

UNDERSTANDING RARE SPECIES IN CALIFORNIA: AN ASSESSMENT OF
CAMATTA CANYON AMOLE (*HOOVERIA PURPUREA* VAR. *REDUCTA*)
AND A META-ANALYSIS OF CALIFORNIA RARE PLANTS
IN LITERATURE

A Thesis
presented to
the faculty of California Polytechnic State University,
San Luis Obispo

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Biology

by
Kieran N. Althaus
June 2022

© 2022
Kieran Norman Althaus
ALL RIGHTS RESERVED

COMMITTEE MEMBERSHIP

TITLE: Understanding Rare Species in California: An Assessment of Camatta Canyon Amole (*Hooveria purpurea* var. *reducta*) and a Meta-Analysis of California Rare Plants in Literature

AUTHOR: Kieran Norman Althaus

DATE SUBMITTED: June 2022

COMMITTEE CHAIR: Dr. Matt Ritter, Ph.D.
Professor of Biological Sciences
Director, Cal Poly Plant Conservatory

COMMITTEE MEMBER: Dr. Jenn Yost, Ph.D.
Associate Professor of Biological Sciences
Director, Robert F. Hoover Herbarium

COMMITTEE MEMBER: Dr. Dena Grossenbacher, Ph.D.
Assistant Professor of Biological Sciences

ABSTRACT

Understanding Rare Species in California: An Assessment of Camatta Canyon Amole (*Hooveria purpurea* var. *reducta*) and a Meta-Analysis of California Rare Plants in Literature

Kieran N. Althaus

California is currently in the midst of a biodiversity crisis. There are approximately 5,000 native species of plants in California, a quarter of which are considered rare. Determining threats to these rare plants is often times difficult. Despite California's botanical resources, we still know very little about much of California's rare plants. San Luis Obispo County is home to 2,000 of California's native plant taxa, one-third of which are rare or endemic to the county. These species are of great local and environmental concern.

In Chapter 1, we attempted to assess the impact of non native species on a threatened species in eastern San Luis Obispo County. We conducted an invasive thatch removal experiment on 10 vegetation plots of Camatta Canyon Amole, *Hooveria purpurea* var. *reducta*. The Camatta Canyon Amole (CCA) is a federally listed "threatened" plant that is only known to occur on 21.15 ha of land on Los Padres National Forest (LPNF). In the 1980s, U.S. Fish and Wildlife Service established 10 plots to monitor the population of CCA. These biologists recorded a decrease in the CCA since the establishment of those plots in the 1980s. One hypothesis for the decline is the absence of cattle grazing from LPNF, which has resulted in the accumulation of a dense thatch layer. We experimentally removed this thatch layer in five of the 1980s vegetation plots to test this hypothesis. While our the experiment was designed to be a long term treatment, from the first 1.5 years, we found no relationship between thatch removal and the amount of CCA in each plot. The effect of our treatment may take many years to materialize.

In Chapter 2, we conducted extensive botanical surveys of the Camatta Ranch, a 32,000 acre cattle ranch in eastern San Luis Obispo County. The goal of these surveys was to estimate the distribution and population size of CCA on private property, which has never before been accessed or surveyed. We did this in two ways: 1) We created a density ratio estimate based off of plot sampling done on the ranch and 2) we created a species distribution model (SDM) to predict the likelihood of presence throughout the ranch. Our surveys of Camatta Ranch, coupled with our SDM suggest that a majority of CCA's preferred habitat is on Camatta Ranch, making the ranch of paramount concern for CCA's protection. Our estimates suggest that 90% of the total population of CCA occurs on Camatta Ranch.

In Chapter 3, we attempted to quantify biases in the literature about California's flora. The California Floristic Province is one of the most biologically diverse floras in the world. Considerable legal and conservation attention is given to rare plants in California. However, there is no information as to the research effort given to rare species in California. Here we ask the question: Is there more research done on rare plants in California than on non-rare species? To answer this question, we quantified the amount of literature available on Google Scholar for California's rare plants, weeds, and non-rare

natives. To account for the differences in species geographic extent, we aggregated occurrence data for each species from GBIF to determine their ranges. We found that rare species were severely under-represented in the literature, even after accounting for the differences in species extent.

Keywords: *Hooveria purpurea* var. *reducta*, Camatta Ranch, thatch removal, floristics, conservation

ACKNOWLEDGMENTS

Thank you to the Morrison Family for allowing me the pleasure to study on Camatta Ranch. It was an absolute joy to spend many weekends out exploring and surveying your ranch for the Camatta Canyon Amole. Thank you to California Native Plant Society San Luis Obispo Chapter, David and Freida Wertman scholarship fund and the US Fish & Wildlife Service for supporting this research.

This research would not have been possible without all those who helped me in the field. Thank you to the Field Botany Class of 2021 for, without complaint, pretended to be cows out at Red Hill Road on a hot spring afternoon. Thank you to Amanda Gersoff, Stephanie Messina, Charollette Miranda, Sophia Jones, Erin Grady, Miguel DiSanzo-Graham, Caitlin Robertson and Maya Avendano for assisting me with field work on Camatta Ranch and Red Hill Road. It's not easy work counting a plant 3 inches tall, but you all did it with enthusiasm. Thank you to both Dave Clandenan and Craig Jacobson for their assistance in securing the supplies I needed to do my field work successfully.

My time at Cal Poly would not be complete without the other students in my life. Thank you to my lab mates, friends and fellow graduate students Cami Pawlak, Annie Meeder, Annie Zell, Annie Ayers, Stephanie Calloway, Emma Fryer, Paul Excoffier and Gage Willey; you've all inspired me and helped me along this journey more than you know. Thank you to Reed Kenny and Thomas Buchloh for constantly challenging me and being willing to just nerd out on plants. Thomas, thank you for keeping me sane during the pandemic and my first year as a graduate student. To all the biology graduate students: your energy, curiosity and intelligence has made the last two years an absolute pleasure. Thank you.

The third chapter of this thesis would not have been possible without my dear friend, Kyle Nessen. Kyle, thank you for teaching me that with enough drive and a bit of code, you can answer the most interesting of questions. I look forward to many more interesting projects with you in the future. To Natalie Love, for being my early morning breakfast companion in 318, and for always taking the time to answer my questions. Your intellect and dedication to science has inspired me to no end.

To my wonderful advisors and friends, Matt Ritter and Jenn Yost, thank you for opening the doors of botany to me. It is no understatement to say that if it weren't for you two, I would not be the person I am today. You have both instilled a sense of curiosity and open-mindedness in me, and I think it is safe to say you've cured me of plant blindness. Matt, you've not only taught me so much about being a good botanist, but also about how to live a fulfilling life. Jenn you were always there when I felt lost in my research. I hope to emulate your passion for teaching and your positive approach to life in my future. And thank you to my committee member, Dena Grossenbacher. Your contribution to my work and development as a field botanist and academic has been immensely valuable. You taught me so much about the nitty gritty of science. I have a feeling I'll be thanking you three for the rest of my academic career.

Thank you to my family and my parents, Lesley and Otto Althaus for supporting me, especially in my dramatic switch from Political Science to Botany. Finally, thank you to all those I couldn't thank on this page. I am grateful to have had your love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES IX

LIST OF FIGURES X

CHAPTER

1. THATCH ACCUMULATION AND ITS EFFECTS ON THE THREATENED CAMATTA CANYON AMOLE (*HOOVERIA PURPUREA* VAR. *REDUCTA*)..... 1

 1.1. INTRODUCTION 1

 1.2. MATERIALS & METHODS 3

 1.2.1. Red Hill Road 3

 1.2.2. *Hooveria purpurea* var. *reducta*..... 3

 1.2.3. Study Plots 5

 1.2.4. Treatment Groups 5

 1.2.5. Treatments..... 7

 1.2.6. Statistical Analysis..... 7

 1.2.7. Difference in Biomass..... 8

 1.2.8. Differences in CCA in Plots 8

 1.3. RESULTS 9

 1.3.1. Biomass..... 9

 1.3.2. Emergent Amole 10

 1.4. DISCUSSION..... 11

2. UNDERSTANDING THE GEOGRAPHIC DISTRIBUTION AND POPULATION SIZE OF CAMATTA CANYON AMOLE (*HOOVERIA PURPUREA* VAR. *REDUCTA*)..... 14

 2.1. INTRODUCTION 14

 2.2. METHODS & MATERIALS 16

 2.2.1. Study Site 16

 2.2.2. Experimental Design..... 16

 2.2.2.1. Spring 2020 Survey..... 16

 2.2.2.2. Spring 2021 Survey..... 17

 2.2.2.3. Creating Survey Polygons..... 17

 2.2.3. Species Niche Model Variables 19

 2.2.4. Ecological Niche-Based Distribution Modeling..... 20

 2.2.5. Comparing SDM to Known Occurrences of CCA 21

 2.3. RESULTS 21

 2.3.1. Ranch Estimates..... 21

 2.3.2. Species Niche Model Results..... 23

 2.4. DISCUSSION..... 24

 2.4.1. Survey Effort..... 24

 2.4.2. Niche Model on Camatta Ranch 25

3. QUANTIFYING RESEARCH EFFORT IN CALIFORNIA’S FLORA..... 28

 3.1. INTRODUCTION 28

 3.2. MATERIALS AND METHODS..... 30

 3.2.1. Scraping 30

 3.2.2. Search List 30

3.2.3. Plant Area.....	31
3.2.4. Statistical Analysis.....	31
3.3. RESULTS	32
3.3.1. Rare, Native, Weed Publishing Bias.....	32
3.3.2. Rare Plant Rank Publishing Bias	33
3.4. DISCUSSION.....	33
BIBLIOGRAPHY	36
APPENDIX	46

LIST OF TABLES

Table	Page
1. Center coordinates and descriptions of 10 circular plots (1 m ²) along Red Hill Road in Los Padres National Forest, San Luis Obispo County. Plots are the same as those used in surveys from USFWS between 2014-2020. Table from Kofron et al. 2021	6
2. The 15 ranch polygons that were surveyed in Spring 2021 along with plot area, CCA counts and relative density ratios. Polygons are arranged in order of smallest to largest density ratios.	53
3. The number of plants in our dataset with various CRPR rankings.....	64

LIST OF FIGURES

Figure	Page
1. Map of the three known populations of <i>Hooveria purpurea</i> . Both Camp Roberts and Fort Hunter Liggett are home to <i>Hooveria purpurea</i> var. <i>purpurea</i> . Red Hill Road is the only location known to have <i>Hooveria purpurea</i> var. <i>reducta</i> , the Camatta Canyon Amole.	4
2. Diagram visualizing the layout of a treatment plot. Dotted line represents a bearing laid out from a marker (usually a tree) to plot center. Bearings with associated markers described in Table 1. The three circular plots are where all CCA counting, percent cover and height measurements took place. The greater rectangle around the three 1 m ² circles is where thatch was collected, including within the three circular plots.	6
3. Biomass of thatch collected across four visits, grouped into two rain years. A majority of the thatch collection was done in November 2020 and February 2021 (first rain year). Thatch did not recover between rain year 1 and 2. Very little thatch was collected in visit 4 (not even visible in the above graph).	10
4. Plot 6 on Camatta Ranch. Shows process for assigning compass bearings and counting CCA along those belt transects. In this example, a total of 41 CCA was counted along 135 m ² . This averages out to ~ 0.3 CCA/ m ² in this plot.	19
5. CCA in each plot over eight years of monitoring. The vertical red line indicates the year thatch removal began. Paired control and treatment plots are organized next to each other and are in the same rectangular border.	46
6. Location of <i>Hooveria purpurea</i> var. <i>reducta</i> as mapped in CNDDDB. The long, skinny polygon is along Red Hill Road. The extent of the polygon that extends onto Camatta Ranch is speculative. Large circular polygons are buffered points.	47
7. Thatch height and cover across three years. While thatch cover and height were evenly variable in the first year between plot types, they had segregated by 2022, where mean height is similar, but the mean cover between groups is different.	48
8. Thatch cover across rain years. Thatch cover decreased in both control and treatment plots between the two rain years.	49
9. Total number of Camatta Canyon Amole counted in all 10 vegetation plots and rainfall (mm) for that rain year (blue line). For this data a rain year begins on November 1 and ends March 1. For example, between November 1, 2013-March 1, 2014, the weather station measured 188.97 mm of rain. Rainfall data is that which we believe would contribute to that growing season. Data from SLO-6 weather station on Shell Creek Road, 8 miles NW of of Red Hill Road.	50
10. Percent cover and Emergent Amole of each plot at the final visit in February 2022 for the ten vegetation plots. Plot 7 had the highest counts for CCA while also having the highest percent cover. Plot 5 had equal cover to plot 8, but with more than twice the number of emergent CCA.	51
11. Total Camatta Canyon Amole from all survey efforts along Red Hill Road. "Historic Red Hill Road" are plots surveyed by David Magney in 1987 and 1988. "Modern Red Hill Road" plots are those reestablished by Kofron + Rutherford in 2014 in the approximate location of Historic Plots. Kofron et al. 2021 treats these	

two sets of plots as the same plots. Because the modern plots are in the approximate location of the historic plots, we treat them as separate. “Random” plots reflects a survey effort of 11 randomly selected plots along Red Hill Road. This happened just once. “Transects” reflect 11 belt transects that occurred over 6 years around Red Hill Road. These transects occurred along areas identified as the densest. All data is recovered from Kofron et al. 2021, Table 3. 52

12. Map of Camatta Ranch. Located directly north of the Carissa Highway (HWY 58) near where Red Hill Road and Carissa Highway intersect. *Hooveria purpurea* var. *reducta* is only known to occur in the very southern portion of the ranch, close to Carissa Highway and Red Hill Road. 53

13. Results of the different models. The upper graph are the average validation AUC and below is the omission rates for the models. The x-axis are the four regularization multipliers, and the color points are the four feature classes. Feature class H and LQH had tied omission rates at rm 7, so our sequential test broke the tie for determining the optimal model by choosing from the highest AUC at that rm, which was fc H. 55

14. Response curves for optimal model. Aspect and slope had low predicting power for CCA, whereas TRI, elevation and NDVI were more strongly predictive of CCA presence. 56

15. Results from 100 null models. Shoes null model AUC and 10% omission rate results, along with the AUC and OR10 of our empirical model. 57

16. Raw prediction of the optimal model on Camatta Ranch and Red Hill Road. 58

17. Species distribution model for CCA in Occurrence 1 (CNDDDB) and along the southern ridges of Camatta Ranch. Because our prediction generally predicts Occurrence 1 along Red Hill Road with a value of 0.5 or higher, we limited the theoretical distribution of the CCA on Camatta Ranch to areas with a value of 0.5 or greater for our population estimates. 59

18. Histogram of literature counts of native plants of California. X-axis is limited to 3000 pieces of literature. The native species mentioned most in literature is *Helianthus annuus*, with 186,293 pieces of literature. The vertical red line separates half of the species in California. 60

19. Boxplot of the literature each species status group had. Median for Non-Rare Natives is 75. Median for Rare Natives is 26. Median for Weeds is 7770. The median number of literature across the three groups is 55. 61

20. Histogram of genbank sequences from the GenBank scrape. The vertical red line represents the median number of sequences: 8 sequences. 62

CHAPTER 1

Thatch Accumulation and its Effects on the Threatened Camatta Canyon Amole

(*Hooveria purpurea* var. *reducta*)

1.1. INTRODUCTION

Invasive species have been shown to impact natural and urban ecosystems alike (Pysek et al. 2012; Blossey 1999). These impacts vary from species to species. Invasive species have been shown to increase soil salinity (MacDonald et al 1988), change fire frequency (Knapp 1992; MacDonald et al 1988), prevent the emergence and recruitment of native plants (Kettenring and Adams 2011), and alter soil chemistry (Facelli and Pickett 1991; Strayer et al. 2006; Strayer 2012; Vitousek and Walker 1989; Ehrenfeld 2003; Levine et al. 2003) Coupled with climate change, which is expected to further exacerbate native/invasive interactions, invasive species are believed to be a driver of rising extinction rates (Wiens 2016; Duenas et al. 2021; Gurevitch and Padilla 2004). Invasive species are, by definition, accomplished at inhabiting novel environments. The shift of species ranges expected with climate change is expected to further the spread already invasive species (Corlett and Westcott 2013; Wiens 2016). While invasive plant species occupy every continent, some places are harder hit than others.

Regions with mediterranean climates have shown significant decline in species richness as a result of invasive species (Pysek et al. 2012; Gaertner et al. 2009).

Mediterranean climates are characterized by cool, wet winters and warm, dry summers (Lionello et al. 2006). There are five mediterranean-climate regions, which together harbor 20% of the worlds plant diversity, while only covering 5% of the total surface area (Cowling et al. 1996). The California Floristic Province (CA-FP), which

encompasses cis-montane California, southern Oregon and northern Baja California, is one of the five regions with a mediterranean climate (Raven and Axelrod 1978). Out of all the mediterranean climates, the occupation of California by invasive remains one of the most substantial, with invasive plant species making up 20% of the state's flora (Bossard and Randall 2007). While all ecosystems in California are comprised of some invasive species, they have particularly come to dominate the landscape of Californias grasslands.

It is believed that 13% of California was once covered by perennial grasslands, which have since shifted to non-native annual grasslands (Keil and Holland 1996; Bossard and Randall 2007). These grasslands were damaged as a result of early cattle grazing and human settlement (Seabloom et al. 2003; Huntsinger et al. 1997). Even after many years post grazing, many grasslands remain dominated by invasive annual grasses (Hayes & Holl, 2003). These invasive species outcompete native competitors and increase the intensity and return-interval of fires, eliminating woody shrub cover and perennial forbs over time (McKenzie and Littell 2011; Keeley 2002; Brooks *et al.* 2004).

Efforts that seek to mitigate the degradation of native grasslands are common in California. The removal of dead leaf litter or the use of herbicide are the two most frequent approaches for grassland restoration (Kettenring and Adams 2011). The affect of litter is variable, but it is commonly associated with altering water, nitrogen, and sunlight availability (Wolkovich, Bolger and Cottingham 2009; Vitousek and Walker 1989, Blossey 1999).

In this study, we test whether thatch accumulation affects the emergence of a perennial geophyte, *Hooveria purpurea* var. *reducta*, the Camatta Canyon Amole. To do this, we revisited 10 vegetation plots established in 1987 (Kofron *et al.* 2021) and removed non-native grass and its thatch. We hope to address the following question: Has the accumulation of thatch between 1988 and 2020 responsible for the decrease of Camatta Canyon Amole.

1.2. MATERIALS & METHODS

1.2.1. Red Hill Road

Field work was carried out November 2020- June 2022 along Red Hill Road in the Los Padres National Forest, in San Luis Obispo County (35.402750, -120.279944). Red Hill Road is a graded dirt road that intersects State Highway 58, and is often used as a launching point for off-road vehicles (Kofron *et al.* 2021). The soil is hard and rocky, and has been mapped as Arbuckle Sandy loam, a soil type that Camatta Canyon Amole associates with. Red Hill Road rests on Red Hill Mesa, which is a flat ridge that is dominated by Blue Oak Savanna (*Quercus douglasii*) and an understory of invasive annual grasses. Stands of chamisal chaparral (*Adenostoma fasciculatum*) and Woolly yerba santa (*Eriodictyon tomentosum*) make up the shrub cover on and near slopes.

