak and edge (bottom).

Actual	**Predicted**	
	Grass	Oak
Grass	38	0
Oak	1	42
Accuracy (%)	97.44	100

Actual	**Predicted**		
	Grass	Oak	Edge
Grass	42	0	1
Oak	0	30	2
Edge	1	2	19
Accuracy (%)	97.67	93.75	86.36

Biomass model development

There were many sources where error was introduced, which prevented creating a biomass-crown area equation. Due to the lack of isolated tree crown polygons next to GPS locations of labeled surveyed trees and ambiguous DBH data, I was only able to select 11 trees, all contributing

varying uncertainties. I expected an increasing relationship between crown area and biomass, but plotting each tree's crown area and interpolated DBH resulted in a scatter of points (Figure 12). Without a clear relationship, I did not develop a biomass-crown area equation.

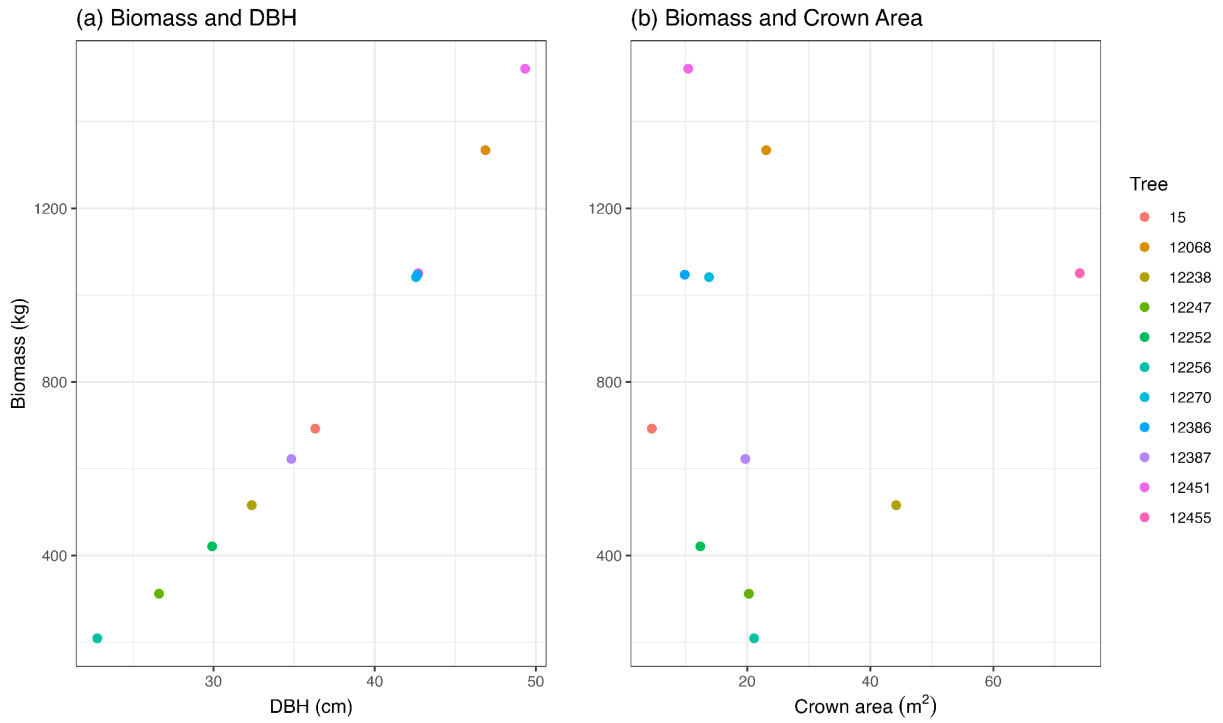

Figure 12. Biomass, diameter at breast height (DBH), and crown area for selected trees. (a) Biomass increases with DBH. (b) Crown area and biomass for the selected trees show no relationship.

Patterns in canopy cover and evapotranspiration

Since segmentation is dependent on the size of the image, I tested more `slic` parameter combinations for the GeOBIA classification of the larger area of Tonzi Ranch, and decided on the following segmentation function and parameters:

$$slic\_segments = slic(img, n\_segments = 100000, compactness = 0.1)$$

The GeOBIA classification of the larger area of Tonzi Ranch was able to classify at high accuracy, but had the same problem of large polygons for clumps of trees. The northeast region has a high density of trees that were clumped together into large tree polygons in the classification, whereas the western and southern regions of the study area have more clearly isolated trees broken up mostly by grass (Figure 13). The classification of the larger area of Tonzi Ranch had high pixel

accuracy for each of the three classes. The random forest classifier was 88.6% accurate in predicting that a segment was grass, 91.1% accurate when predicting that a segment was oak tree, and 82.4% accurate when predicting that a segment was edge (Table 4).

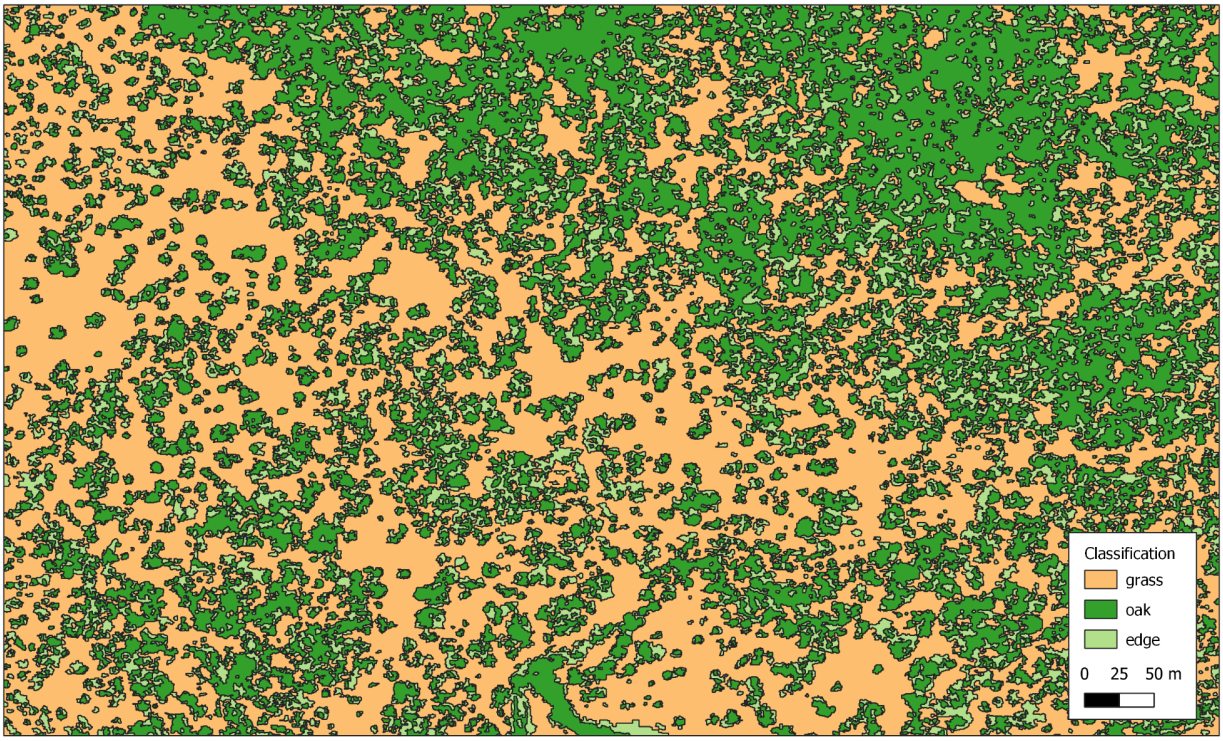

Figure 13. GeOBIA classification results of the larger area of Tonzi Ranch in July of 2018. Image was classified using three truth data classes: grass, oak, and edge.

Table 4. GeOBIA classification pixel-based validation for the larger area of Tonzi Ranch. Confusion matrix and accuracy for classification with three classes: grass, oak, and edge.

Actual	**Predicted**		
	Grass	Oak	Edge
Grass	39	0	1
Oak	2	41	2
Edge	3	4	14
Accuracy (%)	88.63	91.11	82.35

The results of upscaling each land cover type (grass, tree, and edge) and calculating the cover percentage were as anticipated with the top right half of the site having high tree percentages and the bottom left half having high grass percentages (Figure 14).

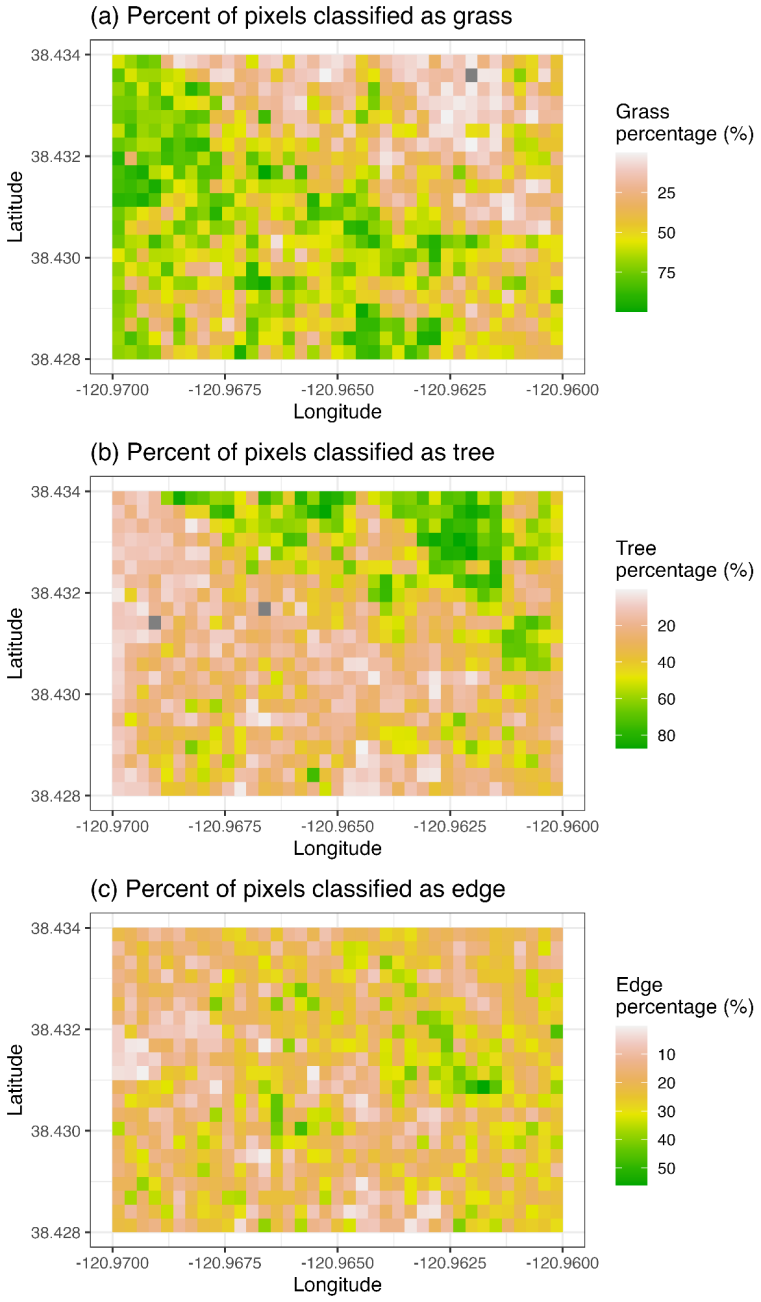

Figure 14. The percentage of each land cover type in each upscaled pixel. (a) Percent of pixels classified as grass. (b) Percent of pixels classified as tree. (c) Percent of pixels classified as edge.

Comparing land cover percentage with evapotranspiration show expected results. Evapotranspiration decreases with more grass cover, having a negative correlation, while ET increases with more tree cover, showing a positive correlation. Percent of edge cover and evapotranspiration also have a positive correlation, but the relationship is not strong with an R^2 of 0.022 (Figure 15).