1.2.2. *Hooveria purpurea* var. *reducta*

Hooveria purpurea var. *reducta* (Asparagaceae) is a recently re-circumscribed species, split from the genus, *Chlorogalum* (Taylor and Keil 2018). The genus *Hooveria* is made up of two species, *Hooveria purpurea* and *Hooveria parviflora* (Hoover 1940; Taylor and Keil 2018).

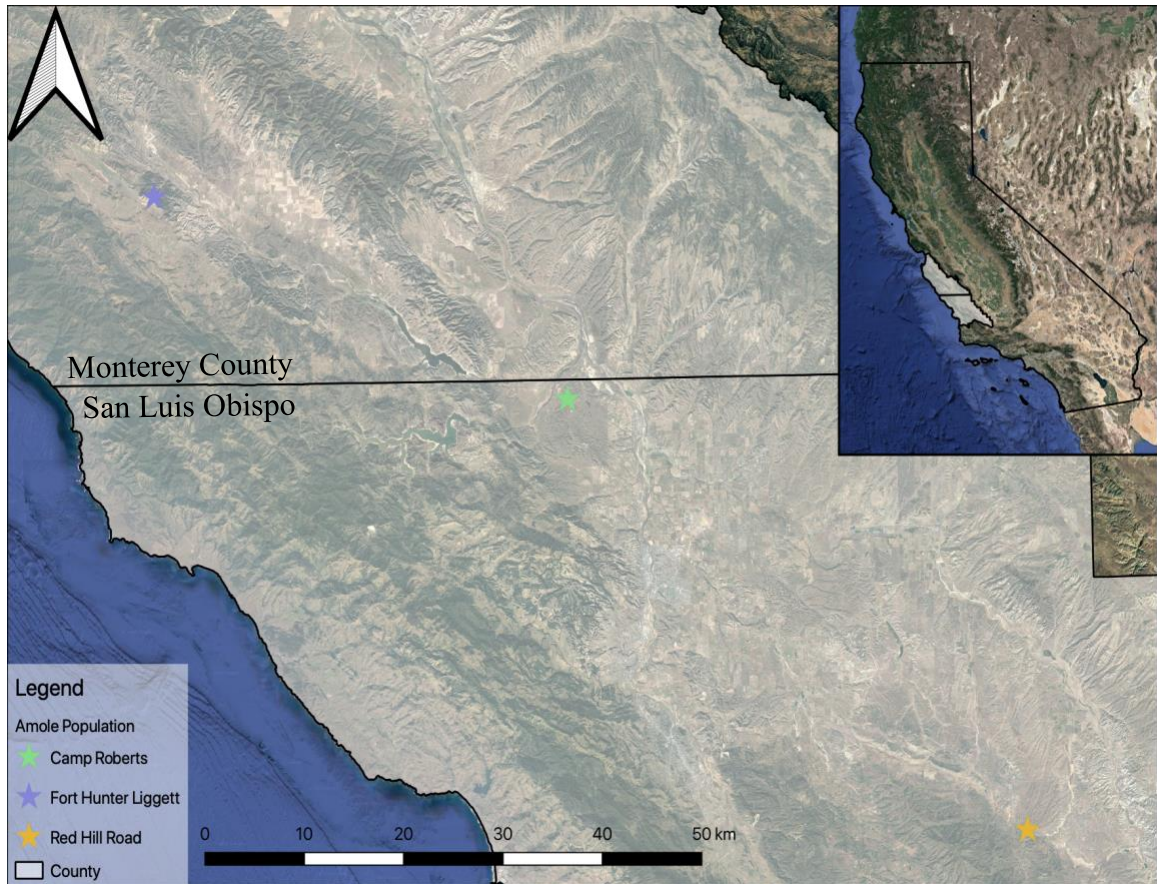


Figure 1: Map of the three known populations of *Hooveria purpurea*. Both Camp Roberts and Fort Hunter Liggett are home to *Hooveria purpurea* var. *purpurea*. Red Hill Road is the only location known to have *Hooveria purpurea* var. *reducta*, the Camatta Canyon Amole.

Hooveria purpurea can be found in both Monterey and San Luis Obispo Counties, with the two varieties separated by 65 km (Taylor and Keil 2018, Figure 1). *Hooveria purpurea* var. *purpurea* and *H. p.* var. *reducta* can be differentiated primarily on size, with var. *reducta* having a much shorter scape and smaller basal rosette. *Hooveria purpurea* is a small geophyte which sends out new leaves in late winter and purple flowers in spring. Its range is restricted to San Luis Obispo County, specifically to Red Hill Mesa in the Los Padres National Forest and some adjacent private property (Taylor and Keil 2018; Kofron *et al.* 2021; Figure 5 [Appendix]). It grows on exposed

ridge-tops in red weathered soils, usually in association with biological soil crusts (Kofron *et al.* 2021; Taylor and Keil 2018).

1.2.3. Study Plots

In the late 1980s, U.S. Forest Service biologists surveyed Red Hill Mesa and established 10 randomly selected, 1 m² plots to monitor CCA on Los Padres National Forest within the CCA's range (Magney 1988). After two years of sampling, these 10 plots ceased being surveyed. These plots were unable to be relocated in 2014 when sampling efforts resumed. Surveyors reestablished those 10 plots based off the original plot descriptions from the 1980s (Magney 1988). These descriptions included latitude and longitude points and descriptions of the plot locations, including bearings from landmarks (mostly trees) to aid in finding the plots. The 10 plots were established with an extra 1 m² buffer plot on either side as an attempt to increase the odds of sampling the same location as the original surveys (Figure 2). Once plot center was found, a 1 m² circular plot was centered around plot center, with the two adjacent 1 m² buffer plots laid out perpendicular to the bearing from the stated landmark (Figure 2).

1.2.4. Treatment Groups

For this project, we paired plots from the group of 10 historical plots. Plots were paired based on geographic proximity (to match microclimate and aspect) and then one of the pair was randomly assigned as a treatment plot and the other a control plot. Control plots were subjected to no thatch removal. Paired plots were: Plots 1 (treatment) and 9, Plots 2 (treatment) and 3, Plots 4 (treatment) and 5, Plots 6 (treatment) and 7, Plots 8 (treatment) and 10. See Table 1 for plot locations.

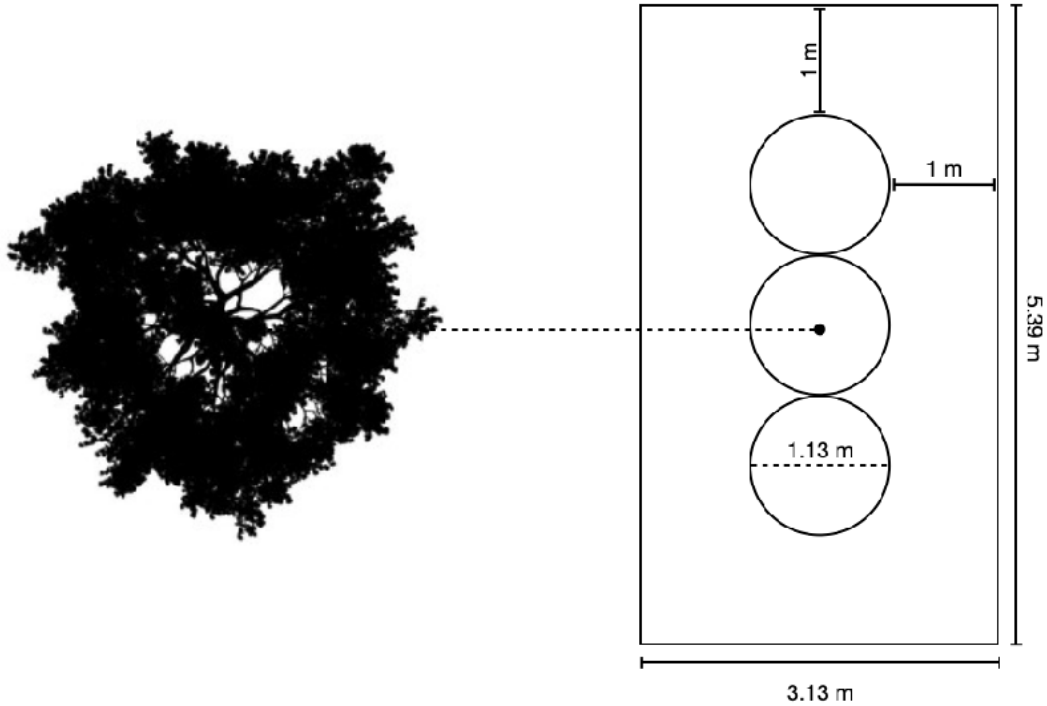


Figure 2: Diagram visualizing the layout of a treatment plot. Dotted line represents a bearing laid out from a marker (usually a tree) to plot center. Bearings with associated markers described in Table 1. The three circular plots are where all CCA counting, percent cover and height measurements took place. The greater rectangle around the three 1 m² circles is where thatch was collected, including within the three circular plots.

Table 1: Center coordinates and descriptions of 10 circular plots (1 m²) along Red Hill Road in Los Padres national forest, San Luis Obispo county. Plots are the same as those used in surveys from USFWS between 2014-2020. Table from Kofron et al. 2021.

Plot	Center coordinates	Description
1	35.40290/−120.27943	Red Hill Mesa E of Red Hill Rd, 4.3 m N of blue oak, bearing 355 ⁰
2	35.40282/−120.27864	Red Hill Mesa E of Red Hill Rd, 6.3 m NW of two small blue oaks (one is dead in 2013), bearing 335 ⁰ , in line with a dead oak in distance
3	35.40295/−120.27846	Red Hill Mesa E of Red Hill Rd, 26.6 m N of two small blue oaks in plot 2, bearing 22 ⁰ , in line with an oak in distance.
4	35.40193/−120.28060	Red Hill Mesa W of Red Hill Rd, 10 m W of small blue oak, original bearing 265 ⁰ , shifted slightly in 2017 due to tree fall, now ~6 m to the N, new bearing 305 ⁰
5	35.40279/−120.28022	Red Hill Mesa W of Red Hill Rd, 69.6 m WSW of blue oak in plot 1; 4.7 m SSW of chamise and bearing 205 ⁰
6	35.38929/−120.28736	Red Hill Ridge W of Red Hill Rd, 9 m NW of three-trunk blue oak, bearing 320 ⁰
7	35.38945/−120.28746	Red Hill Ridge W of Red Hill Rd, 4.6 m NW of two-trunk blue oak, bearing 300 ⁰
8	35.40336/−120.27943	Red Hill Mesa E of Red Hill Rd, 3.75 m SE of pipe fence (white paint mark)
9	35.40305/−120.27956	Red Hill Mesa E of Red Hill Rd, 12 m SE of U.S. Forest Service sign
10	35.40331/−120.27970	Red Hill Mesa W of Red Hill Rd, 12 m S of another U.S. Forest Service sign, next to road

1.2.5. Treatments

Prior to any treatment, we measured ocular plant cover of all plants and averaged it across the three 1 m² circular plots (Figure 2). Additionally, we recorded mean thatch height by placing a ruler at ground level haphazardly four times throughout each circular plot, taking the mean of the 12 total measurements.

The first round of thatch removal was performed on November 21, 2020. Thatch was removed from around the three plots, buffered by 1 m on all sides to create a 5.39 m x 3.13 m treatment area (Figure 2). Thatch was removed everywhere within the 5.39 m x 3.13 m rectangular area. We reduced annual grass height to 3.5 cm by hand with shears and removed loose thatch with rakes. We collected dry thatch in trash bags to weigh. Once collected, all biomass was placed in dryers at 100°F for 5 days. We visited plots every two weeks between November and February 2021 to check if the Camatta Canyon Amole had emerged. The same process was repeated for Year 2: November 2021-February 2022, where we collected thatch in November 2021 and February 2022.

We counted all CCA within the three circular plots in March 2021 and March 2022. The total number of CCA in these plots was counted and divided by 3 to establish the number of CCA/ m². The three plots were buffered by 1 m on all sides to create a 5.39 m x 3.13 m treatment area. Thatch was removed everywhere within the 5.39 m x 3.13 m rectangular area.

1.2.6. Statistical Analysis

All statistical analyses were performed in R version 4.1.2 (R Development Core Team 2022) using the R packages “stats” (R Core Team 2021), “car” (Fox and Weisberg

2019), and “lme4” (Bates *et al.* 2015) . All plots were made using “ggplot2” (Wickham 2021).

1.2.7. Difference in Biomass

To test for the difference in thatch height and density between treatment and control plots, we used a one-way anova. We ran one anova comparing height between plot types and another comparing density between plot types. We tested our one-way anova to see if it met all assumptions for homogeneity of variance, normality and independence. Both Anovas met all the assumptions.

To compare the differences in thatch collected over time, we ran a one-way ANOVA using collected thatch as our response variable and rain year as our grouping variable. We use rain year as opposed to calendar year for consistency throughout our analyses. The goal of this step was to see if our treatments were affective at eliminating invasive across years.

1.2.8. Differences in CCA in Plots

To test for the differences in the emergence of CCA, we used a generalized linear models (GLM). The count data we collected is heavily right skewed, so we used poisson distributions for our GLM. The first model used emergent amole as the response variable, with rain year and plot type (control/treatment) as the explanatory variables. All assumptions for a poisson distribution were met.

To come to a complete picture of the trends in CCA over time, we ran a GLM of CCA counts between 2014-2022 as our response variable, and rain (mm) as our explanatory variable. We obtained CCA counts between 2014-2020 that were presented in Kofron *et al.* 2021, Table 3. Rainfall data was collected from the Shell Creek weather

station (SLO-6)¹. For this analysis, we thinned the data down to just plots 5, 6, 7 and 8 between 2014-2022 because the other 6 plots had almost 0 CCA in those eight years.

To analyze yearly totals between 2014-2022, total CCA in a year was the response variable, with rain fall (mm) as an explanatory variable and year as a random effect. The count data fit a poisson distribution.

1.3. RESULTS

1.3.1. Biomass

We collected thatch four times, across two rain seasons: between November 2020- February 2021 and November 2021- February 2022. Figure 3 shows the biomass of thatch collected across two rain years. We collected significantly more thatch in the first rain year than in the second rain year (ANOVA, $p = 0.0428$, $F = 4.901$, $DF = 1$), meaning that thatch was slow to accumulate between collections. We collected 24x the amount of thatch between November 2020- February 2021 than we did between November 2021- February 2022.

Prior to the first round of treatment, thatch height (ANOVA, $p = 0.277$, $F = 1.363$, $DF = 1$) and cover (ANOVA, $p = 0.277$, $F = 1.4$, $DF = 1$) were similar between treatment and control plots. Thatch cover did not recover between rain years (ANOVA, $p = 0.0152$, $F = 7.204$, $DF = 1$) (Figure 7, appendix). In 2020, treatment plots and control plots had a similar amount of cover, but by 2022 treatment plots had lower thatch cover than control (Figure 6, appendix).

¹ <https://ucce-slo.westernweathergroup.com>

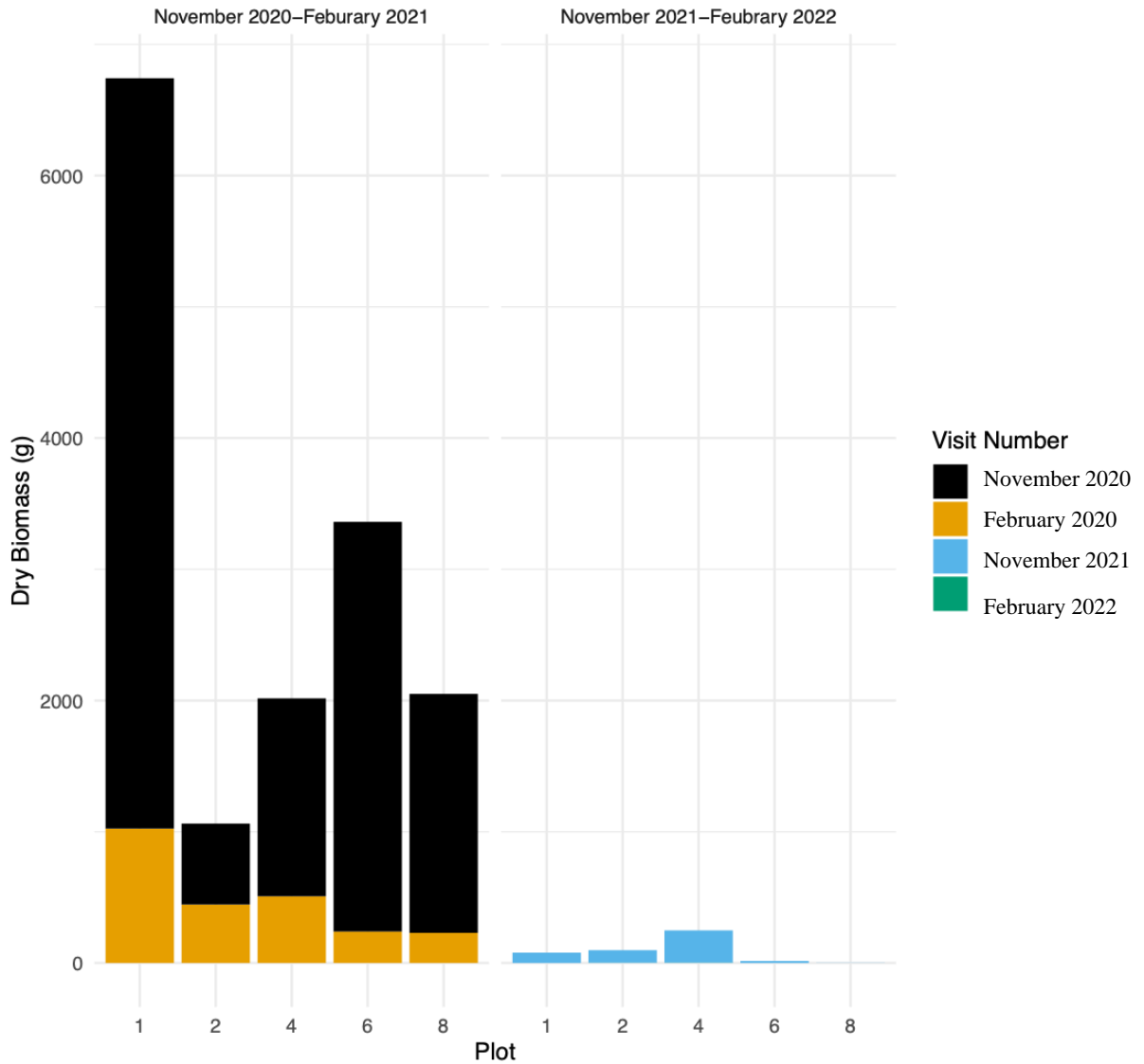


Figure 3: Biomass of thatch collected across four visits, grouped into two rain years. A majority of the thatch collection was done in November 2020 and February 2021 (first rain year). Thatch did not recover between rain year 1 and 2. Very little thatch was collected in visit 4 (not even visible in the above graph).

1.3.2. Emergent Amole

We counted the number of CCA in ten vegetation plots twice over the course of two rain years. Treatment had no effect on CCA emergence in plots between 2020 and 2022. (GLM, $p_{\text{rainyear}} = 0.530$, $z_{\text{rainyear}} = 0.629$, $p_{\text{plot type}} = 0.788$, $z_{\text{plot type}} = 0.269$, $DF = 11$).

After removing plots that had 0 CCA in them, rain (mm) was found to be positively correlated with a higher number of CCA between 2014 and 2022 ($n = 4$, GLM, $p = 0.00174$, $z = 3.132$, $DF = 39$).

We found no relationship between number of CCA and year in the 2014-2020 data (GLM; $p = 0.591$, $z = 0.537$, $DF = 6$). We also found no relationship between rain (mm) and total Amole (GLM, $p = 0.591$, $z = 0.537$, $DF = 6$) Similarly, there is no relationship between year and number of CCA in the belt transect data (LM; $p = 0.159$, $t = 1.728$, $DF = 4$).

1.4. DISCUSSION

Thatch removal in 2021 and 2022 had no affect on Camatta Canyon Amole emergence. Even when CCA counts were aggregated by year, and compared across 8 years, we still found relationship between thatch removal and CCA emergence. Perhaps CCA emergence is better explained by other factors.

For instance, there is robust research supporting the hypothesis that the primary effect of invasive annuals is the suppression of seedlings (Mordecai 2012; Grime *et al.* 1981; Wolkovich, Bolger and Cottingham 2009; Lenz *et al.* 2003). Plots with the highest number of CCA were also those with the highest amount of thatch prior to treatment, indicating that direct competition was not responsible for CCA decline. Perhaps mature CCA are dying at their standard rate, but the recruitment of young individuals has halted due to the competitive pressures of invasive annuals.

Perhaps our window of observation was too narrow to observe any change in CCA emergence. A majority of studies that test for the affect of invasive annuals on native perennials only test for one year, and track only the affect of the removal of said

invasive (Kettenring and Adams 2011). Our study of two years may still have been too small an observational window. We still understand so little about the biology and life-history of the Camatta Canyon Amole, let alone other geophytic species. For instance, there is evidence to suggest that between 25-50% of a species of a geophyte population experiences prolonged dormancy in any given year (Lesica and Steele 1994; Tatarenko 2019; Shefferson, Kull, Tali, 2005; Shefferson et al. 2018). Prolonged dormancy of this kind is common for 1-2 years, but has been reported for as long as 5 years (Lesica and Crone 2007). Perhaps the confluence of both a large dormancy period and exceptionally low rainfall in 2020 and 2022 resulted in low CCA observations. It is for the above reasons why this research has been set up as a long-term monitoring study, in which data will be recorded yearly in all 10 plots. With time, clearer trends in CCA will emerge.

Broadly speaking, plants of different lifeforms are likely to differ greatly in their response not only to the removal of thatch, but the way in which thatch is removed. The simulated grazing presented in this study had no effect on CCA emergence, but other removal measures might. In South Africa, where geophyte diversity is high, fire is the constraining effect on geophyte abundance (Manning *et al.* 2002; Proches *et al.* 2006). In California, it is thought that geophytes bloom the most vigorously the first spring following a fire (Keeley 1988). Perhaps fire-removal studies are a more adequate way to measure perennial bulb response to thatch removal. In which case it is the removal of both the direct competition of the invasive plant as well as structural properties of the thatch layer that stimulates geophyte restoration. There is some work with perennial grassland communities that suggest a combination of disturbance regimes (herbicide, fire,

seeding) may be the most effective at increasing native perennial density (Stanley *et al.* 2011; Adams *et al.* 2020).

In conclusion, our study was unable to determine the effect of thatch removal on the number of CCA in our plots. While thatch removal was successful, in that invasive grasses did not reemerge in force year-after-year, it remained ineffective in spurring an increase in CCA. These results may stem from the unknowns in the life-history of the Camatta Canyon Amole, or the time frame in which the project was done. Perhaps future efforts in conserving geophytic species use alternative measures at stimulating emergence, such as fire, which has proven affective with other geophytes throughout the world.

CHAPTER 2

2. Understanding the geographic distribution and population size of Camatta Canyon

Amole (*Hooveria purpurea* var. *reducta*)

2.1. INTRODUCTION

California is one of the world's 36 biodiversity hotspots, which are collectively home to 50% of the world's plant species (Mittermeier *et al.* 2011). These hotspots are also those most under threat from anthropogenic degradation, in which we expect to see a 30% reduction in species abundance due to direct human intervention (Bellard *et al.* 2014). This threat has spurred a variety of conservation organizations and scientists to create protocols and monitor these hotspots to lessen the threat of human disturbance (Price 1994)

Quantitative data regarding populations we are interested in conserving is necessary for the mission of protecting vulnerable species. The U.S. Endangered Species Act requires preliminary information about said species, including population and range estimates (USFWS 2009a). A species' range and population size estimates are important because plants with a narrow distribution or with few populations are considered especially vulnerable (Leimu *et al.* 2006). Furthermore, accurate data on population declines are paramount in understanding the long-term viability of populations in the face of further human disruption, and in creating adequate steps in conservation (Parmesan & Yohe 2003; Pereira & Cooper 2006). However, not all plant populations are created equal.

Small populations of plants are more vulnerable to extinction, and are more genetically homogenous than larger, more abundant populations (Shaffer 1981; Lande

1988; Ellstrand & Elam 1993). Because threats are felt most strongly when populations are small, it is no surprise that highly endemic species, often with small populations, make up a large proportion of the Red List and other rare plant lists (Schatz 2009; CNPS 2022). There are many drivers of species endemism. “Ecological endemics”, for example, are species that have evolved very strict ecological requirements (Daubenmire 1987). One-third of California’s rare plants would be considered “ecological endemics”, which are restricted to specific soil substrates (Skinner and Pavlik 1994; Safford *et al.* 2005; Safford 2011; Safford and Miller 2020). The combination of both threatened soil edaphic species and threatened soils (Amundson, Guo & Gong 2003; Drohan and Farnham 2006) highlights the need for comprehensive approaches to conservation.

An approach to conservation growing more and more common is the utilization of private land. Agriculture and ranching poses a great risk to unique soil types and endangered species in California (Amundson, Guo & Gong 2003; Drohan and Farnham 2006; Reiner & Craig 2011; Uematsu *et al.* 2010; Wilcove *et al.* 1998). It is also believed that many populations of rare plants exist on private property, especially in the West, which have gone completely un-surveyed (Lovett-Doust *et al.* 2003). In fact, it has become the primary means of conservation for many land trusts to acquire cattle ranches and other private land for species protection (Merenlender *et al.* 2004). While the success of these smaller reserves depends on the species that fall under their protection, they can be functional for narrowly edaphic species (Shaffer 1981; Parker 2012).

H. purpurea var. *reducta*, the Camatta Canyon Amole, is one of those narrowly endemic, soil specialist species (Magney 1988; Kofron *et al.* 2021) While the majority of its known distribution is in Los Padres National Forest, biologists currently believe that

much of its range occupies nearby private property (Kofron *et al.* 2021; USFWS 2020; Figure 5). Camatta Ranch, a cattle ranch directly north of Red Hill Road, is thought to host many Camatta Canyon Amole (CCA). Camatta Ranch has never been accessible for surveying for CCA. The goal of this study is to survey Camatta Ranch to better understand the range of the Camatta Canyon Amole, and estimate population size on the ranch. Additionally, we will utilize species niche modeling to estimate and compare the suitable habitat for the CCA on Camatta Ranch and Los Padres National Forest.

2.2. METHODS & MATERIALS

2.2.1. Study Site

Camatta Ranch is a 32,000 acre cattle ranch in the interior San Luis Obispo County. The ranch is made up of rolling hills dominated by Blue Oak (*Quercus douglasii*) savanna. It experiences a typical California mediterranean climate, with cool wet winters and dry hot summers. Nearly all rain occurs between October and March. The southern extent of the ranch has been mapped as Arbuckle Sand Loam, the Camatta Canyon Amole's preferred substrate (Kofron *et al.* 2021; USFWS 2020). According to the owners of the ranch, CCA occupies three ridges at the southern end of the ranch, just north of Highway 58 and the entrance to Red Hill Road. The three ridges that CCA occupy are regularly grazed by cattle, which has been ongoing since the 1840s, and are often devoid of above-ground biomass by grazing.

2.2.2. Experimental Design

2.2.2.1. Spring 2020 Survey

To get a sense of the distribution of the CCA on the southern part of the ranch, we surveyed the three ridge tops the landowners identified as being the main extent of CCA.

In Spring 2020, we surveyed three ridges that had been mapped as Arbuckle Sandy Loam. These surveys were done using a meandering transect. We collected point and polygon data on ArcGIS Collector app with a Samsung Galaxy Tab E and iPhone 11's (Goff *et al.* 1982; Arc GIS Collector, version 21.0.2, ESRI). Points were collected haphazardly throughout groups of CCA. The goal of these points was to capture when clusters of CCA began and ended along the ridge.

Polygons were collected when the group of CCA could be seen from a central location. Meaning, when a group of CCA was close enough to see the entire group, we walked the circumference of that group and drew polygons. All paths walked along the three ridges were recorded.

2.2.2.2. Spring 2021 Survey

The next round of surveys were conducted to create polygons that would eventually be used in estimating the population of CCA on Camatta Ranch. In Spring 2021, we created points every 30 m along each of the three ridges on Camatta Ranch. At each 30 m point along the ridge, we recorded if CCA was present within 15 m of the point or not. If CCA was not found within 2 minutes of searching, we moved on to the next survey point. These surveys took place on March 17, 2021.

2.2.2.3. Creating Survey Polygons

We used the point data collected in the Spring 2020 survey to create polygons of the CCA patches. Only “presence” points were used to create these polygons. Each point was buffered by 15 m, creating a 706 m² search area. This represents the search area of the Spring 2021 Survey. Any points or polygons from the Spring 2020 survey that

were within 30 m of these circular polygons were merged into new polygons. Merging polygons created successive linear polygons or amorphous polygons.

Each survey polygon was assigned two random bearings (1-360°) generated in R. Once in the field, we navigated to the polygon centroid, took our bearings and laid out a transect tape from the centroid to the edge of the polygon (distances varied - See Figure 10). To count CCA, we walked a belt transect along the two randomly assigned bearings and counted all CCA found within .5m to either side of the transect line until the edge of the survey polygon. Belt transect lengths were unequal due to the irregular shape of the plots. These surveys resulted in number of CCA / m² for each polygon. In total, we created 37 polygons of CCA patches and randomly selected 5 polygons per ridge line to survey. 15 of 37 polygons were surveyed on May 8th, 2021.

In an attempt to validate the population estimates in these plots, we counted the all observable CCA in two randomly selected plots. A large team set out with pin flags and walked transects through these plots, and set a pin flag per CCA individual. We then compared the observed density of plots with the estimated density to validate our estimates.

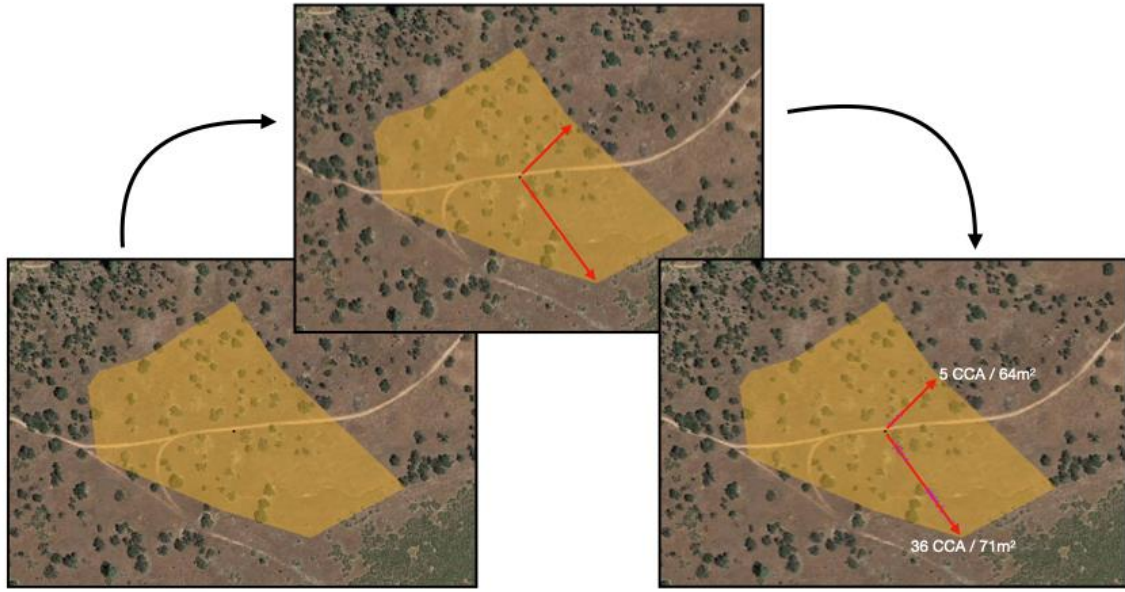


Figure 4: Plot 6 on Camatta Ranch. Shows process for assigning compass bearings and counting CCA along those belt transects. In this example, a total of 41 CCA was counted along 135 m². This averages out to ~ 0.3 CCA/ m² in this plot.

2.2.3. Species Niche Model Variables

Typical models utilize environmental variables to characterize the species' ecological niche, however, our area of interest is incredibly narrow. With a known range of ~ 1 km², variables with coarse data are of little practical use to model creation. Because we know from literature and field surveys that the CCA occupies the flat area of ridge tops, we selected variables related to topography. We downloaded a digital elevation model (DEM) from USGS with a spatial resolution of 1 m (USGS 2021), and calculated topographic roughness index (TRI) , slope and aspect from the DEM at both 1m and 15m resolutions. Additionally, calculated the normalized difference vegetation index (NDVI) using Planet (Planet Team 2017) satellite imagery. We used presence points gathered from field work done in Spring 2020 (n = 187) and generated 2000 pseudo-absence points.

2.2.4. Ecological Niche-Based Distribution Modeling

To model the species distribution of *H. p. var. reducta*, we used Maxent, a widely used method to model species distribution from presence-only data (Elith *et al.* 2011; Radosavljevic and Anderson 2014). After thinning points, models were calibrated with the remaining 127 presence points and 2000 pseudo-absence points. We used the *set.seed()* function to make pseudorandom points repeatable. Our seed was set to “15”. The region surrounding all presence points was buffered by 11 m, as that represents a reasonable land area that the CCA could disperse and occupy on the ranch. We used linear, quadratic, product and hinge (L, Q, P, H) feature classes to model the relationship between occurrence in a cell and the values of environmental predictors. We tested the model on four regularization parameters—1, 4, 7, 10—and a 4-fold cross validation.

To evaluate our model, we calculated the area under the receiver operating characteristic curve (AUC) test statistic (*dismo*, Hijmans and Elith, 2017). To test the validity of our best model, we calculated null models with random data. We used the *ENMnulls()* function to create 100 iterations of our best model (Raes & Steege, 2007). The *ENMnulls()* function automatically generates random point data. We used the *set.seed()* function to initialize the pseudorandom numbers that *ENMnulls()* would pull from. Our seed was set to “15”. We were then able to compare the results of the null model to our empirical model. If the mean AUC of the 100 null models were close to that of the empirical model, we have reason to doubt the validity of our empirical model. If the mean AUC is significantly different than our empirical model, then we can confirm that our data has higher predicting power than random data.

2.2.5. Comparing SDM to Known Occurrences of CCA

In order to compare the distribution of CCA on Camatta Ranch to the population on Los Padres National Forest, we needed a polygon that encapsulates a theoretical distribution of the plant. On Los Padres National Forest, the California Natural Diversity Database (CNDDDB) reports the distribution of CCA on Red Hill Road that covers an area of 297,528 m². This polygon represents extensive field surveys conducted in the 1980s in Los Padres National Forest, and is the only polygon that CNDDDB has for CCA that was created as a result of a specific survey. We calculated the average cell value from our SDM within this CNDDDB polygon, and filtered SDM cells on the ranch by that value.

2.3. RESULTS

2.3.1. Ranch Estimates

Fifteen of 37 plots were surveyed for Camatta Canyon Amole. The ratio estimate of CCA (plants per area) in plots varied from 0 CCA in plot 26 to 0.641 CCA / m² in plot 34. To adequately estimate density from this survey, which estimates unequal area belt transects, we calculate the ratio estimator, which is the estimated ratio between two sample means (Steham and Salzer 2000). Because plots were of unequal length, we calculated the density ratio for each plot individually, and then got the mean for those plots. The formula is shown below:

$$D = \frac{Y}{A} = \frac{2.61428571}{13.3714286}$$

where D is the ratio estimate, Y is the sample mean number of CCA per transect, and A is the mean transect area (m²). The standard error of our estimate is 0.0398. This approach

to evaluating density between belt-transects of various areas more closely estimates a “true” population size.

To ground truth our estimates, we counted the total number of above-ground amole in plots 22 and 33. Given their respective density ratios, we predicted that plot 22 would have 712 CCA and plot 33 to have 161 CCA. The actual number of CCA in plot 22 was 831. The actual number of CCA in plot 33 was 2117. For plot 22, our prediction was incorrect by a factor of 1.16, while for plot 33 we were off by a factor of 13.15. This suggests that we routinely underestimated the amount of CCA on Camatta Ranch. To adjust our estimated counts, all estimates were multiplied by an average correction factor of 7.15.

Our density ratio is 0.19551282 CCA / m², with a standard error of 0.398. The area of the 15 survey plots is 96,071 m². The density ratio was applied to the area of the 15 survey plots to get a population of 18,783 ±4,904 plants (95% CI [13,879; 23,682]). The area under all 37 plots, which includes those not surveyed, is 134,000 m². Applying the same density ratio, we calculate 26,198 ±6,945 (95% CI [19,253; 33,143]). As a point of comparison, the occurrence of CCA along Red Hill Road is approximately 297,528 m². Using our density ratio, we estimate that there are 57,720 ±15,190 plants (95% CI [42,530; 72,910]).

To get a complete estimate of the population of CCA across all three ridges, we required a definitive boundary or polygon that encapsulates the entirety of its distribution. To do this, we used our species niche model. Our species niche model predicts the cells within the CNDDDB polygon along Red Hill Road (Occurrence 1) as having ≥ 0.5 probability of containing CCA. To create a polygon that represents the distribution on

Camatta Ranch, we removed all cells from our SDM with values < 0.5 that occurred on the ranch. We chose cells with < 0.5 probability because our SDM predicts a known population of CCA along Red Hill Road with a value of ≥ 0.5 . We hope that this method lends itself best to comparing two populations of the same plant that occur in cells with equal predictions. This resulted in a predicted 3,397,530 m² of habitat on the ranch. Using the density ratio calculated above, we estimate that there are $\sim 664,260 \pm 173,460$ CCA on Camatta Ranch (95% CI [490,800; 837,720]).