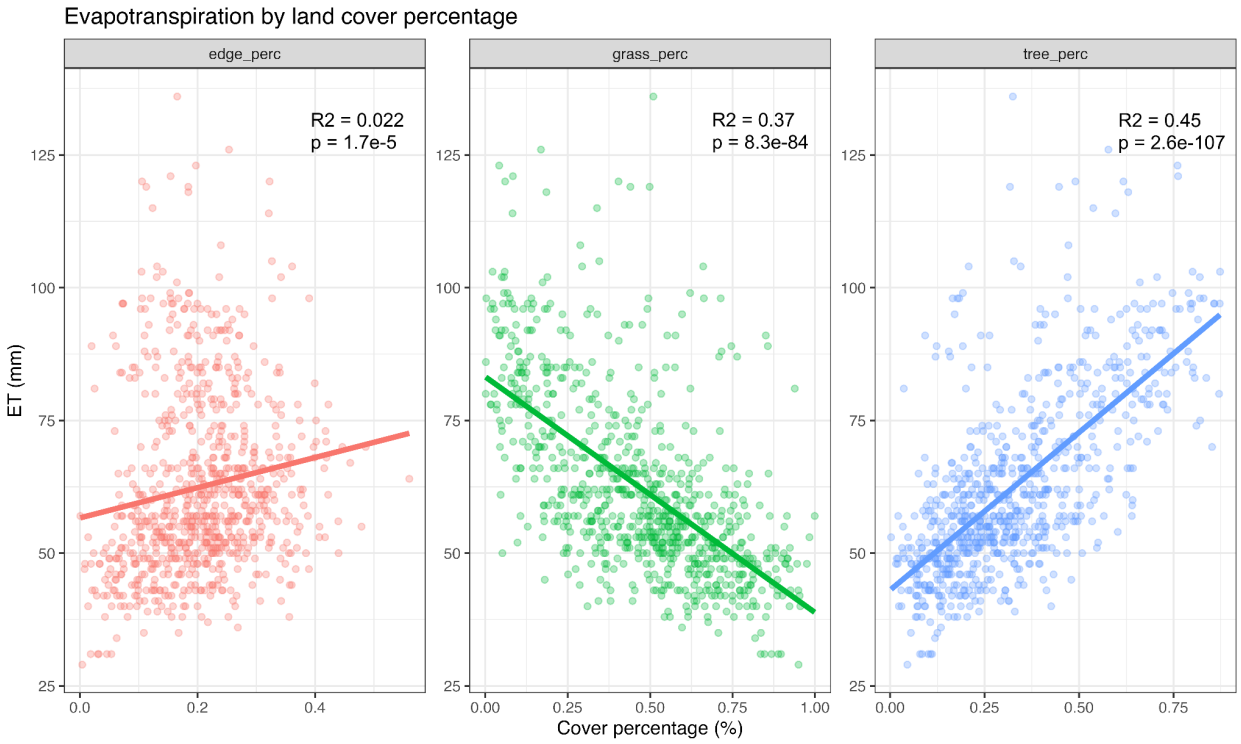

Figure 15. Relationship between evapotranspiration and cover percentage separated by land cover type. Tree and edge percent have positive relationships with evapotranspiration, and grass percent has a negative relationship with evapotranspiration.

After testing different variables in regular general additive models, I determined that land cover percentage can be used to predict evapotranspiration. The errors of the models ranged from 13 to 17 mm, and including tree percentage as a parameter improved the RMSE than when edge and grass percentage were used as the sole variable (Figure 16). The spread of the residuals is similar between models, especially for GAMs with two or more predictors and ones that include tree percentage (Figure 16). There are more outlier residuals present for the model with grass

percentage and the model with edge percentage as the parameter, but all the models have mean residuals close to 0 (Figure 16).