2.3.2. Species Niche Model Results

Our model of *Hooveria purpurea* var. *reducta* generated predictions for feature classes (fc): H (Hinge), L (linear), LQ (linear + quadratic) and LQH (linear + quadratic + hinge). To select the optimal model, we used a sequential method that cross-validates results by selecting for the lowest average test omission rate, and then the highest AUC (Radosavljevic & Anderson 2014; Kass *et al.* 2020).

Our optimal model has an omission rate = 0.09333333 and an AUC = 0.9130344 (Figure 12). Variable response curves indicate that areas of low terrain ruggedness (tri) and an elevation of about 600 ft represent favorable environments for *H. p.* var. *reducta* (Figure 13). Slope and aspect have little to no effect on the presence of CCA. Low NDVI values were preferable to CCA as well, as high values are likely trees and shrubs that CCA avoids growing by (Figure 13). We quantified model overfitting by comparing threshold-dependent omission rates. To use these omission rates, we compared the observed omission rates to theoretical omission rates. For an optimal model, we expect zero omission of evaluation localities using the lowest presence threshold, and 10 % omission for the 10th percentile omission rate (Fielding & Bell, 1997; Peterson *et al.*

2011). The 10% training omission rate, for our optimum model is <0.08 . This metric indicates the proportion of pseudo absence points with values lower than that which excludes the 10% of training data with the lowest suitability scores (Fielding & Bell, 1997; Peterson and Soberon 2012).

To get an accurate picture of the true performance of our model, we built null models with random data. The point of a null model is to test if, with completely random data, we get different AUC scores than our empirical model (Raes & Steege 2007). We ran 100 null iterations of our preferred H model with a regularization multiplier, $rm = 7$. That results in a mean AUC for the Null model of 0.5, which is significantly different than our empirical AUC of 0.913 (Figure 14). The prediction output can be found in Figure 15.

The map of suitable habitat spans the length of the three ridge tops of Camatta Ranch, and predicts the CNDDDB polygon along Red Hill Road as high suitability (> 0.5) (Figure 16). Additionally, areas of Los Padres National Forest along the Burnout OHV trail that were surveyed after model creation were predicted as suitable (> 0.7) where Amole was subsequently found.

2.4. DISCUSSION

2.4.1. Survey Effort

Surveying plots on Camatta Ranch should be done annually in order to capture the variance of emergence in Amole along each transect. As prefaced earlier, CCA experiences a dormancy period of unknown length (Koch, unpublished), so the number of Amole available for counting changes annually. Understanding approximately how many Amole are dormant in any given year is necessary for more accurately estimating

population sizes. Additionally, the population estimate provided here is an important measure of comparison for future conservation success. Thus, research that tracked individual plants over time would aid in our understanding of CCA's life history, and the populations long-term viability (Lande 1988).

The coupled survey-model design presented here will only increase the predicting power of future models and population estimates (Guisan *et al.* 2005). However, designs like ours aren't without bias. A completely random polygon selection process could have reduced any bias in the survey effort. However, this would have required more time and resources to complete a comprehensive survey, with more samples to overcome variability between plots. We believe that our current approach was the most efficient for sampling occurrence data for a patchily-distributed rare plant.

2.4.2. Niche Model on Camatta Ranch

Here we combine the use of species distribution modeling and field techniques to get a complete look at the distribution and population of the Camatta Canyon Amole. Kofron *et al.* 2021 maintained that it was likely that a majority of the species occurred on Camatta Ranch (Kofron *et al.* 2021; USFWS 2020; USFWS 2022). Our findings suggest that the population of Camatta Canyon Amole on Camatta Ranch is likely much larger than previously expected. Based on our estimate, somewhere between 80%-90% of CCA occurs on Camatta Ranch. The population size of any endangered species is of primary concern in designating a rarity status (He and Gaston 2000). The importance of this research is often not well reflected in endangered species programs (Schemske *et al.* 1994; Bayliss *et al.* 2013), where expert opinion alone has historically been the only source of information (Hurlbert & Jetz 2007; Lacher *et al.* 2012; Fitzgerald *et al.* 2021).

Using SDM's to narrow down search areas for hard-to-find rare plants may be useful for the future of conservation and reserve creation (Guisan *et al.* 2014; Franklin 2013). A draft recovery plan for the Camatta Canyon Amole includes a proposed \$10,000,000 to protect currently unprotected habitat (USFWS 2022). Our model predictions has helped delineate what does and does not need to be protected on currently unprotected land, primarily on Camatta Ranch. Small reserves like this may provide adequate long-term protection for rare plants with narrow edaphic restrictions, such as the Camatta Canyon Amole (Parker 2012; Shaffer 1981). The acquisition of easements designed to protect private property, where many rare species now reside, should be a major activity of conservation agencies; public and private alike (Merenlender *et al.* 2004; Lovett-Doust *et al.* 2003).

Modeling rare plant distributions without climate data will become more and more important as the race to protect rare species increases, as most climate data is still too coarse to apply to narrow endemics. The spatial scale that is required for modeling their distribution is often unavailable for most non-topographic data. Worldclim, which is a set of commonly used environmental variables, has a minimum resolution of 4.5 km. With a dearth of fine-resolution environmental variables, using topographic data to map species distributions is typical for rare plants. Here, slope and terrain ruggedness proved to be the most important variables in the model, and are confirmed by what we already know about the CCA; that it grows on flat ridge tops. Because many rare plants have a narrow preference for certain soil types (Wamelink *et al.* 2014), future modeling should attempt to include soil data.

In conclusion, our data suggests that 90% of the geographic range of the Camatta Canyon Amole occurs on Camatta Ranch, just adjacent to Red Hill Road. Our model maps the CCAs distribution almost perfectly with the soil map of Arbuckle Sandy Loam on ridge tops, which confirms past evidence of CCAs growing preferences. The accuracy of our counts, while underestimates, provide a good starting place for evaluating the presence of rare plants in large areas. As a result, this research provides a template for counting rare plant populations and measuring their distribution on previously upsampled properties.

CHAPTER 3

3. Quantifying Research Effort in California's Flora

3.1. INTRODUCTION

Rare species have been the subject of much discussion and research since the advent of modern biology. With the reality of global climate change and its expected impact on biodiversity, cataloging species diversity has never been more important. It is expected that 40% of species will experience a significant loss in range in the coming years as a result of climate change and human disturbance (Ceballos, Ehrlich & Dirzo, 2017). Additionally, many rare species have small scattered populations, which makes conservation decisions difficult (Leimu *et al.* 2006; Schemske *et al.* 1994; Holsinger 1991; Rabinowitz 1981).

California is home to the highest number of rare and endangered plant in the nation (Baldwin *et al.* 2017), with 286 listed threatened or endangered (CNDDDB 2022). While not reflected in law, the California Native Plant Society (CNPS) lists over 1,000 native plants as meeting the standards set forth in the California Endangered Species Act (CNPS 2022).

Protecting rare species in California is big business. Environmental consulting is a \$3 billion industry in California, and operates with the goal of mitigating risk posed to species of plants and animals (IBISWorld, 2019). One facet of risk mitigation for these organizations is funding research and species assessments for species of interest. It is, however, important to note that any and all research regarding rare plants is beneficial to their conservation.

The kind of information generated by this research is also important. Research into rare plant species often fails to accurately and consistently collect data that is key to their protection, such as population sizes and demographics (Bevill & Louda 1999; Schemske et al. 1994). Focusing our research attention towards these data deficient groups is important in meeting the goals of biology and conservation (Okuyama 2010). As highlighted previously, many conservation organizations in California focus a large portion of time and money into California's rare plants. With approximately a quarter of California's flora considered rare (CNPS 2022), and the threat of these plants from human disturbance increasing (Baldwin *et al.* 2017; Ceballos, Ehrlich & Dirzo, 2017), the barriers to rare plant research aren't becoming any easier to traverse. Additionally, studying some rare plants is often made difficult by low/sparse population sizes (Guisan *et al.* 2006; Rabinowitz 1981), which results in often considerable gaps in our knowledge (Lyons *et al.* 2005).

These gaps aren't just reflected at the species level. There are other biases in publishing as well, with certain genera (Okuyama 2010), morphological characteristics (Yang *et al.* 2021), and higher level taxonomic groups (Mammides 2019) that have more research and conservation emphasis than others. It is likely then, that discrepancies in publishing spans many ways of categorizing plants, such as species status, habit, flowering time, and much more.

The question this paper is attempting to answer is: How much research effort has gone into researching rare plants in California, and how does this compare to other natives and weedy species?

To answer this question, we gathered two sets of data to use as proxies for research effort: (1) the amount of literature a species appears in, and (2) the number of sequence data publicly available for each species. We then compared the literature and sequence counts between rare plants, weedy plants and non-rare native plants to assess research discrepancies.

3.2. MATERIALS AND METHODS

3.2.1. Scraping

We used python libraries “pandas” (McKinney *et. al.* 2021), “bs4” (Richardson 2007), “re” (Van Rossum 2020) and “requests”² to scrape Google Scholars query results number. Google Scholars search algorithm combs through any available piece of text for associated literature including paper titles, abstracts and the main text of a paper.

3.2.2. Search List

To assess differences in research effort of rare vs. non-rare taxa within California's flora, we collected total search results for individual species using Google Scholar. Our study uses the Jepson flora Project (2021) as the basis for our search list, and includes native status and any taxonomic synonyms. To classify rare species, we used the CNPS Inventory of Rare and Endangered Plants of California (2021). While there are many other ranking systems for rare plants (I.e. NatureServe), the CNPS list was the most targeted list we found. For invasive species, we include only those tracked by the California Invasive Plant Council (Cal-IPC, 2021), which excludes most naturalized species. In total, our search list includes 6,200 plant species that were featured in the Jepson flora Project and Cal-IPC lists.

² <https://docs.python-requests.org/en/latest/>

The same species list was used for gathering sequence data in GenBank. Using the same process as described above, we gathered the total number of sequences (cRNA, genomic DNA/RNA, mRNA, ncRNA, rRNA, tRNA, transcribed RNA) from GenBank for each species. We recorded the total number of query results given our search parameters. This may include multiple submissions for the same gene.

3.2.3. Plant Area

To control for the difference in the amount of land a species occupies, we aggregated GBIF data for each species, and binned points into hexagonal tiles that cover California. We used hexagonal tiles from the Uber H3 geospatial indexing system. The hexagonal grid was set to a scale of 32 km². Once points were binned into the H3 tiles, we calculated the area under the hexagons. The reasoning behind this step was to control for the variability the amount of literature published between widespread species and California endemics. We were really concerned with weedy species, which by definition occur in other places throughout the world. It would thus be no surprise that there is more literature on these species.

Five-thousand five hundred and fifteen taxa resulted from pulling geographic data from GBIF. Some species were automatically lumped from their sub-species level taxonomy (varieties and subspecies) into species level (e.g. *Zygophyllum fabago* var. *brachycarpum* -> *Zygophyllum fabago*). This was a result of taxonomic discrepancies between GBIF and Jepson flora project. Area was recorded in km².

3.2.4. Statistical Analysis

The literature count data from Google Scholar and GenBank are right skewed. We ran a dispersion test in R that tested for the over dispersion of our count data. We found

that the count data was over dispersed, and decided to use a Negative binomial regression, which is used for over-dispersed count data. We used the R package “MASS” (Venables & Ripley 2002) and the function *glm.nb()* to run our negative binomial regression analyses.

To test for the relationship between the species status (Non-Rare Natives, Rare Natives and Weeds) and the amount of found research, we used a negative binomial regression with literature results as the response variable, species status and geographic area as explanatory variables along with the interaction between area and species status.

To test for the relationship between species status and the amount of sequence data found, we used the same model as above, with found sequence data as the response variable.

We then compared the literature found based on CNPS rare plant ranking. California Rare Plant Rankings (CRPR) ranges from plants that’s presumed extinct (1A) to plants now on a watch list (4A). We used a negative binomial regression with literature results as the response variable, CRPR ranking and geographic area as explanatory variables along with the interactions between ranking and species status.

3.3. RESULTS

3.3.1. Rare, Native, Weed Publishing Bias

Rare plants have a lower likelihood of being published about than both non-rare natives (negative binomial regression, $z = -33.73$, $df = 6161$, $p < 2 \times 10^{-16}$, mean = 343.967) and weedy species (negative binomial regression, $z = 21.53$, $df = 6161$, $p < 2 \times 10^{-16}$, mean = 19134.895), even after accounting for area. The overall distribution of the data is skewed towards zero, with weedy species making up most of the tail (Figure 17,

appendix). Half of all native plants have below 30 search results from google scholar (Figure 17, appendix; Figure 18, appendix). Rare plants made up a majority of plants published below the median.

The same general pattern exists for genetic data. Rare plants have the lowest likelihood of having any sequence data available on GenBank (negative binomial regression, $z = -43.21$, $df = 6161$, $p < 2 \times 10^{-16}$, $mean = 141.273$), and weeds the highest (negative binomial regression, $z = 24.80$, $df = 6161$, $p < 2 \times 10^{-16}$, $mean = 43961.484$). The median is more extreme in this case, with half of all species in California have below 8 sequences on Genbank (Figure 19).

3.3.2. Rare Plant Rank Publishing Bias

For this analysis we excluded CNPS rank 4.1 because it consists of only two plant species in our data set. CRPR ranks and the number of plants associated with those ranks are reported in Table 3 (appendix). CNPS ranks 1B.2 (negative binomial regression, $z = -5.584$, $df = 6161$, $p = 2.35 \times 10^{-8}$, $mean = 31.826$) and 1B.3 (negative binomial regression, $z = -7.374$, $df = 6161$, $p = 1.66 \times 10^{-13}$, $mean = 18.896$) had significantly less literature published on them than any other group. Assuming that all other predictor variables are held equal, being in CRPR 4.3 has the greatest effect of any CRPR (negative binomial regression, $z = 5.545$, $df = 6161$, $p = 2.95 \times 10^{-8}$, $mean = 196.82759$).

3.4. DISCUSSION

Here we make the first attempt to quantify the amount of research effort made for studying rare plants in California. We measured this by getting data on publishing effort in literature and in sequence data. Our results suggest that rare plants are studied less than both non-rare natives and species weedy to the states. These differences are

reinforced at the Rare plant ranking levels too, with rarer plants studied the least. Despite this discrepancy, these results should come as a good sign. Nearly all of the 5,000 native species in California are represented in some way in the knowledge base. While one or two pieces of literature is not enough for entire management decisions to be made off of, it provides an incredibly helpful base for future research and conservation plans. Coupled with Californias robust network of tools used to identify (Jepson flora Project 2022) and locate (California Natural Diversity Database) rare plants, the barriers to making beneficial conservation decisions have never been lower.

While our data illuminates the differences between research effort, we cannot draw conclusions as to *why* rare plants are studied the least. Studying rare species is difficult for a variety of reasons. They are often in hard-to-reach areas and difficult to detect (Lesica, Yurkewycz and Crone 2006). Additionally, it can be bureaucratically difficult to obtain the proper permits and gain adequate access to study high-risk species. On top of logistical reasons for the discrepancy in research, there are likely other reasons for the differences. For example, it's likely that different species are represented differently based on research topic (i.e. ecology, evolution, etc.) Additionally, inequity likely exists based on flower color, habit, or any other morphological character (Yang *et al.* 2021; Mammides 2019). Whether or not these preferable characteristics all coalesce in weedy species is unknown, but it is now undeniable that weedy species receive the most research attention.

Part of this discrepancy is likely explained by the generalist nature of weeds; there isn't a one-to-one relationship between weeds and rare plants. One species of weed can often affect many native species (Mangla and Callaway 2008), and doing research on

weeds is a more cost-effective way in studying threats to rare plant populations. This strategy comes at great cost to our knowledge on impacted species, as one of the major shortcomings of literature on rare species is the lack of demographic or life-history information (Schemske *et al.* 1994), which is of prime importance when considering a listing status. With such basic information not widely available, it seems obvious where we, as researchers, should begin to pivot our attention.

Our research is the first broad-scale approach at understating patterns in publication among plant species in California. We can decisively say that rare plants receive less research attention than non-rare natives and weedy species. It is only with regular assessments of our knowledge that we can reorient research attention elsewhere, and reevaluate where research attention is best spent.

BIBLIOGRAPHY

- Adams, S. N., Jennings, S., & Warnock, N. (2020). Plant invasion depresses native species richness, but control of invasive species does little to restore it. *Plant Ecology & Diversity*, *13*(3–4), 257–266. <https://doi.org/10.1080/17550874.2020.1817998>
- Amundson, R., Guo, Y., & Gong, P. (n.d.). *Soil Diversity and Land Use in the United States*. 13.
- Baldwin, B. G., Thornhill, A. H., Freyman, W. A., Ackerly, D. D., Kling, M. M., Morueta-Holme, N., & Mishler, B. D. (2017). Species richness and endemism in the native flora of California. *American Journal of Botany*, *104*(3), 487–501. <https://doi.org/10.3732/ajb.1600326>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7*.
- Bayliss, H., Stewart, G., Wilcox, A., & Randall, N. (2013). A perceived gap between invasive species research and stakeholder priorities. *NeoBiota*, *19*, 67–82. <https://doi.org/10.3897/neobiota.19.4897>
- Bellard, C., Leclerc, C., Leroy, B., Bakkenes, M., Veloz, S., Thuiller, W., & Courchamp, F. (2014). Vulnerability of biodiversity hotspots to global change. *Global Ecology and Biogeography*, *23*(12), 1376–1386. <https://doi.org/10.1111/geb.12228>
- Bevill, R. L., & Louda, S. M. (1999). Comparisons of Related Rare and Common Species in the Study of Plant Rarity. *Conservation Biology*, *13*(3), 493–498. <https://doi.org/10.1046/j.1523-1739.1999.97369.x>
- Blossey, B. (n.d.). *Before, During and After: The Need for Long-term Monitoring in Invasive Plant Species Management*. 11.
- Bossard, C. C., & Randall, J. M. (2007). Non-native plants of California. *Terrestrial Vegetation of California. 3rd Edition. University of California Press, Berkeley, CA*, 107–123.
- Brooks, M. L., D’Antonio, C. M., Richardson, D. M., Grace, J. B., Keeley, J. E., DiTOMASO, J. M., Hobbs, R. J., Pellant, M., & Pyke, D. (2004). Effects of Invasive Alien Plants on Fire Regimes. *BioScience*, *54*(7), 677. [https://doi.org/10.1641/0006-3568\(2004\)054\[0677:EOIAP0\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0677:EOIAP0]2.0.CO;2)
- California Native Plant Society, Rare Plant Program. 2022. *Rare Plant Inventory (online edition, v9-01 1.5)*. Website <https://www.rareplants.cnps.org>.

- Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences*, 114(30). <https://doi.org/10.1073/pnas.1704949114>
- Corlett, R. T., & Westcott, D. A. (2013). Will plant movements keep up with climate change? *Trends in Ecology & Evolution*, 28(8), 482–488. <https://doi.org/10.1016/j.tree.2013.04.003>
- Cowling, R. M., Rundel, P. W., Lamont, B. B., Kalin Arroyo, M., & Arianoutsou, M. (1996). Plant diversity in mediterranean-climate regions. *Trends in Ecology & Evolution*, 11(9), 362–366. [https://doi.org/10.1016/0169-5347\(96\)10044-6](https://doi.org/10.1016/0169-5347(96)10044-6)
- Daubenmire, R. (1978). *Plant Geography*. Academic Press.
- Drohan, P. J., & Farnham, T. J. (2006). Protecting Life's Foundation. *SOIL SCI. SOC. AM. J.*, 70, 11.
- Dueñas, M.-A., Hemming, D. J., Roberts, A., & Diaz-Soltero, H. (2021). The threat of invasive species to IUCN-listed critically endangered species: A systematic review. *Global Ecology and Conservation*, 26, e01476. <https://doi.org/10.1016/j.gecco.2021.e01476>
- Ehrenfeld, J. G. (2003). Effects of Exotic Plant Invasions on Soil Nutrient Cycling Processes. *Ecosystems*, 6(6), 503–523. <https://doi.org/10.1007/s10021-002-0151-3>
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. *Diversity and Distributions*, 17(1), 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Ellstrand, N. C., & work(s);, D. R. E. R. (1993). Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annual Review of Ecology and Systematics*, 2, 217–242.
- Environmental Consulting in California—Market Research Report*. (2019, February 28). IBISWorld. <https://www.ibisworld.com/us/industry/california/environmental-consulting/11158/>
- Facelli, J. M., & Pickett, S. T. A. (1991). Plant litter: Its dynamics and effects on plant community structure. *The Botanical Review*, 57(1), 1–32. <https://doi.org/10.1007/BF02858763>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*,

24(1), 38–49. <https://doi.org/10.1017/S0376892997000088>

Fish, C. D. of & Game. (2022). *Special vascular plants, bryophytes, and lichens list*. Calif. Dept. Fish Game Sacramento.

Fitzgerald, D. B., Smith, D. R., Culver, D. C., Feller, D., Fong, D. W., Hajenga, J., Niemiller, M. L., Nolfi, D. C., Orndorff, W. D., Douglas, B., Maloney, K. O., & Young, J. A. (2021). Using expert knowledge to support Endangered Species Act decision-making for data-deficient species. *Conservation Biology*, 35(5), 12.

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., & Bolker, B. (2019).

Car: Companion to Applied Regression. R package version 3.0-3. *Website* [https://CRAN.R-Project.Org/Package= Car](https://CRAN.R-Project.Org/Package=Car) [Accessed 17 March 2020].

Franklin, J. (2013). Species distribution models in conservation biogeography: Developments and challenges. *Diversity and Distributions*, 19(10), 1217–1223. <https://doi.org/10.1111/ddi.12125>

Gaertner, M., Den Breeyen, A., Cang Hui, & Richardson, D. M. (2009). Impacts of alien plant invasions on species richness in Mediterranean-type ecosystems: A meta analysis. *Progress in Physical Geography: Earth and Environment*, 33(3), 319–338. <https://doi.org/10.1177/0309133309341607>

Goff, F. G., Dawson, G. A., & Rochow, J. J. (1982). Site examination for threatened and endangered plant species. *Environmental Management*, 6(4), 307–316. <https://doi.org/10.1007/BF01875062>

Goldblatt, P., & Manning, J. C. (2002). Plant diversity of the cape region of Southern Africa. *Annals of the Missouri Botanical Garden*, 89(2), 281–302.

Grime, J. P., Mason, G., Curtis, A. V., Rodman, J., & Band, S. R. (1981). A comparative study of germination characteristics in a local flora. *The Journal of Ecology*, 1017–1059.

Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C., & Kueffer, C. (2014). Unifying niche shift studies: Insights from biological invasions. *Trends in Ecology & Evolution*, 29(5), 260–269. <https://doi.org/10.1016/j.tree.2014.02.009>

Gurevitch, J., & Padilla, D. (2004). Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution*, 19(9), 470–474. <https://doi.org/10.1016/j.tree.2004.07.005>

Hayes, G. F., & Holl, K. D. (2003). Cattle Grazing Impacts on Annual Forbs and Vegetation Composition of Mesic Grasslands in California. *Conservation Biology*,

17(6), 1694–1702. <https://doi.org/10.1111/j.1523-1739.2003.00281.x>

- He, F., & Gaston, K. J. (2000). Estimating species abundance from occurrence. *The American Naturalist*, 156(5), 553–559.
- Hijmans, R. J., Steven Phillips, John Leathwick, & Elith, J. (2017). Package “dismo.” <ftp://ftp.gr.xemacs.org/mirrors/CRAN/web/packages/dismo/dismo.pdf>
- Holland, V. L., Keil, D. J., & Schierenbeck, K. A. (1996). California Vegetation. *Madrono*, 43(1), 99–99.
- Holsinger, K. E., and L. D. Gottlieb. 1991. Conservation of rare and endangered plants: Principles and prospects. Pages 195-208 in D. A. Falk and K. E. Holsinger, editors. *Genetics and conservation of rare plants*. Oxford University Press, New York, New York, USA. (n.d.).
- Hoover, R. F. (1940). A MONOGRAPH OF THE GENUS CHLOROGALUM. *Madroño*, 5(5), 137–147.
- Huntsinger, L., Buttolph, L., & Hopkinson, P. (1997). Ownership and management changes on California hardwood rangelands: 1985 to 1992. *Rangeland Ecology & Management/Journal of Range Management Archives*, 50(4), 423–430.
- Hurlbert, A. H., & Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Sciences*, 104(33), 13384–13389.
- Kass, J. M., Anderson, R. P., Espinosa-Lucas, A., Juárez-Jaimes, V., Martínez-Salas, E., Botello, F., Tavera, G., Flores-Martínez, J. J., & Sánchez-Cordero, V. (2020). Biotic predictors with phenological information improve range estimates for migrating monarch butterflies in Mexico. *Ecography*, 43(3), 341–352.
- Keeley, J. E. (1988). Bibliographies on chaparral and the fire ecology of other Mediterranean systems. *Report-California Water Resources Center, University of California (USA)*.
- Keeley, J. E. (2002). Native American impacts on fire regimes of the California coastal ranges. *Journal of Biogeography*, 29(3), 303–320. <https://doi.org/10.1046/j.1365-2699.2002.00676.x>
- Kettenring, K. M., & Adams, C. R. (2011). Lessons learned from invasive plant control experiments: A systematic review and meta-analysis: Invasive plant control experiments. *Journal of Applied Ecology*, 48(4), 970–979. <https://doi.org/10.1111/j.1365-2664.2011.01979.x>

- Knapp, P. A. (1992). Secondary plant succession and vegetation recovery in two western Great Basin Desert ghost towns. *Biological Conservation*, 60(2), 81–89. [https://doi.org/10.1016/0006-3207\(92\)91158-O](https://doi.org/10.1016/0006-3207(92)91158-O)
- Kofron, C. P., Rutherford, C., Magney, D. L., Borchert, M., & Simpson, L. G. (2021). Camatta Canyon amole *Hooveria purpurea* var. *reducta* (Agavaceae): A Threatened Plant in La Panza Range, San Luis Obispo County, California. *Bulletin, Southern California Academy of Sciences*, 120(1). <https://doi.org/10.3160/0038-3872-120.1.26>
- Lacher, T. E., & Boitani, L. (2012). The IUCN global assessments: Partnerships, collaboration and data sharing for biodiversity science and policy. *Conservation Letters*, 7.
- Lande, R. (2022). *Genetics and Demography in Biological Conservation*. 241, 7.
- Leimu, R., Mutikainen, P., Koricheva, J., & Fischer, M. (2006). How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology*, 94(5), 942–952. <https://doi.org/10.1111/j.1365-2745.2006.01150.x>
- Lenz, T. I., Moyle-Croft, J. L., & Facelli, J. M. (2003). Direct and indirect effects of exotic annual grasses on species composition of a South Australian grassland. *Austral Ecology*, 28(1), 23–32. <https://doi.org/10.1046/j.1442-9993.2003.01238.x>
- Lesica, P., & Crone, E. E. (2007). Causes and consequences of prolonged dormancy for an iteroparous geophyte, *Silene spaldingii*. *Journal of Ecology*, 95(6), 1360–1369. <https://doi.org/10.1111/j.1365-2745.2007.01291.x>
- Lesica, P., Yurkewycz, R., & Crone, E. E. (2006). Rare plants are common where you find them. *American Journal of Botany*, 93(3), 454–459. <https://doi.org/10.3732/ajb.93.3.454>
- Levine, J. M., Vilà, M., Antonio, C. M. D., Dukes, J. S., Grigulis, K., & Lavorel, S. (2003). Mechanisms underlying the impacts of exotic plant invasions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1517), 775–781. <https://doi.org/10.1098/rspb.2003.2327>
- Lionello, P., Malanotte-Rizzoli, P., Boscolo, R., Alpert, P., Artale, V., Li, L., Luterbacher, J., May, W., Trigo, R., & Tsimplis, M. (2006). The Mediterranean climate: An overview of the main characteristics and issues. In *Developments in earth and environmental sciences* (Vol. 4, pp. 1–26). Elsevier.
- Lovett-Doust, J., Biernacki, M., Page, R., Chan, M., Natgunarajah, R., & Timis, G. (n.d.). *Effects of land ownership and landscape-level factors on rare-species richness in natural areas of southern Ontario, Canada*. 14.

- Lyons, K. G., Brigham, C. A., Traut, B. H., & Schwartz, M. W. (2005). Rare Species and Ecosystem Functioning. *Conservation Biology*, 19(4), 1019–1024. <https://doi.org/10.1111/j.1523-1739.2005.00106.x>
- Macdonald, I. A. W., Graber, D. M., DeBenedetti, S., Groves, R. H., & Fuentes, E. R. (1988). Introduced species in nature reserves in Mediterranean-type climatic regions of the world. *Biological Conservation*, 44(1–2), 37–66. [https://doi.org/10.1016/0006-3207\(88\)90004-3](https://doi.org/10.1016/0006-3207(88)90004-3)
- Magney, D. (1988). Results of second year population dynamics of the state-listed rare *Camatta amole* (*Chlorogalum purpureum* var. *Reductum*). *Report to Los Padres National Forest, Goleta, Calif.*
- Mangla, S., & Callaway, R. M. (2007). Exotic invasive plant accumulates native soil pathogens which inhibit native plants. *Journal of Ecology*, 0(0), 071031082432001-??? <https://doi.org/10.1111/j.1365-2745.2007.01312.x>
- McKenzie, D., & Littell, J. S. (2011). *Climate Change and Wilderness Fire Regimes*. 8.
- McKinney, W. & others. (2011). pandas: A foundational Python library for data analysis and statistics. *Python for High Performance and Scientific Computing*, 14(9), 1–9.
- Merenlender, A. M., Huntsinger, L., Guthey, G., & Fairfax, S. K. (2004). Land Trusts and Conservation Easements: Who Is Conserving What for Whom? *Conservation Biology*, 18(1), 65–76. <https://doi.org/10.1111/j.1523-1739.2004.00401.x>
- Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: The critical role of hotspots. In *Biodiversity hotspots* (pp. 3–22). Springer.
- Mordecai, E. A. (2012). Soil moisture and fungi affect seed survival in California grassland annual plants. *PLoS One*, 7(6), e39083.
- Okuyama, Y. (n.d.). *Which Genus to Study? In Search of Plant Genera Underrepresented or Overrepresented in the Research from the Flora of Japan*. 9.
- Parker, S. (2012). Small Reserves Can Successfully Preserve Rare Plants Despite Management Challenges. *Natural Areas Journal*, 32(4), 403–411. <https://doi.org/10.3375/043.032.0409>
- Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *421*, 6.
- Pereira, H. M., & Cooper, H. D. (2006). *Towards the global monitoring of biodiversity*

change. 7.

- Peterson, A. T., & Soberón, J. (2012). Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right. *Natureza & Conservação*, 10(2), 102–107. <https://doi.org/10.4322/natcon.2012.019>
- Planet Team (2017). *Planet Application Program Interface: In Space for Life on Earth*. San Francisco, CA. (n.d.). [Map].
- Price, M. (1994). Ecopolitics and environmental nongovernmental organizations in Latin America. *Geographical Review*, 42–58.
- PROCHEŞ, Ş., Cowling, R. M., Goldblatt, P., Manning, J. C., & Snijman, D. A. (2006). An overview of the Cape geophytes. *Biological Journal of the Linnean Society*, 87(1), 27–43.
- Pyšek, P., Jarošík, V., Hulme, P. E., Pergl, J., Hejda, M., Schaffner, U., & Vilà, M. (2012). A global assessment of invasive plant impacts on resident species, communities and ecosystems: The interaction of impact measures, invading species 'traits and environment. *Global Change Biology*, 18(5), 1725–1737. <https://doi.org/10.1111/j.1365-2486.2011.02636.x>
- Radosavljevic, A., & Anderson, R. P. (2014). Making better Maxent models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography*, 41(4), 629–643. <https://doi.org/10.1111/jbi.12227>
- Raes, N., & ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, 30(5), 727–736. <https://doi.org/10.1111/j.2007.0906-7590.05041.x>
- Raven, P. H., & Axelrod, D. I. (1978). *Origin and Relationships of the California Flora*. University of California Press.
- Reiner, R., & Craig, A. (2011). Conservation Easements in California Blue Oak Woodlands: Testing the Assumption of Livestock Grazing as a Compatible use. *Natural Areas Journal*, 31(4), 408–413. <https://doi.org/10.3375/043.031.0411>
- Richardson, L. (2007). Beautiful soup documentation. *April*.
- Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'mass.' *Cran r*, 538, 113–120.
- Safford, H. D. (2011). Serpentine endemism in the California flora. *Fremontia*, 32.
- Safford, H., & Miller, J. E. D. (2020). AN UPDATED DATABASE OF SERPENTINE

ENDEMISM IN THE CALIFORNIA FLORA. *Madroño*, 67(2).
<https://doi.org/10.3120/0024-9637-67.2.85>

- Safford, H., Viers, J., & Harrison, S. (2005). Serpentine endemism in the California flora: A database of serpentine affinity. *Madroño*, 52(4), 222–257.
- Schatz, G. E. (2009). Plants on the IUCN Red List: Setting priorities to inform conservation. *Trends in Plant Science*, 14(11), 638–642.
- Schemske, D. W., Husband, B. C., Ruckelshaus, M. H., Goodwillie, C., Parker, I. M., & Bishop, J. G. (1994). Evaluating Approaches to the Conservation of Rare and Endangered Plants. *Ecology*, 75(3), 584–606. <https://doi.org/10.2307/1941718>
- Seabloom, E. W., Borer, E. T., Boucher, V. L., Burton, R. S., Cottingham, K. L., Goldwasser, L., Gram, W. K., Kendall, B. E., & Micheli, F. (2003). COMPETITION, SEED LIMITATION, DISTURBANCE, AND REESTABLISHMENT OF CALIFORNIA NATIVE ANNUAL FORBS. *Ecological Applications*, 13(3), 575–592.
[https://doi.org/10.1890/1051-0761\(2003\)013\[0575:CSLDAR\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2003)013[0575:CSLDAR]2.0.CO;2)
- Shaffer, M. L. (1981). Minimum Population Sizes for Species Conservation. *BioScience*, 31(2), 131–134. <https://doi.org/10.2307/1308256>
- Shefferson, R. P., Kull, T., Hutchings, M. J., Selosse, M.-A., Jacquemyn, H., Kellett, K. M., Menges, E. S., Primack, R. B., Tuomi, J., & Alahuhta, K. (2018). Drivers of vegetative dormancy across herbaceous perennial plant species. *Ecology Letters*, 21(5), 724–733.
- Shefferson, R. P., Kull, T., & Tali, K. (2005). ADULT WHOLE-PLANT DORMANCY INDUCED BY STRESS IN LONG-LIVED ORCHIDS. *Ecology*, 86(11), 3099–3104. <https://doi.org/10.1890/05-0586>
- Skinner, M. W., & Pavlik, B. M. (1994). *California Native Plant Society's inventory of rare and endangered vascular plants of California*. California Native Plant Society.
- Stanley, A. G., Kaye, T. N., & Dunwiddie, P. W. (2011). Multiple Treatment Combinations and Seed Addition Increase Abundance and Diversity of Native Plants in Pacific Northwest Prairies. *Ecological Restoration*, 29(1–2), 35–44. <https://doi.org/10.3368/er.29.1-2.35>
- Stehman, S. V., & Salzer, D. W. (2000). ESTIMATING DENSITY FROM SURVEYS EMPLOYING UNEQUAL-AREA BELT TRANSECTS. *Wetlands*, 20(3), 512–519. [https://doi.org/10.1672/0277-5212\(2000\)020<0512:EDFSEU>2.0.CO;2](https://doi.org/10.1672/0277-5212(2000)020<0512:EDFSEU>2.0.CO;2)
- Strayer, D. L. (2012). Eight questions about invasions and ecosystem functioning.

- Ecology Letters*, 15(10), 1199–1210.
<https://doi.org/10.1111/j.1461-0248.2012.01817.x>
- Strayer, D. L., Eviner, V. T., Jeschke, J. M., & Pace, M. L. (2006). Understanding the long-term effects of species invasions. *Trends in Ecology & Evolution*, 21(11), 645–651. <https://doi.org/10.1016/j.tree.2006.07.007>
- Tatarenko, I. (2019). Having a break: Prolonged dormancy observed in a rare species, *Fritillaria meleagris*. *SOCIALNO-ECOLOGICHESKIE TECHNOLOGII*, 9(3), 302–324. <https://doi.org/10.31862/2500-2961-2019-9-3-302-324>
- Taylor, D. W., & Keil, D. J. (2018). *HOOVERIA, A NEW GENUS LIBERATED FROM CHLOROGALUM (AGAVACEAE SUBF. CHLOROGALOIDEAE)*. 6.
- Team, R. C. (n.d.). *Version 4.1. 2. R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria; 2021.
- Uematsu, Y., Koga, T., Mitsuhashi, H., & Ushimaru, A. (2010). Abandonment and intensified use of agricultural land decrease habitats of rare herbs in semi-natural grasslands. 6.
- U.S. Fish and Wildlife Service. (2020). *Species Status Assessment for Purple Amole (Hooveria purpurea [Chlorogalum purpureum])*.
- U.S. Fish and Wildlife Service (USFWS). (2009a). *Listing a species as threatened or endangered: Section 4 of the Endangered Species Act*. USFWS, Washington, D.C.
- U.S. Geological Survey, 20211116, USGS 1/3 Arc Second n36w121 20210610: U.S. Geological Survey. (n.d.). [Map].
- Van Rossum, G., 2020. *The Python Library Reference, release 3.8. 2, Python Software Foundation*. (n.d.).
- Vitousek, P. M., & Walker, L. R. (1989). Biological Invasion by *Myrica Faya* in Hawai'i: Plant Demography, Nitrogen Fixation, Ecosystem Effects. *Ecological Monographs*, 59(3), 247–265. <https://doi.org/10.2307/1942601>
- Wamelink, G. W. W., Goedhart, P. W., & Frissel, J. Y. (2014). Why Some Plant Species Are Rare. *PLoS ONE*, 9(7), e102674. <https://doi.org/10.1371/journal.pone.0102674>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>
- Wiens, J. J. (2016). Climate-Related Local Extinctions Are Already Widespread among

Plant and Animal Species. *PLOS Biology*, 14(12), e2001104.
<https://doi.org/10.1371/journal.pbio.2001104>

Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A., & Losos, E. (2022). *Quantifying Threats to Imperiled Species in the United States*. 10.

Wolkovich, E. M., Bolger, D. T., & Cottingham, K. L. (2009). Invasive grass litter facilitates native shrubs through abiotic effects. *Journal of Vegetation Science*, 20(6), 1121–1132.

Worldwide, R. C. T. and C. (2019). *The R stats package*.

Yang, W., Liu, D., You, Q., Chen, B., Jian, M., Hu, Q., Cong, M., & Ma, K. (2021). Taxonomic bias in occurrence information of angiosperm species in China. *Science China Life Sciences*, 64(4), 584–592.
<https://doi.org/10.1007/s11427-020-1821-x>

APPENDIX

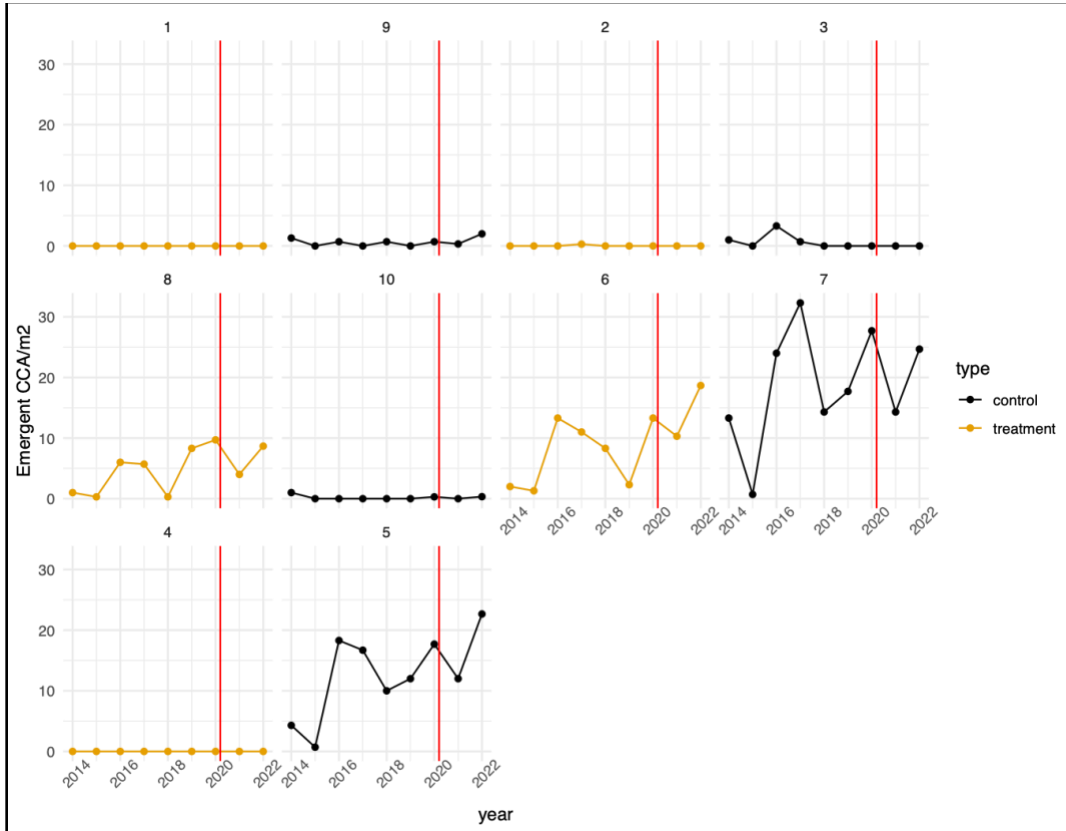


Figure 5: CCA in each plot over eight years of monitoring. The vertical red line indicates the year thatch removal began. Paired control and treatment plots are organized next to each other and are in the same rectangular border.

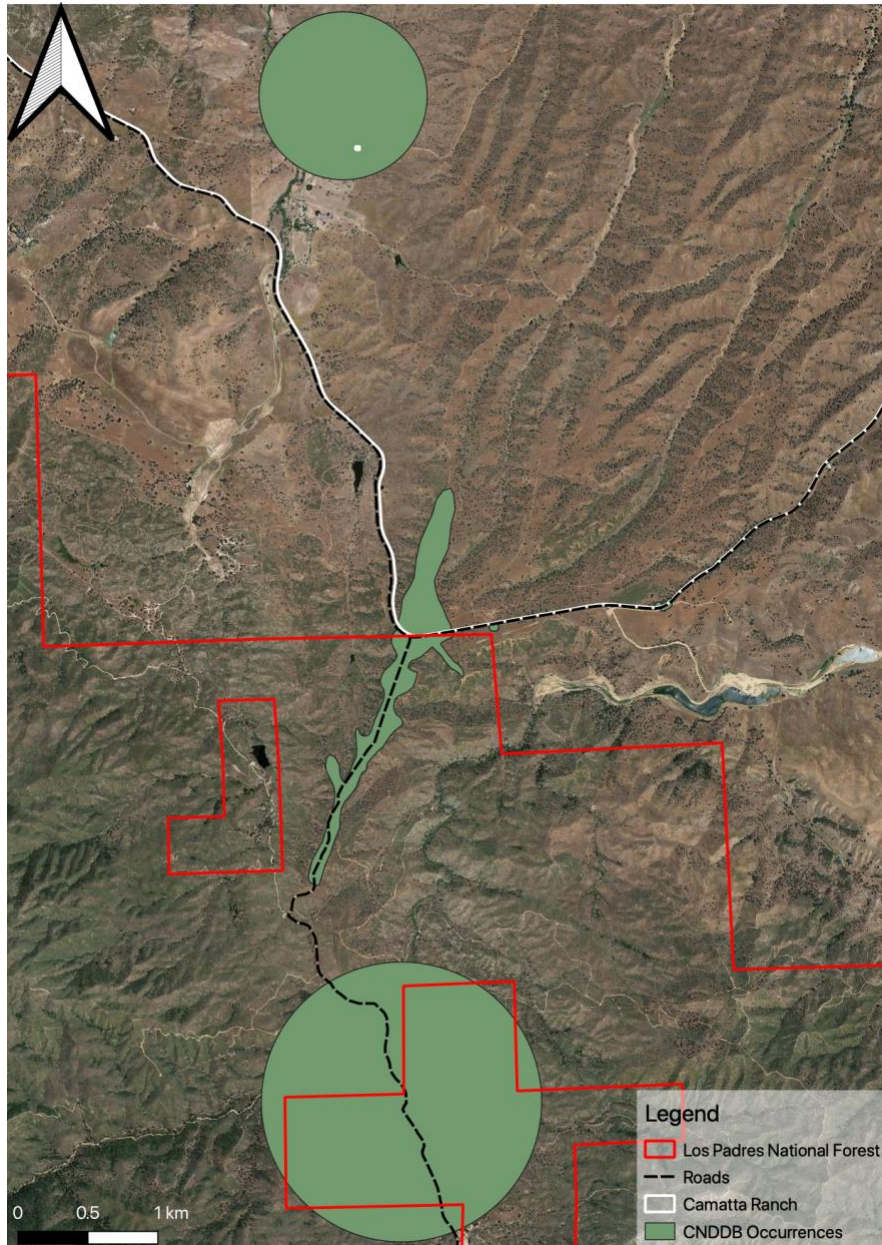


Figure 6: Location of *Hooveria purpurea* var. *redcuta* as mapped in CNDDDB. The long, skinny polygon is along Red Hill Road. The extent of the polygon that extends onto Camatta Ranch is speculative. Large circular polygons are buffered points.

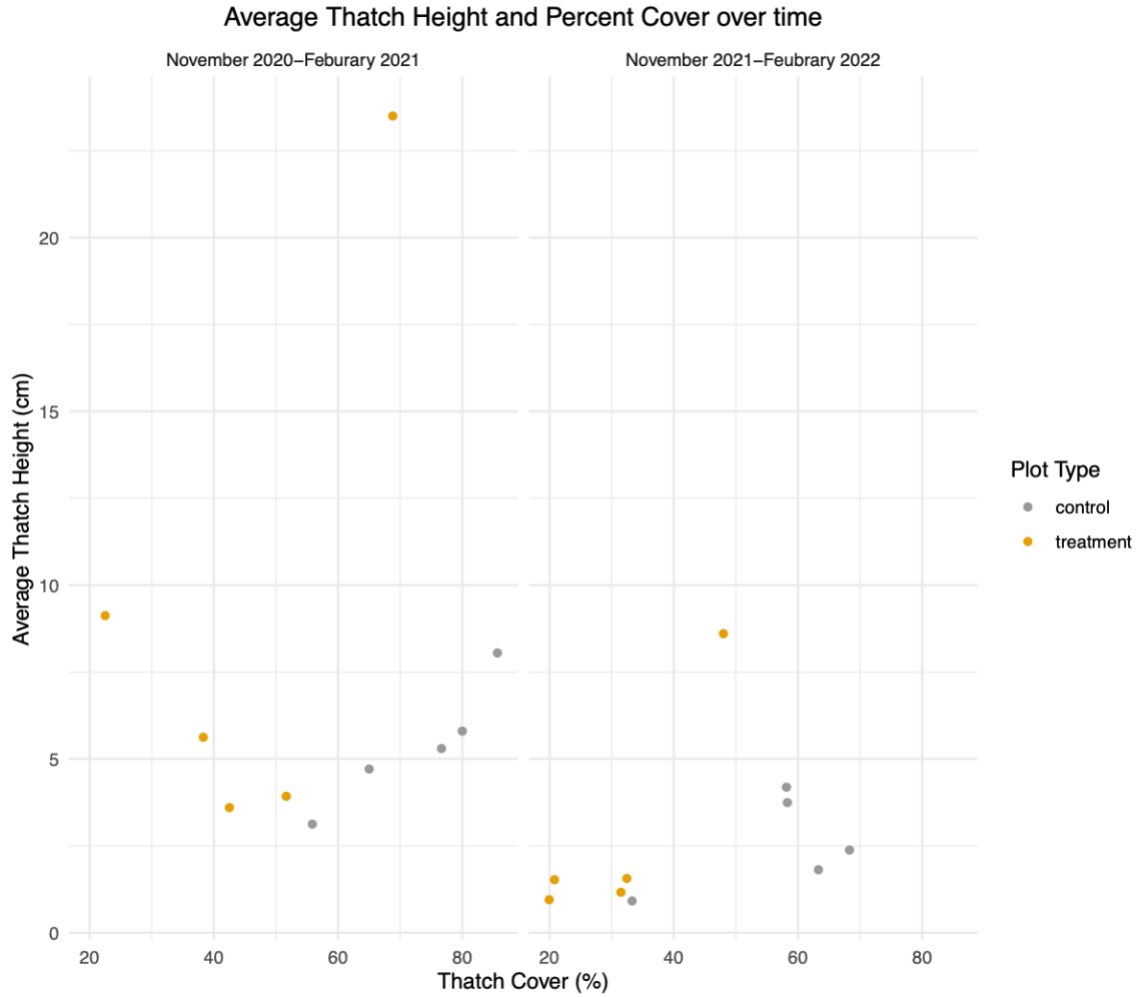


Figure 7: Thatch height and cover across three years. While thatch cover and height were evenly variable in the first year between plot types, they had segregated by 2022, where mean height is similar, but the mean cover between groups is different.

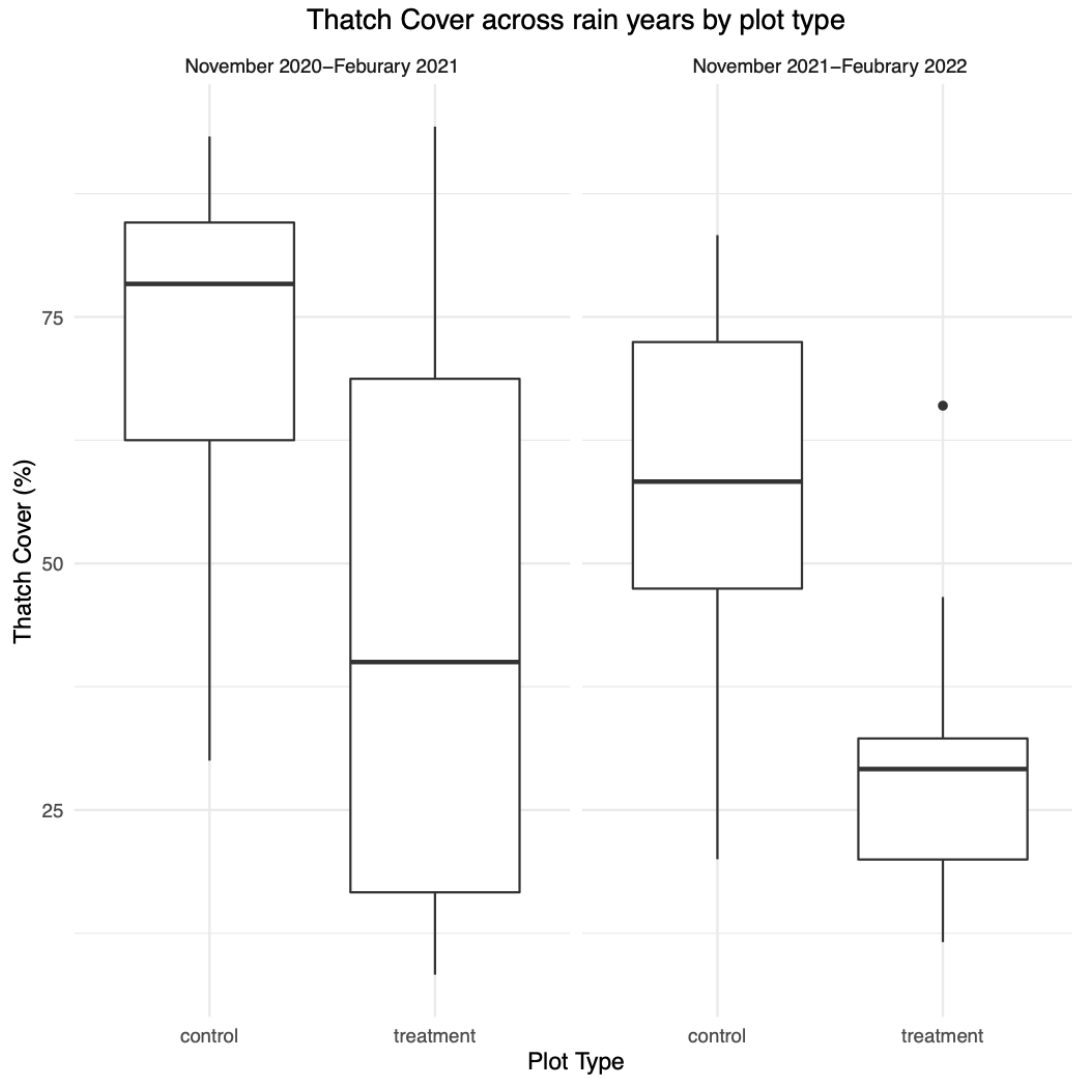


Figure 8: Thatch cover across rain years. Thatch cover decreased in both control and treatment plots between the two rain years.

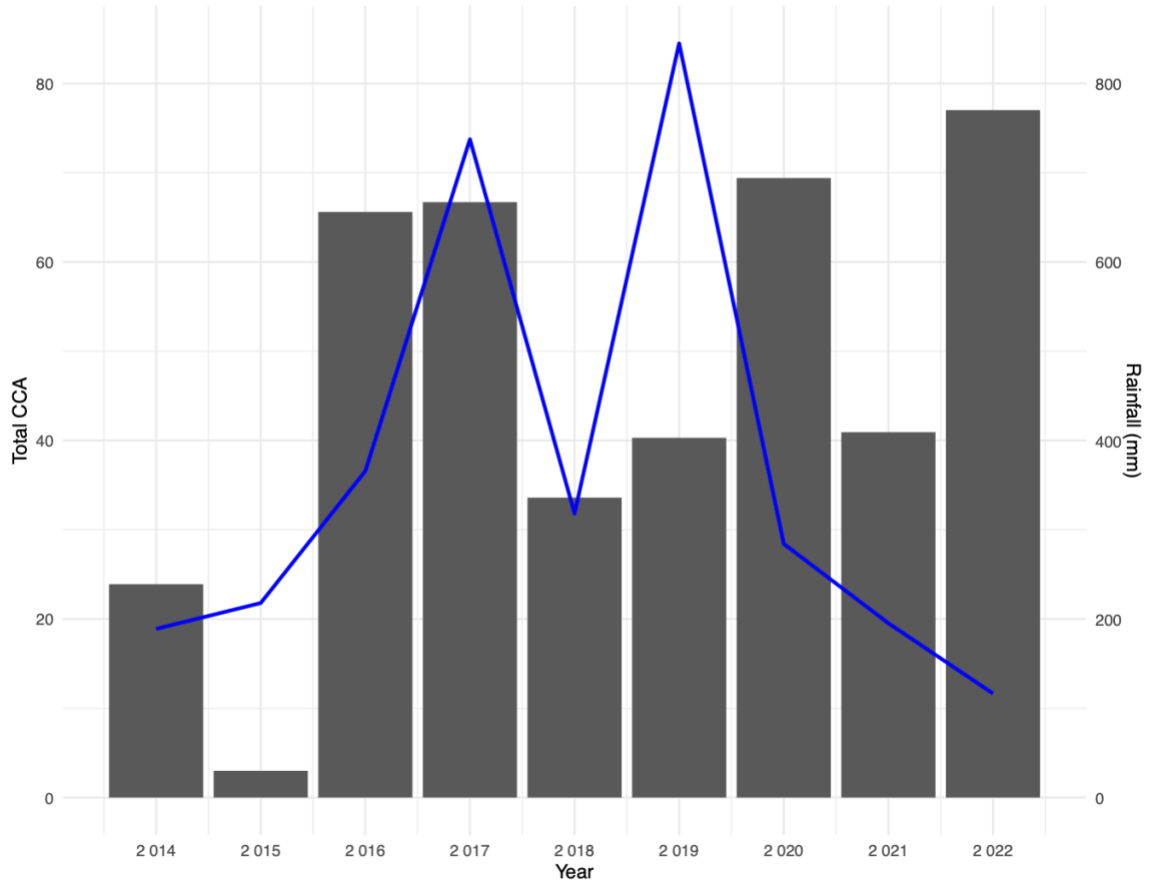


Figure 9: Total number of Camatta Canyon Amole counted in all 10 vegetation plots and rainfall (mm) for that rain year (blue line). For this data a rain year begins on November 1 and ends March 1. For example, between November 1, 2013-March 1, 2014, the weather station measured 188.97 mm of rain. Rainfall data is that which we believe would contribute to that growing season. Data from SLO-6 weather station on Shell Creek Road, 8 miles NW of of Red Hill Road.