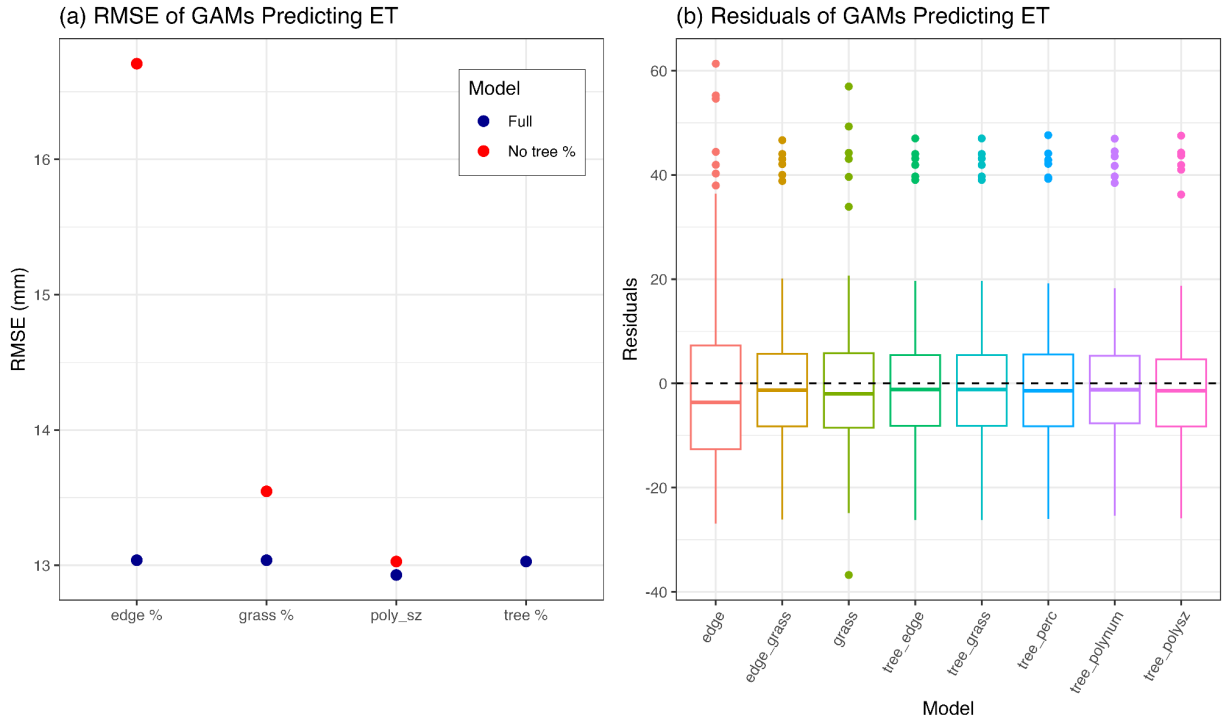

Figure 16. Results of generalized additive models (GAMs) predicting evapotranspiration (ET) using different combinations of land cover predictor variables. (a) Root Mean Squared Error (RMSE) comparing models that include and exclude tree cover percentage. (b) Residuals of GAMs all have means close to 0.

Furthermore, I focused on four generalized additive models: tree percentage, edge percentage, grass percentage, and tree and edge percentage. When conducting Student's unpaired t-tests to compare the residuals to 0, all models had a p-value greater than 0.39, meaning that there was not a significant difference between the predicted and actual values of evapotranspiration (Appendix Table A2). Residuals for the four GAMs over- and underestimate in similar places (Figure 17, Appendix Figure A1). Grass percentage has the widest residual range, and edge percentage overestimates in the top right region of the large study site where there is a high density of trees (Figure 17). Overall, the various GAMs performed similarly and in agreement between the model predictions and actual values.

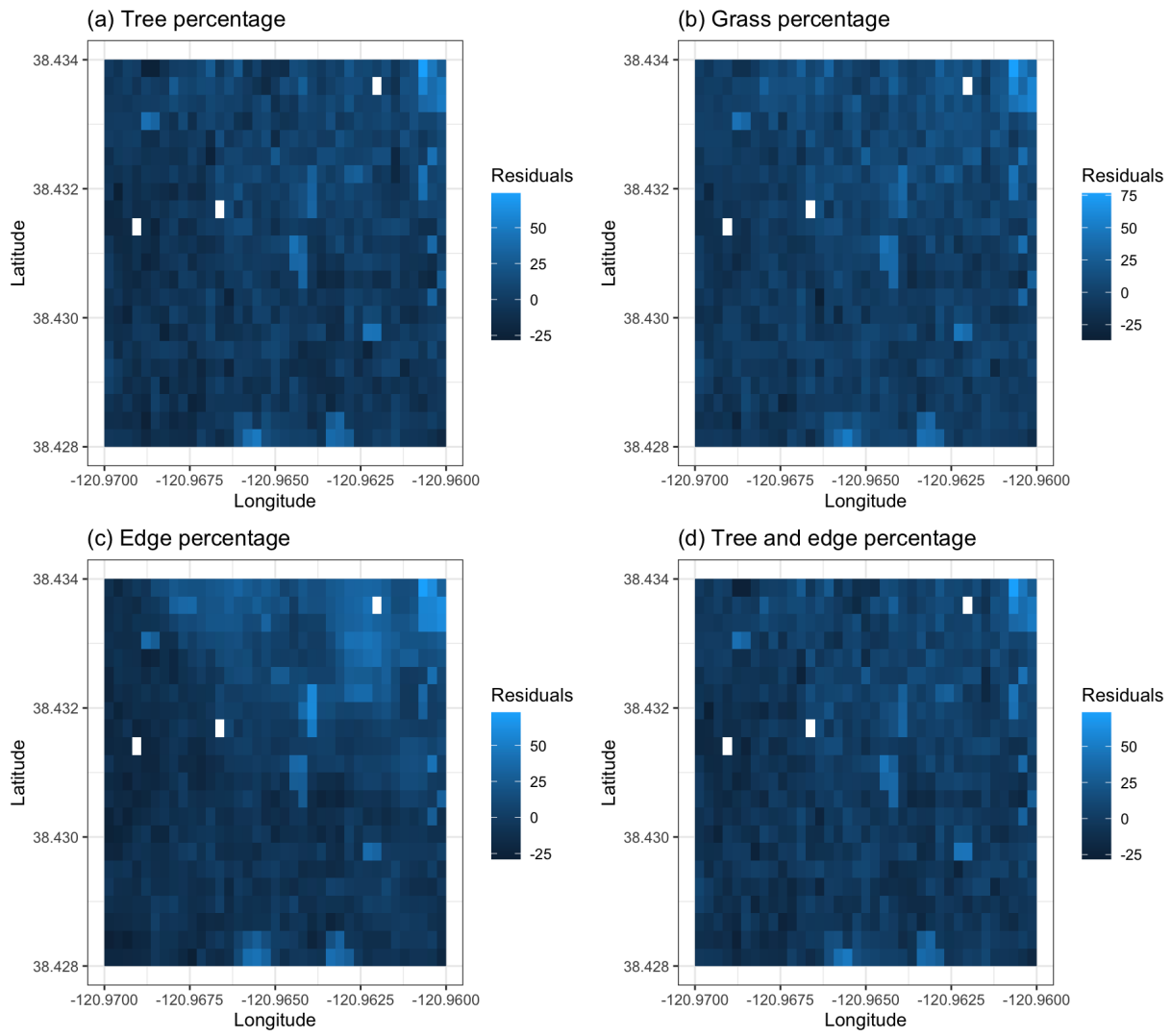

Figure 17. Residual rasters of generalized additive models (GAMs) predicting evapotranspiration. (a) GAM with tree percentage predictor variable. (b) GAM with grass percentage predictor variable. (c) GAM with edge percentage predictor variable. (d) GAM with tree and edge percentage predictor variable.

DISCUSSION

Despite the difficulties of classifying oak savanna landscapes, parts of the geographic object-based image analysis methodology can be altered to improve results. The `slic` segmentation function was more responsive to the spectral characteristics than `quickshift`. Clumps of trees can be disaggregated with the addition of a ground truth 'edge' class, but other computational functions could be used to reduce the loss of tree crown area. Ground truth data on crown area and precise tree location would increase the accuracy of remote crown area

measurements and will work towards creating a biomass-crown area equation. Land cover percentage, derived from a GeOBIA classification, as parameters in generalized additive models is a predictor of evapotranspiration. With an applicable tree crown detection methodology, there are many implications to improve the work of forest managers and agricultural planners, and verify nature-based climate solutions.