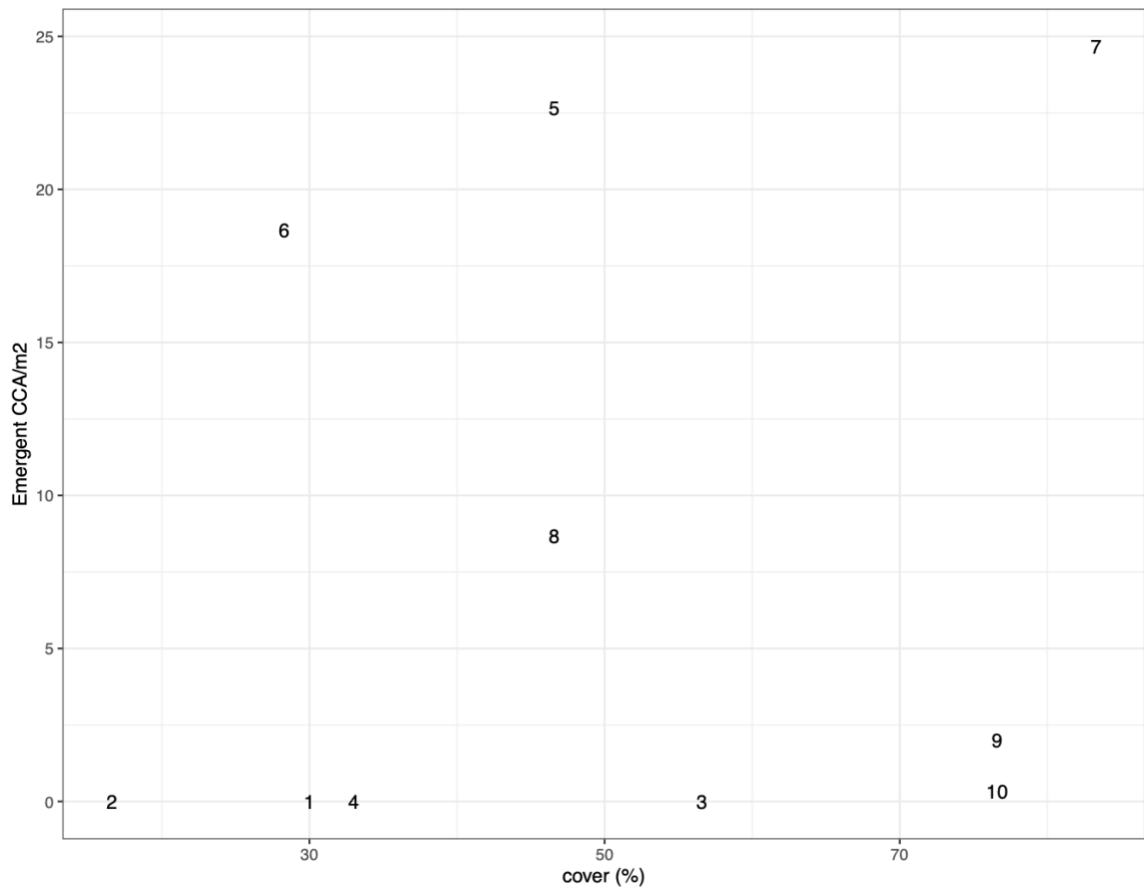


Figure 10: Percent cover and Emergent Amole of each plot at the final visit in February 2022 for the ten vegetation plots. Plot 7 had the highest counts for CCA while also having the highest percent cover. Plot 5 had equal cover to plot 8, but with more than twice the number of emergent CCA.

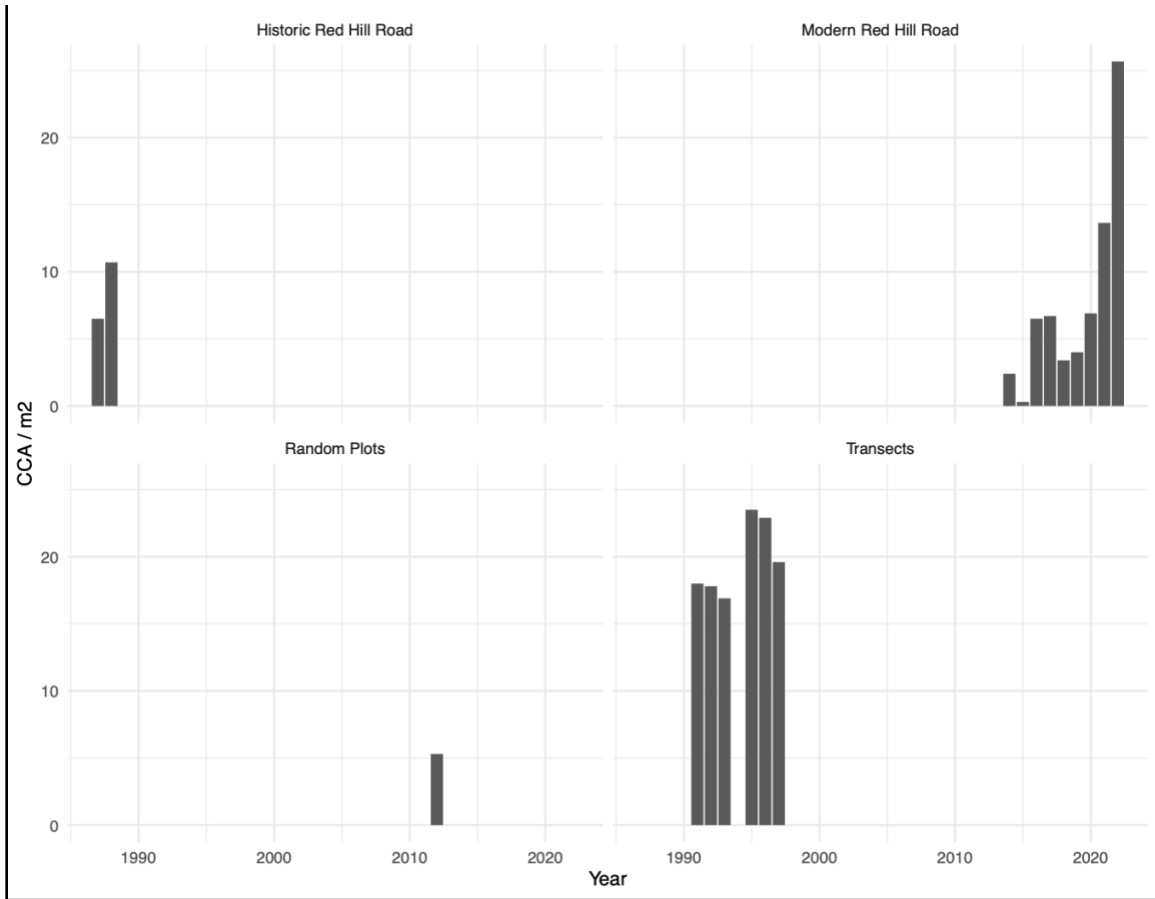


Figure 11: Total Camatta Canyon Amole from all survey efforts along Red Hill Road. "Historic Red Hill Road" are plots surveyed by David Magney in 1987 and 1988. "Modern Red Hill Road" plots are those reestablished by Kofron + Rutherford in 2014 in the approximate location of Historic Plots. Kofron et al. 2021 treats these two sets of plots as the same plots. Because the modern plots are in the approximate location of the historic plots, we treat them as separate. "Random" plots reflects a survey effort of 11 randomly selected plots along Red Hill Road. This happened just once. "Transects" reflect 11 belt transects that occurred over 6 years around Red Hill Road. These transects occurred along areas identified as the densest. All data is recovered from Kofron et al. 2021, Table 3.

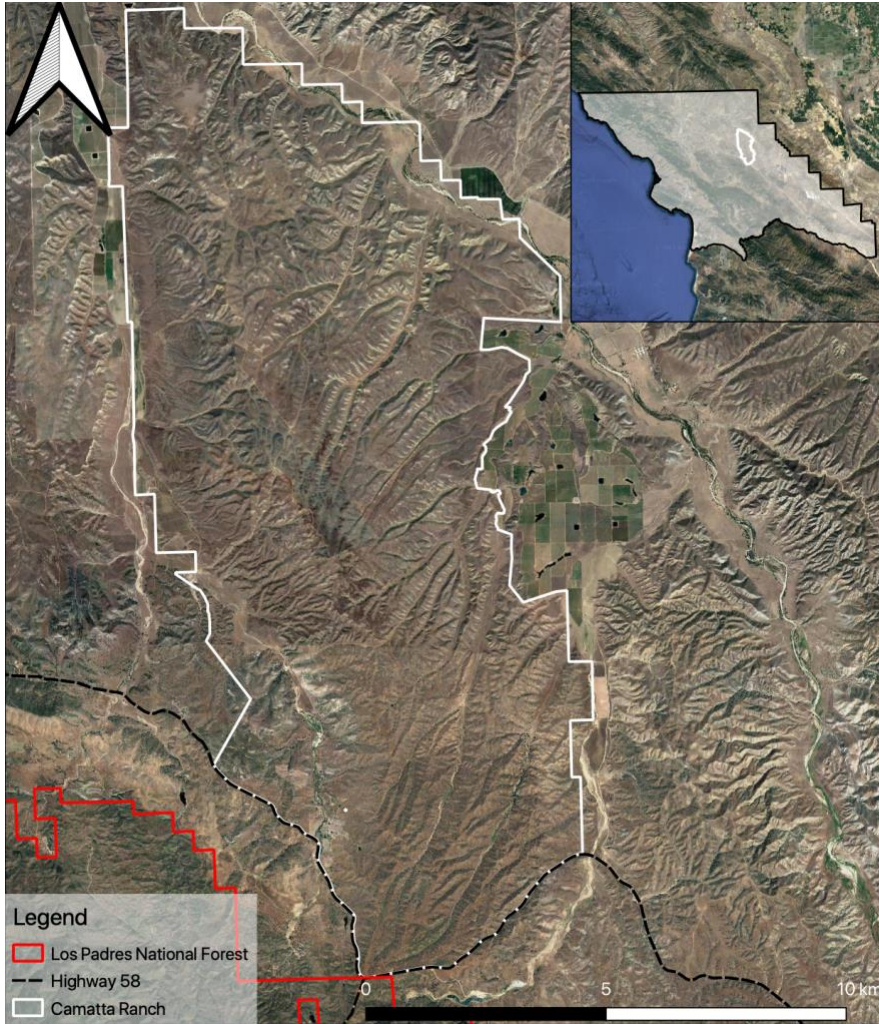


Figure 12: Map of Camatta Ranch. Located directly north of the Carissa Highway (HWY 58) near where Red Hill Road and Carissa Highway intersect. *Hooveria purpurea* var. *reducta* is only known to occur in the very southern portion of the ranch, close to Carissa Highway and Red Hill Road.

Table 2: the 15 ranch polygons that were surveyed in spring 2021 along with plot area, cca counts and relative density ratios. Polygons are arranged in order of smallest to largest density ratios.

Polygon	Belt transect area (m²)	CCA	CCA/m²
26	33	0	0
28	115	5	0.04347826
33	88	4	0.04545455
3	30	2	0.06666667
9	44	3	0.06818182
24	53	4	0.0754717
8	90	9	0.1
29	30	3	0.1
22	66	8	0.12121212
35	47	6	0.12765957
2	58	16	0.27586207
6	135	41	0.3037037
32	25	11	0.44
7	55	28	0.50909091
34	67	43	0.64179105
TOTAL	936	183	0.194

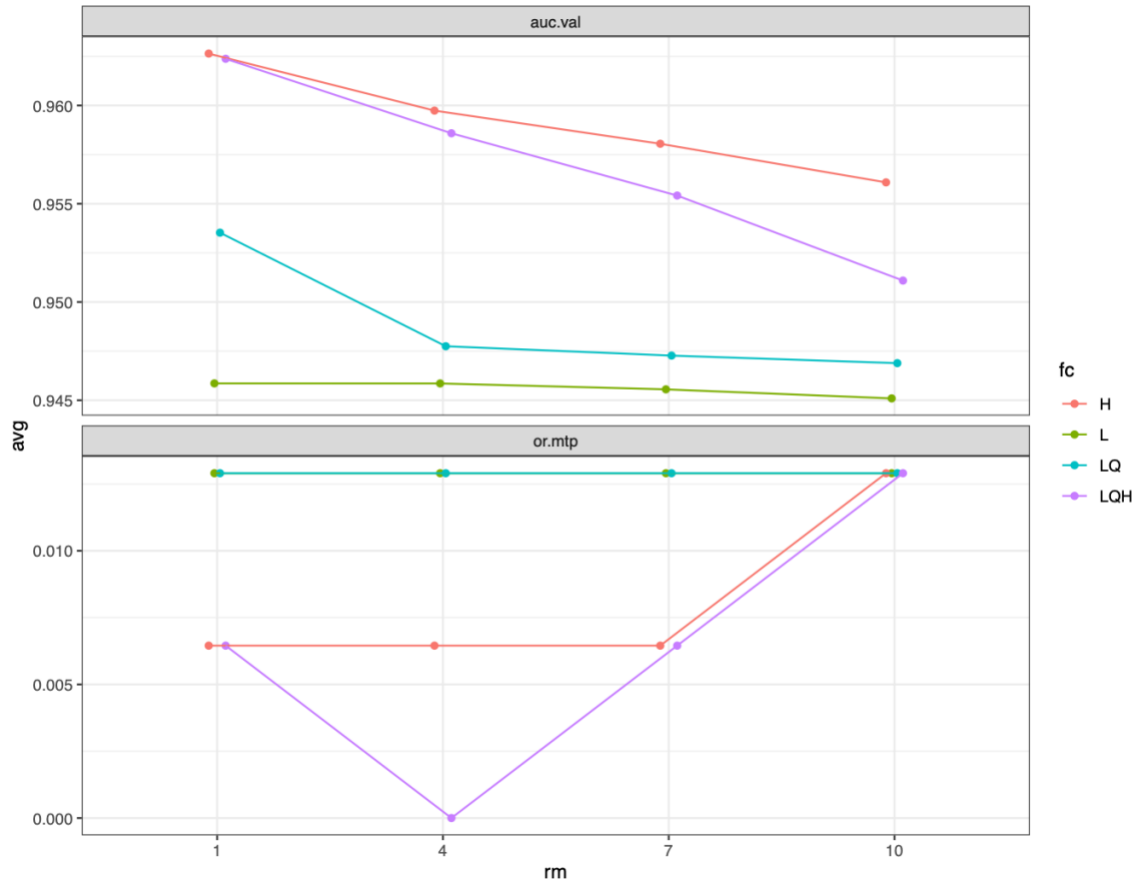


Figure 13: Results of the different models. The upper graph are the average validation AUC and below is the omission rates for the models. The x-axis are the four regularization multipliers, and the color points are the four feature classes. Feature class H and LQH had tied omission rates at rm 7, so our sequential test broke the tie for determining the optimal model by choosing from the highest AUC at that rm, which was fc H.

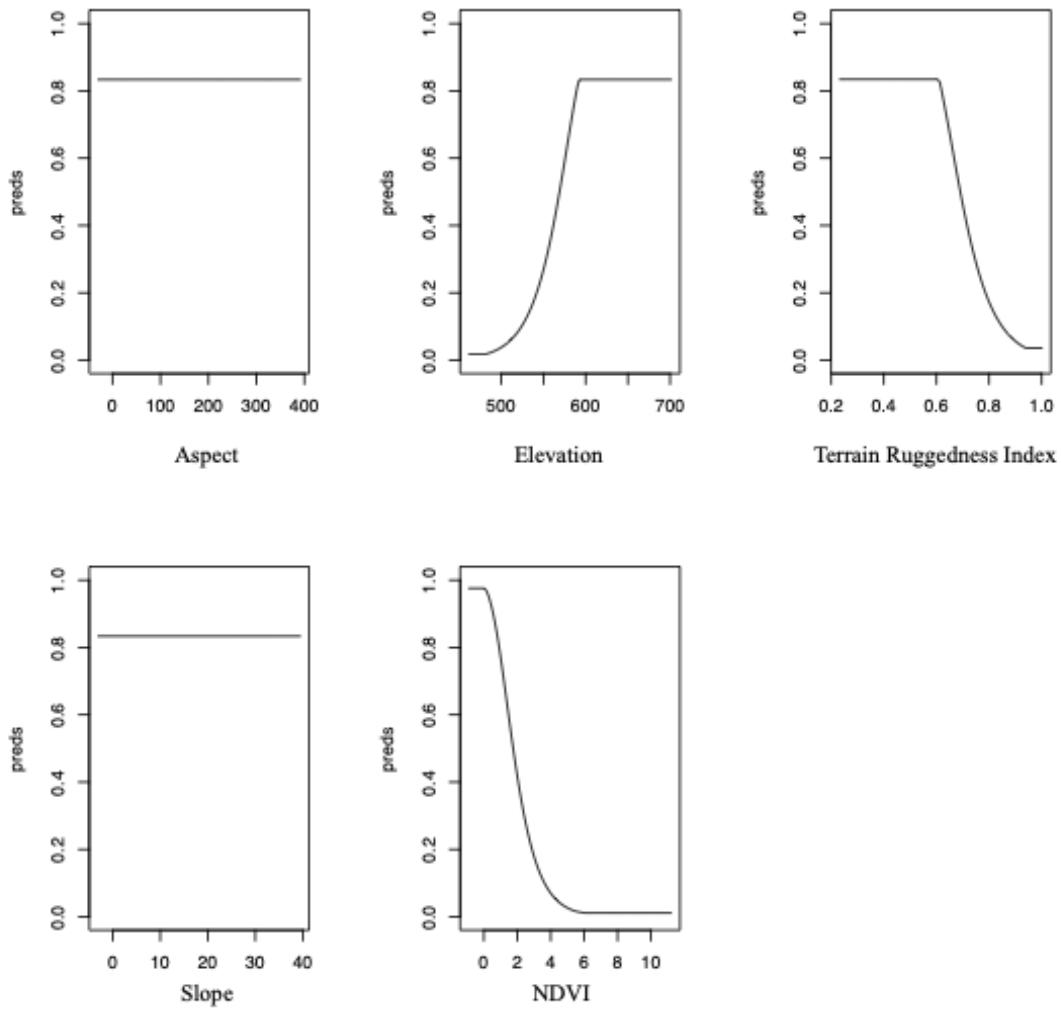


Figure 14: Response curves for optimal model. Aspect and slope had low predicting power for CCA, whereas TRI, elevation and NDVI were more strongly predictive of CCA presence.

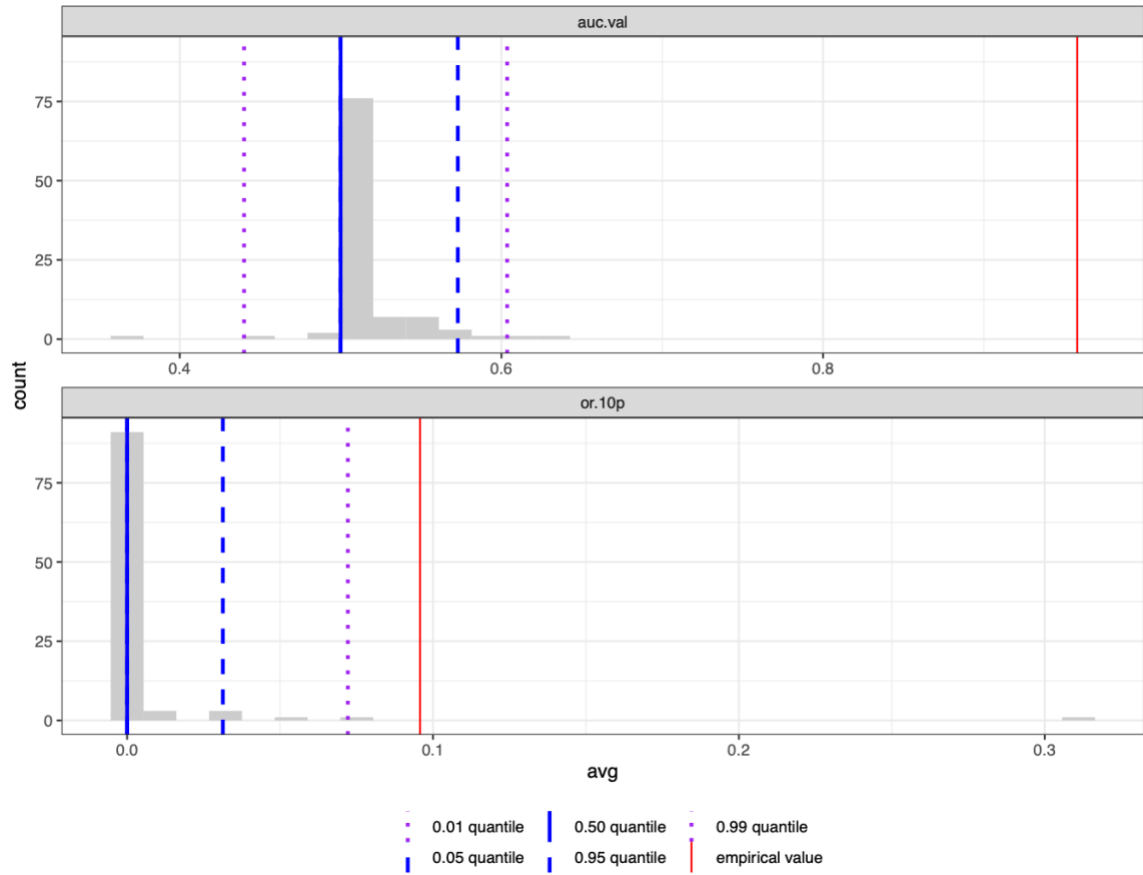


Figure 15: Results from 100 null models. Shows null model AUC and 10% omission rate results, along with the AUC and OR10 of our empirical model.

fc.H_rm.7

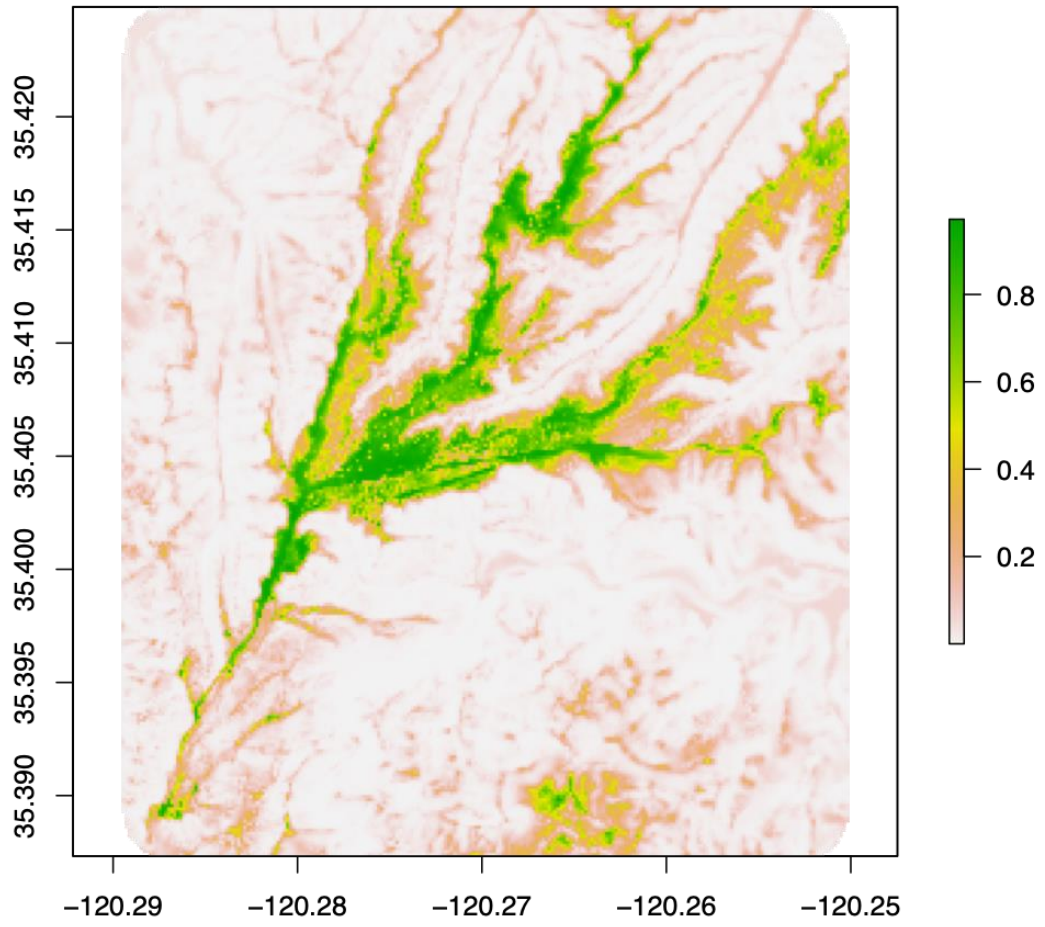


Figure 16: Raw prediction of the optimal model on Camatta Ranch and Red Hill Road.

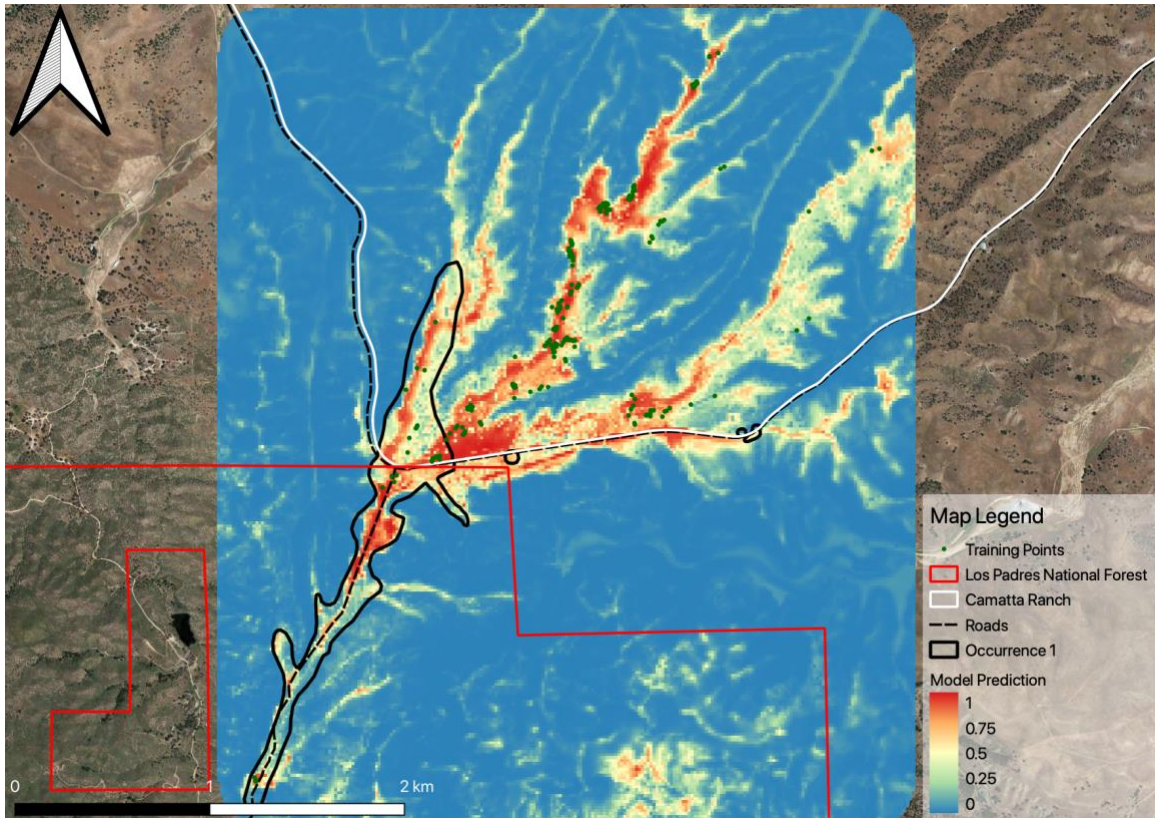


Figure 17: Species distribution model for CCA in Occurrence 1 (CNDDDB) and along the southern ridges of Camatta Ranch. Because our prediction generally predicts Occurrence 1 along Red Hill Road with a value of 0.5 or higher, we limited the theoretical distribution of the CCA on Camatta Ranch to areas with a value of 0.5 or greater for our population estimates.

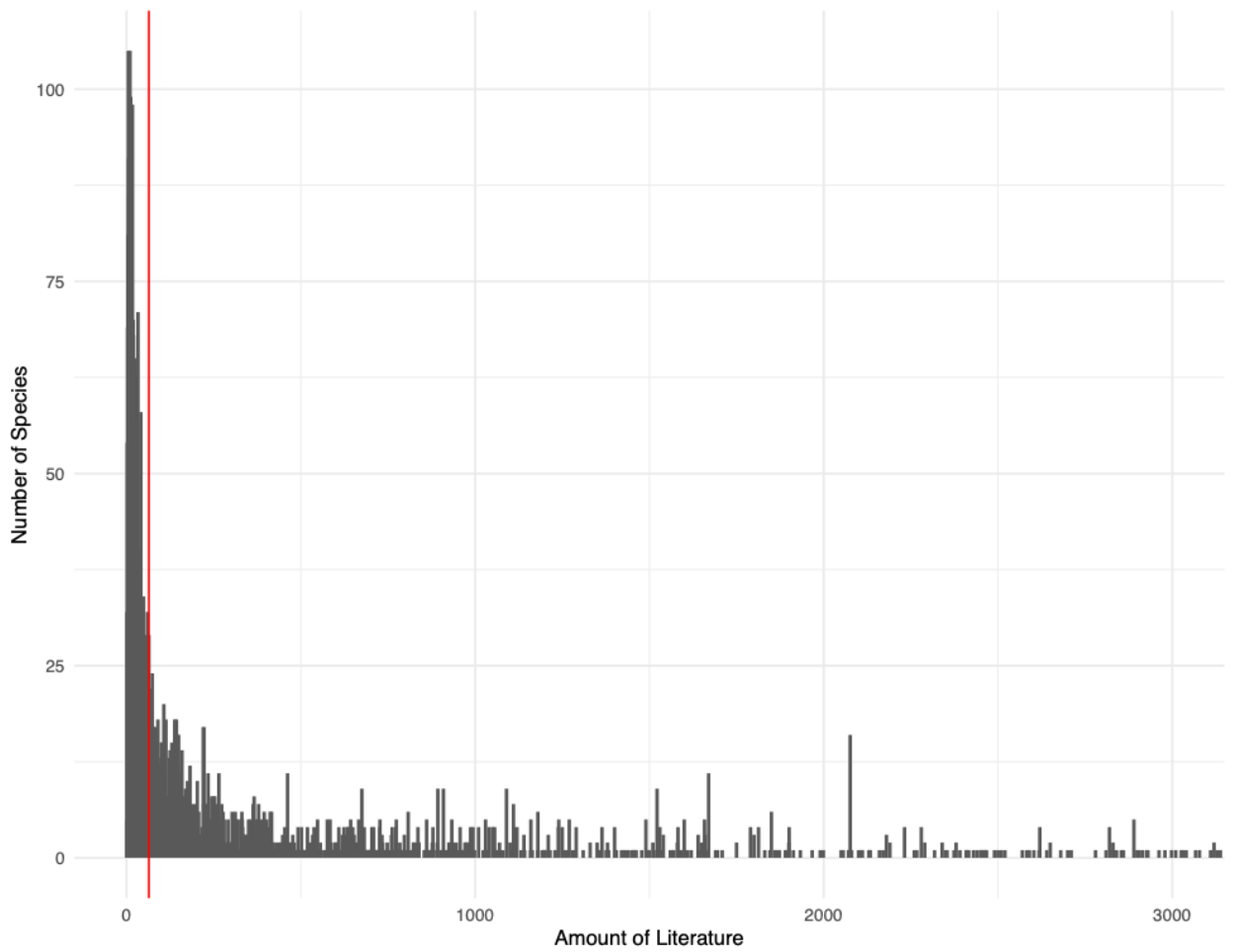


Figure 18: Histogram of literature counts of native plants of California. X-axis is limited to 3000 pieces of literature. The native species mentioned most in literature is *Helianthus annuus*, with 186,293 pieces of literature. The vertical red line separates half of the species in California.

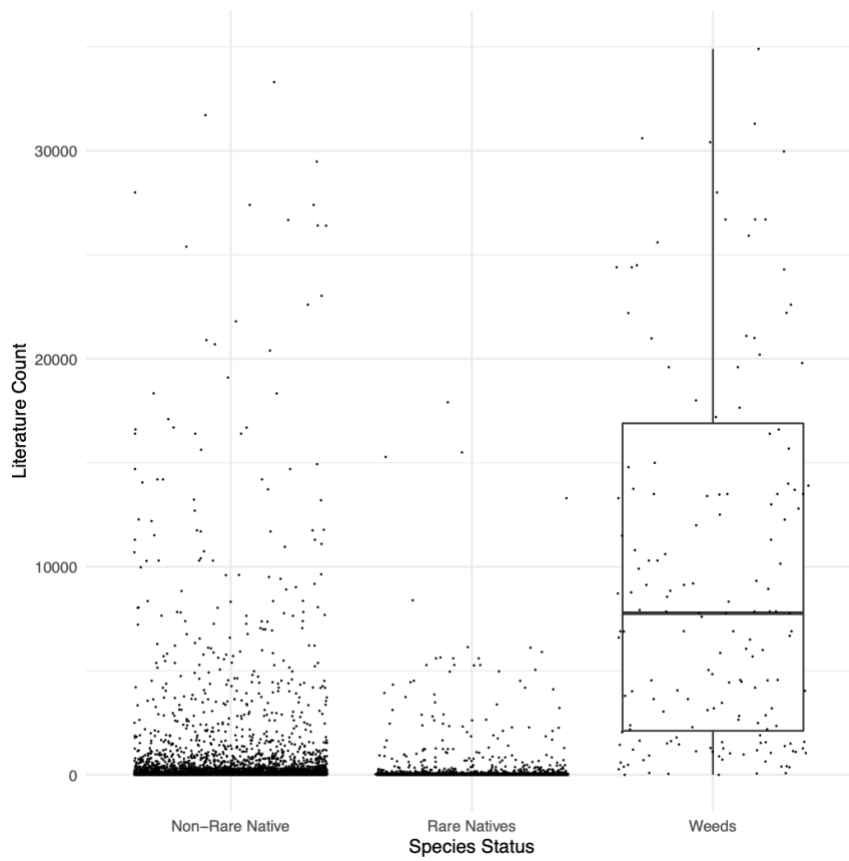


Figure 19: Boxplot of the literature each species status group had. Median for Non-Rare Natives is 75. Median for Rare Natives is 26. Median for Weeds is 7770. The median number of literature across the three groups is 55.

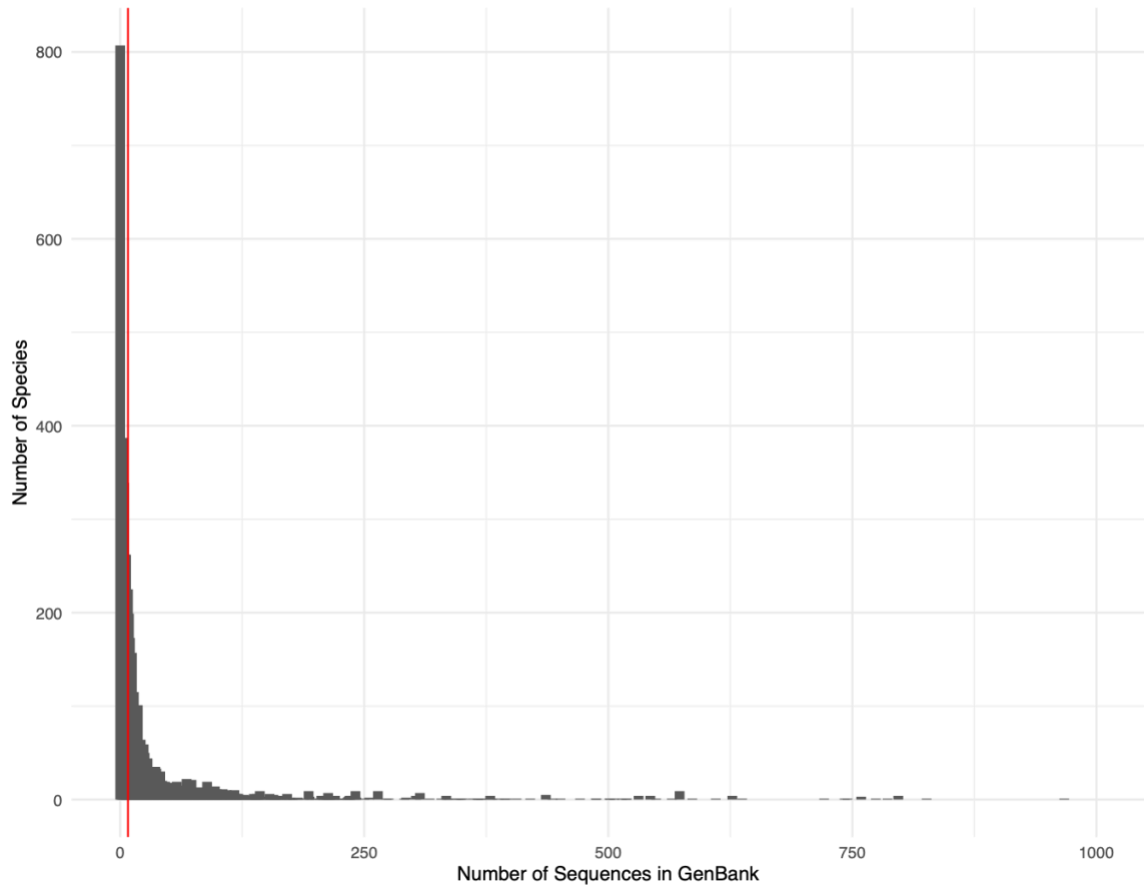


Figure 20: Histogram of genbank sequences from the GenBank scrape. The vertical red line represents the median number of sequences: 8 sequences.

Table 3: The number of plants in our dataset with various CRPR rankings.

CRPR	n
1B.1	251
1B.2	374
1B.3	154
2B.1	60
2B.2	160
2B.3	175
4.1	2
4.2	145
4.3	291
non-rare native	4326
weed	224