Tree crown detection with GeOBIA

Segmentation parameters

The shape of the `slic` segments were distinct depending on if the function as segmenting grassy regions or trees, suggesting that the segmentations resulting from `slic` were more responsive to the spectral characteristics in the oak savanna remote sensing image. When the `compactness` parameter was 0.1, the `slic` function segmented grassy areas into uniformly sized rectangles and the segments for areas with trees were a different clumpy shape (Figure 8). Contrastingly, decreasing the `max_dist` parameter closer to 1 in `quickshift` made the segments smaller and contain fewer pixels, but the shape of the segments and included pixels appeared random, not reflecting the difference in tree and grass in the images.

These patterns suggest that `slic` would similarly segment for other sparse and open-canopy landscapes, where some features, like tree crowns, are textured and the regions in between are spectrally homogenous. Studies that segment images into superpixels in an effort to isolate trees from the background find agreement in the resulting classification, as SLIC is able to delineate between features well depending on the number of segments and compactness parameters (Zimudzi et al. 2019, Correa Martins et al. 2021). Though, improvement is needed when the image is highly heterogeneous or strongly affected by shadows (Correa Martins et al. 2021). With careful selection of parameter values and depending on the background landscape, segmentation can improve the definition and classification of trees.

Tree clump disaggregation

I found that the addition of an edge truth data class disaggregated larger polygons of clumps of trees during classification than when an edge class was not included. Studies highlighted the space in-between trees and touching cells in microscopic imagery, during the learning process, giving them a larger weight in the model (Ronneberger et al. 2015, Brandt et al. 2020, Tucker et al. 2023). Like including an edge class, this approach emphasizes the gaps that need to be distinguished from the desired objects.

Determining the best band or combination of bands is another method to enhance differences. The principal components of a principal component analysis can set the spectral threshold for mapping tree objects apart from the understory (Kamal et al. 2015). Additionally, the ecosystem type determines the band data used to inform segmentation. A study looking at denser plots of Norway spruce and hardwood trees tested each band and determined that the green band distinctly differentiated the tree crowns from the shadows present between trees (Ke and Quackenbush 2011). Comparatively, the African drylands study found it sufficient to only use the NDVI band, as it made the tree crowns most distinct compared to the non-vegetated background (Brandt et al. 2020, Tucker et al. 2023).

Additional image segmentation algorithms on the initial GeOBIA result can isolate tree crowns. Region growing starts at a crown seed location, which can be decided on by a local maximum filter, and merges with neighboring pixels (Wulder et al. 2000, Kamal et al. 2015). Otherwise, binary morphological watershed segmentation of a distance transform was implemented to separate the rows of trees planted in urban areas, suggesting that the strategy is possibly landscape-dependent (Ardila et al. 2012). Then, morphological opening operations can refine edges, preserving both the area and shape of the trees, while the approach used in my study resulted in smaller and less circular tree polygons. Disaggregation approaches, including considering the best band data, region growing, local maximal filtering, and watershed algorithms, could have been layered to isolate oak tree crowns.

Biomass model development

While the absence of ground truth data to verify crown area estimates measured from remote sensing images prevented this study from creating crown area-derived biomass estimates and equations, past research has used crown characteristics as predictors for biomass. Results

have shown how diameter at breast height (DBH) is the strongest predictor of biomass, and the addition of crown characteristics, like width, length, and ratio, improves aboveground biomass models, but more so for tree components, suggesting that crown area and biomass may not have the strongest relationship to create a biomass-crown area relationship from (Cienciala et al. 2008). Therefore, more predictors can be used in addition to crown characteristics, until accurate remote crown area measurements can be taken. Apart from tree characteristics, an important predictor of biomass to include in a model with crown area is stand density, particularly when quantifying belowground biomass as well (Dahlhausen et al. 2017). Blue oak trees have the highest stand density among California oak trees, and it can also have implications on competition for water, the amount of water that is evaporated and transpired, and the presence of hydraulic lift among clumped oak trees (Harold Mooney and Erika Zavaleta 2016). While diameter at breast height is a strong predictor of biomass, integrating crown area and stand density can improve biomass estimation models. This approach remains crucial until accurate remote crown area measurements are feasible and can alone estimate biomass.

Patterns in canopy cover and evapotranspiration

The positive relationship between amount of tree cover and evapotranspiration, and negative relationship between amount of grass cover and evapotranspiration found is well expected. Land cover and land use significantly affect hydrological processes; in addition to evapotranspiration, there are strong effects of land cover on soil moisture and water table depth (Zhang and Schilling 2006). It is important to understand these relationships as climate change alters our ecosystems. On the leaf scale, warmer temperatures and lower humidity will increase the evaporative demand of the environment, increasing rates of transpiration and providing a cooling effect for the leaf surface (Raschke 1960, Baldocchi et al. 2021). Increased photosynthesis due to higher CO_2 concentrations could cause increased transpiration, but other studies conclude that CO_2 would induce stomatal closure to maintain a certain ratio between leaf internal and atmospheric carbon dioxide concentrations (Baldocchi et al. 2021). Comprehending the intricate interplay between land cover, hydrological processes, and climatic factors is imperative in navigating the evolving landscape of our ecosystems under climate change.

While the purpose of the edge class was to disaggregate large tree vector polygons, the edge class had a wider range of residuals and higher RMSE than the models predicting evapotranspiration with other land cover percentages, suggesting possible structural differences in the perimeter of the canopy or leaf morphology. Decreased leaf area index (LAI), amount of leaf material in a unit of the canopy, at the edges of trees, which are regions that were classified as tree when an edge class was not included, could explain why the model is overestimating evapotranspiration slightly more. Additionally, phenotypic plasticity in leaves may mean that leaves at the edges of a tree canopy could have different characteristics that allow them to persist through environmental conditions, such as increased temperatures from lower LAI and more exposure to direct sunlight. Oak trees can acclimate and grow leaves with different morphologies depending on the current needs, like varying leaf shape, size, thickness, and age (Ramírez-Valiente et al. 2010, Viscosi 2015). The discrepancies in predicting evapotranspiration with the edge class compared to other land cover percentages suggest variations in canopy structure or leaf morphology.

Conclusion

There is a lot of potential in using classifications and delineations resulting from geographic object-based image analysis. Results and literature review display how remote detection is possible, but likely requires a multi-step approach of image processing algorithms, depending on the available data and landscape characteristics. Estimating environmental variables using values measured remotely is necessary, but field measurements are initially essential in order to develop the relationships needed to fill gaps in data. Predicting evapotranspiration with land cover percentage as parameters in generalized additive models showed how the current methodology is adequate for use in modeling through the upscaling and aggregating of data from a GeOBIA classification.

Limitations

The uncertainty surrounding the validity of the classification limited the study, especially around areas with many trees clumped together, which proved to be a substantial challenge.

Although we found that adding an 'edge' class helped with disaggregating large tree polygons, it is unclear whether the removal of area classified as tree into edge area gives a correct measurement of tree crown size without the availability of ground truth crown area data, so I could not assess the accuracy of crown area values.

There was imprecision in the DBH measurements, ambiguous GPS information for individual trees at the study site, and a lack of disaggregated tree crowns. The inability to match crowns to surveyed trees caused some guesswork on my part, making it difficult to select which trees to consider and leaving an insufficient number of data points to create a crown area-biomass equation. These sources of uncertainties extend to any estimations and models developed from the classification and remotely detected measurements, but the errors could not be clearly quantified.

As for generalizability, because a blue oak-specific allometric DBH-biomass equation was used for biomass estimates, the results would not have been applicable to all oak trees or other woody plants.

Future directions

The potential for the geographic object-based image analysis workflow to detect and measure tree biomass in more ecosystems is dependent on the ability to validate the estimates. The robustness of this study would be greatly improved by incorporating ground-truth data in the form of crown area field measurements or biomass estimates derived from lidar surveys, as well as more certainty over tree locations and their respective DBHs.

Additionally, this study depended on pixel-based validation as the method to evaluate the GeOBIA classification, however object-based validation is the most appropriate form of assessment, as the segments used to collect training data in GeOBIA are meaningful objects that should be validated with objects. Object-based validation, or polygon accuracy, involves creating and rasterizing validation polygons in a geospatial software, like QGIS, and comparing the proportion of area within each polygon that was classified by GeOBIA. This assessment type allows us to generate a confusion matrix based on area of objects, instead of tally, which is more fitting for pixel-based classification (Congalton and Green 2019). Ground truth data and polygon accuracy would help us understand the applicability and usefulness of the geographic object-based image analysis methodology.

With more vigorous testing of parameters and certainty in the truth data and field measurements, this research can be expanded to all California oak savannas. Either a general biomass equation for oak trees can be used by aggregating existing allometric equations for multiple species or species-specific biomass equations after conducting species classification can estimate biomass. Additionally, the methodology can be tested on other sparse ecosystems, such as open canopy forests, woodlands, and the tundra.

Once a robust and accurate model is built, the methodology could also classify images from various years to create a time series that analyzes changes in tree crown size, canopy shape or position, and biomass and carbon stock as environmental conditions fluctuate.

Broader implications

Many tools are available that will help transition away from expensive and laborious field forest inventories and towards automatic detection of land cover types and species classification in remote sensing imagery. With sufficient spatial, spectral, and temporal resolution and some ground truth data for validation, there are numerous applications for tree crown detection. Despite the difficulties, like the complex objects in images and large variability in spectral characteristics between urban tree types, that exist when classifying trees in urban areas, the geographic object-based image analysis method expands research into measuring health associated with green spaces and allows for monitoring by forestry institutes (Ardila et al. 2012). There are environmental implications for forest and oak savanna managers who can use this research to understand the growth or decline of trees in their landscape. Automatic detection of tree counts and sizes helps agricultural planners manage orchards and estimate yields from their crops (Koc-San et al. 2018, Dong et al. 2020).

Vegetation built up by nature-based solutions to combat climate change requires verification. Environmental changes resulting from increased carbon dioxide levels, extreme temperatures, and fluctuating precipitation will cause alterations in water use by plants. Therefore, the amount of data that could be made available through remote detection warrants continued research.

ACKNOWLEDGEMENTS

Thank you to everyone who supported me throughout this entire process, especially to my mentors, Sophie Ruehr and Eric Romero, for patiently guiding me past challenges and offering recommendations during numerous meetings and through emails. I appreciate Trevor Keenan and the Quantitative Dynamics Group for welcoming me into the lab, and Joe Verfaillie for providing me with diameter at breast height data and allowing me to do field work at Tonzi Ranch. I acknowledge the Geospatial Innovation Facility (GIF) at UC Berkeley for allowing me to use the computers and softwares that were essential in my work. Thank you Konrad from Open Source Options (OSO) for their geospatial video tutorials and code. With funding awarded by the RCNR Student-Initiated SPUR program, I am grateful to have been able to connect and share my research at AGU Fall Meeting 2023 in San Francisco. Thank you to the ESPM 175A/B and ESPM 100ES instructors, including Patina Mendez, John Battles, Melissa von Mayrhauser, and Rebecca Serata, for their time and help in developing research questions over office hours. Finally, big shoutout to all my friends, family, and Environmental Science class of 2024!

REFERENCES

Ardila, J. P., W. Bijker, V. A. Tolpekin, and A. Stein. 2012. Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. International Journal of Applied Earth Observation and Geoinformation 15:57–69.

Baldocchi, D., S. Ma, and J. Verfaillie. 2021. On the inter- and intra-annual variability of ecosystem evapotranspiration and water use efficiency of an oak savanna and annual grassland subjected to booms and busts in rainfall. Global Change Biology 27:359–375.

Brandt, M., C. J. Tucker, A. Kariryaa, K. Rasmussen, C. Abel, J. Small, J. Chave, L. V. Rasmussen, P. Hiernaux, A. A. Diouf, L. Kergoat, O. Mertz, C. Igel, F. Gieseke, J. Schöning, S. Li, K. Melocik, J. Meyer, S. Sinno, E. Romero, E. Glennie, A. Montagu, M. Dendoncker, and R. Fensholt. 2020. An unexpectedly large count of trees in the West African Sahara and Sahel. Nature 587:78–82.

Cienciala, E., J. Apltauer, Z. Exnerová, and F. Tatarinov. 2008. Biomass functions applicable to oak trees grown in Central-European forestry. Journal of Forest Science 54:109–120.

Congalton, R. G., and K. Green. 2019. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Third Edition. Third edition. CRC Press, Boca Raton.

Correa Martins, J. A., G. Menezes, W. Gonçalves, D. A. Sant'Ana, L. P. Osco, V. Liesenberg, J. Li, L. Ma, P. T. Oliveira, G. Astolfi, H. Pistori, and J. M. Junior. 2021. Machine learning and SLIC for Tree Canopies segmentation in urban areas. Ecological Informatics 66:101465.

Dahlhausen, J., E. Uhl, M. Heym, P. Biber, M. Ventura, P. Panzacchi, G. Tonon, T. Horváth, and H. Pretzsch. 2017. Stand density sensitive biomass functions for young oak trees at four different European sites. Trees 31:1811–1826.

Dong, X., Z. Zhang, R. Yu, Q. Tian, and X. Zhu. 2020. Extraction of Information about Individual Trees from High-Spatial-Resolution UAV-Acquired Images of an Orchard. Remote Sensing 12:133.

Garcia, M., S. Saatchi, A. Ferraz, C. A. Silva, S. Ustin, A. Koltunov, and H. Balzter. 2017. Impact of data model and point density on aboveground forest biomass estimation from airborne LiDAR. Carbon Balance and Management 12:4.

GISGeography. 2017, May 9. What is NDVI (Normalized Difference Vegetation Index)? https://gisgeography.com/ndvi-normalized-difference-vegetation-index/.

Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561–1575.

Hafen, K. 2020a, February 29. Python: Geographic Object-Based Image Analysis (GeOBIA) – Part 1: Image Segmentation – OpenSourceOptions. https://opensourceoptions.com/python-geographic-object-based-image-analysis-geobia/.

Hafen, K. 2020b, March 25. Python: Geographic Object-Based Image Analysis (GeOBIA) – Part 2: Image Classification – OpenSourceOptions. https://opensourceoptions.com/python-geographic-object-based-image-analysis-geobia-part-2-image-classification/.

Harold Mooney and Erika Zavaleta. 2016. Ecosystems of California. University of California Press, Oakland, California.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. National-Scale Biomass Estimators for United States Tree Species. Forest Science 49:12–35.

Kamal, M., S. Phinn, and K. Johansen. 2015. Geographic object based image analysis (GEOBIA) for mangrove tree crown delineation using WorldView-2 image data.

Karlik, J. F., and D. C. Chojnacky. 2014. Biomass and carbon data from blue oaks in a California oak savanna. Biomass and Bioenergy 62:228–232.

Ke, Y., and L. Quackenbush. 2011. A comparison of three methods for automatic tree crown detection and delineation from high spatial resolution imagery. International Journal of

Remote Sensing 32:3625–3647.

Koc-San, D., S. Selim, N. Aslan, and B. T. San. 2018. Automatic citrus tree extraction from UAV images and digital surface models using circular Hough transform. Computers and Electronics in Agriculture 150:289–301.

Long, J. W., M. K. Anderson, L. Quinn-Davidson, R. W. Goode, F. K. Lake, and C. N. Skinner. 2016. Restoring California black oak ecosystems to promote tribal values and wildlife. Page PSW-GTR-252. U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, Albany, CA.

Ma, S., L. Xu, J. Verfaillie, and D. Baldocchi. 2016. AmeriFlux AmeriFlux US-Ton Tonzi Ranch.

NAIP: National Agriculture Imagery Program | Earth Engine Data Catalog. 2018. . https://developers.google.com/earth-engine/datasets/catalog/USDA_NAIP_DOQQ.

Ozdemir, E., E. Makineci, E. Yilmaz, M. Kumbasli, S. Caliskan, V. Beskardes, A. Keten, H. Zengin, and H. Yilmaz. 2019. Biomass estimation of individual trees for coppice-originated oak forests. European Journal of Forest Research 138:623–637.

Parresol, B. R. 1999. Assessing Tree and Stand Biomass: A Review with Examples and Critical Comparisons. Forest Science 45:573–593.

Ramírez-Valiente, J. A., D. Sánchez-Gómez, I. Aranda, and F. Valladares. 2010. Phenotypic plasticity and local adaptation in leaf ecophysiological traits of 13 contrasting cork oak populations under different water availabilities. Tree Physiology 30:618–627.

Raschke, K. 1960. Heat Transfer Between the Plant and the Environment. Annual Review of Plant Biology 11:111–126.

Ronneberger, O., P. Fischer, and T. Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Pages 234–241 *in* N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham.

Tucker, C., M. Brandt, P. Hiernaux, A. Kariryaa, K. Rasmussen, J. Small, C. Igel, F. Reiner, K. Melocik, J. Meyer, S. Sinno, E. Romero, E. Glennie, Y. Fitts, A. Morin, J. Pinzon, D. McClain, P. Morin, C. Porter, S. Loeffler, L. Kergoat, B.-A. Issoufou, P. Savadogo, J.-P. Wigneron, B. Poulter, P. Ciais, R. Kaufmann, R. Myneni, S. Saatchi, and R. Fensholt. 2023. Sub-continental-scale carbon stocks of individual trees in African drylands. Nature 615:80–86.

Viscosi, V. 2015. Geometric morphometrics and leaf phenotypic plasticity: assessing fluctuating asymmetry and allometry in European white oaks (Quercus). Botanical Journal of the Linnean Society 179:335–348.

Woodall, C. W., L. S. Heath, G. M. Domke, and M. C. Nichols. 2011. Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the U.S. forest inventory, 2010. Gen. Tech. Rep. NRS-88. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 30 p. 88:1–30.

Wulder, M., K. O. Niemann, and D. G. Goodenough. 2000. Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. Remote Sensing of Environment 73:103–114.

Zhang, Y.-K., and K. E. Schilling. 2006. Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: A Field observation and analysis. Journal of Hydrology 319:328–338.

Zimudzi, E., I. Sanders, N. Rollings, and C. Omlin. 2019. Segmenting mangrove ecosystems drone images using SLIC superpixels. Geocarto International 34:1648–1662.

APPENDIX

Table A1. Attribute table used to analyze the relationship between biomass and crown area. The table lists the surveyed tree number, crown area measured from the GeOBIA classification in QGIS, interpolated DBH for 2018, biomass estimated from biomass-DBH allometric equation, and notes on choosing selected trees.

Tree	Area (m^2)	DBH (cm)	Biomass (kg)	Notes
12455	74.053952	42.708	1050.437	probably 2+ trees in polygon duo, or sideways canopy
12451	10.418264	49.339	1521.734	
15	4.505195	36.312	692.256	GPS somewhat close to polygon
12270	13.797230	42.565	1041.336	
12068	23.089247	46.873	1333.857	dead as of 2023
12256	21.118227	22.783	209.045	GPS somewhat close to polygon
12247	20.273629	26.618	311.806	
12387	19.710476	34.832	622.316	
12252	12.389446	29.914	420.889	GPS somewhat close to polygon
12386	9.855246	42.657	1047.047	
12238	44.207844	32.372	515.833	

Table A2. Results of Student's unpaired t-tests for residuals from generalized additive models. The table lists the predictor variables used in the GAMs, t-values, and p-values.

Variable(s)	T-value	P-value
Tree percentage	-0.85	0.3947
Grass percentage	-0.74	0.4586
Edge percentage	-0.18	0.8595
Tree and edge percentage	-0.83	0.4075

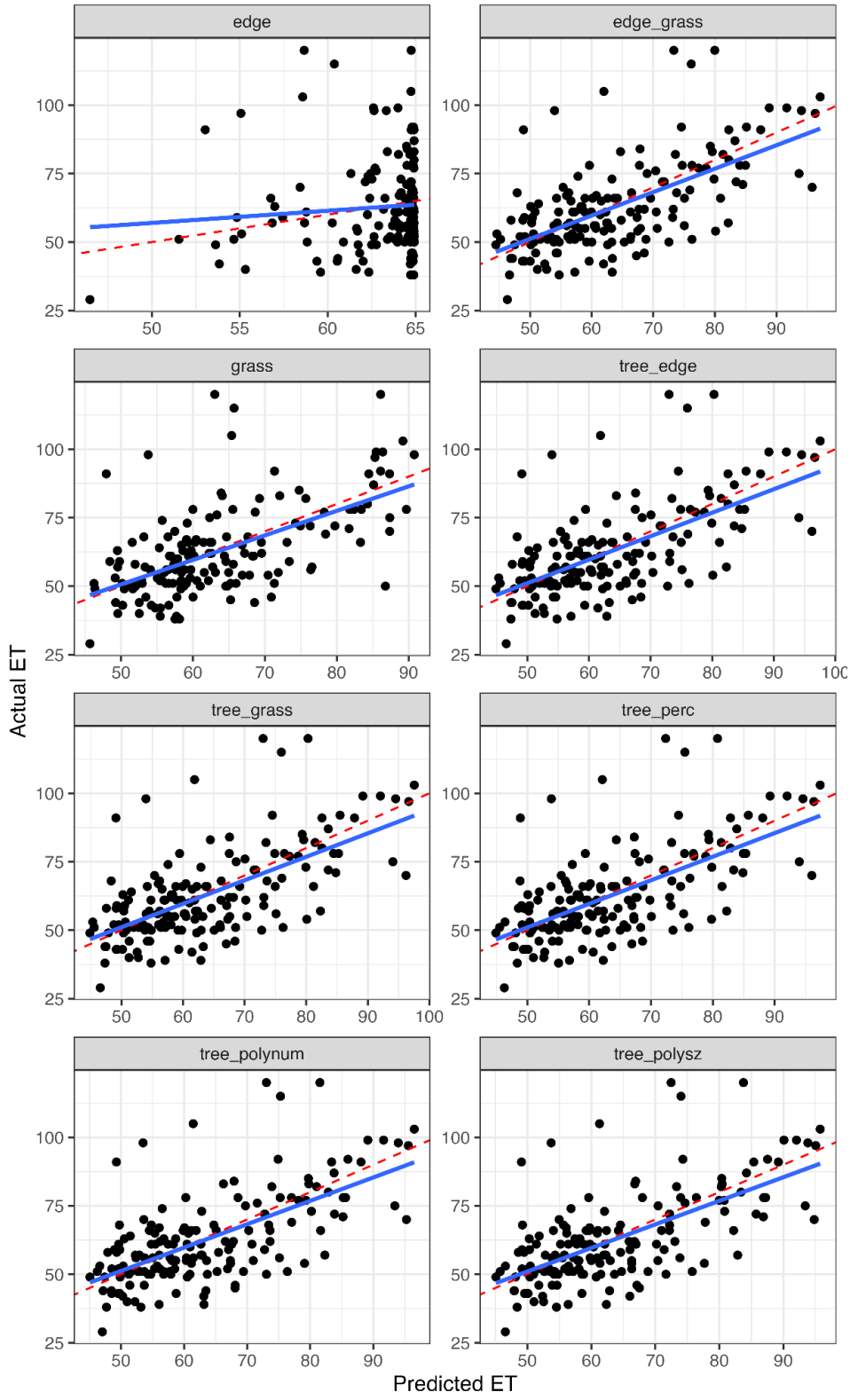

Figure A1. Evapotranspiration values predicted by GAMs compared to actual measurements. Eight different predictor variable combinations.