San Jose State University

SJSU ScholarWorks

Summer 2013

Predicting Slender False Brome (Brachypodium sylvaticum) Invasion in the Santa Cruz Mountains, California

Janine Ellen Bird
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

PREDICTING SLENDER FALSE BROME (*BRACHYPODIUM SYLVATICUM*)

INVASION IN THE SANTA CRUZ MOUNTAINS, CALIFORNIA

A Thesis

Presented to

The Faculty of the Department of Geography

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Janine E. Bird

August 2013

PREDICTING SLENDER FALSE BROME (*BRACHYPODIUM SYLVATICUM*)
INVASION IN THE SANTA CRUZ MOUNTAINS, CALIFORNIA

by

Janine E. Bird

APPROVED FOR THE DEPARTMENT OF GEOGRAPHY

SAN JOSÉ STATE UNIVERSITY

August 2013

Dr. Richard Taketa   Department of Geography

Dr. M. Kathryn Davis   Department of Geography

Cynthia Powell   Calflora

ABSTRACT

PREDICTING SLENDER FALSE BROME (*BRACHYPODIUM SYLVATICUM)*
INVASION IN THE SANTA CRUZ MOUNTAINS, CALIFORNIA

by Janine E. Bird

Early detection of an invasive species facilitates control and eradication. Slender false brome *(Brachypodium sylvaticum)* was first discovered in the Santa Cruz Mountains of Central California in 2003 as a non-native grass in redwood forests, competing with native vegetation. The current infestation in the Santa Cruz Mountains, estimated to be 300 acres, is concentrated in San Mateo County and could be eradicated. This study sought to determine most likely locations of slender false brome in the Santa Cruz Mountains by assessing environmental attributes of known presence locations using species distribution modeling and Maxent software. The study used 1,320 species presence points collected in field surveys conducted from 2009 to 2012, GIS environmental layers covering a 940 km² study area, and the machine-learning program Maxent to identify slender false brome habitat at a 30 m resolution in the Santa Cruz Mountains. Maxent models successfully identified locations of potential distribution of slender false brome (training AUC = 0.961, test AUC = 0.960). Annual precipitation, average annual maximum or minimum temperature, and soils were the most important predictors. An independent dataset corroborated the performance of the Maxent model. Maxent could be used by land managers for targeting field surveys by predicting most likely *B. sylvaticum* habitat in the Santa Cruz Mountains.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Slender false brome (*Brachypodium sylvaticum*) is a perennial bunchgrass native to Europe and northern Africa. It was cultivated as an ornamental in North America beginning in the early 20th century but became naturalized, eventually reaching invasive status in Oregon and later Washington State. These northwestern infestations spread through seed dispersal to eventually cover thousands of acres in Oregon (Oregon Department of Agriculture, 2013). Slender false brome ("BRSY" is the United States Department of Agriculture symbol for *B. sylvaticum*) spreads readily in western North America where conditions are favorable. An outlier population of BRSY was found in San Mateo County, California, in 2003 (Institute for Applied Ecology, 2013) and has since been estimated to have reached at least 300 acres in size (Midpeninsula Regional Open Space District staff, personal communication, January 2013) . This California infestation inhabits the redwood ecosystem of the San Mateo County portion of the Santa Cruz Mountains in the Pacific Coast Range and is displacing the native vegetation of the understory. If the spread of BRSY is not controlled, California may experience the level of infestation found in the Pacific Northwest. Using species distribution modeling to determine the most likely locations of BRSY populations can aid land managers in planning effective surveys, ensuring that likely BRSY habitat is checked so that the grass can be controlled. This study was conducted to predict locations of slender false brome in the Santa Cruz Mountains by assessing environmental attributes of known presence locations using species distribution modeling.

Species distribution modeling (SDM) is an analytical tool used by biologists to analyze a landscape and determine where a species is likely to occur, based on the characteristics of that species' realized niche. One SDM model, Maxent (maximum entropy), uses machine learning to process presence-only data and environmental attributes to predict potential distribution. Use of presence-only data is an appropriate modeling method for an invasive species where distribution is unstable and absence data does not necessarily indicate unsuitable habitat (Jarnevich & Reynolds, 2011). The slender false brome presence data in this study were collected from 2009 to 2012 by the staff of a local land management agency (Midpeninsula Regional Open Space District) and consist of geographic coordinates recorded with GPS units where BRSY was observed growing. We surveyed for slender false brome for the purpose of treating and eliminating any infestations. Although we recorded only the presence of the grass, we subjectively observed environmental characteristics, such as vegetation class, or percent of canopy cover, that seemed to favor existence and spread of slender false brome, and noted that these similarities of what appeared to be ideal BRSY habitat could be used to efficiently target future surveys. Maxent allows an analysis of these environmental attributes by analyzing geographic information system (GIS) layers, where the study area is divided into a grid (in this study, a 30 m by 30 m cell size). Each cell of the study area holds a value for every environmental layer (e.g., soil type, maximum annual precipitation, canopy cover, or vegetation class). Maxent compares the environmental layer values for each cell where a presence point for a species has been recorded, and uses these combined attributes to indicate other cells where the environmental values are

similar, indicating suitable habitat for that species and hence target survey locations. Maxent showed that annual precipitation, average annual minimum temperature, and soils were the most important predictor layers for suitable slender false brome habitat.

Slender False Brome (*Brachypodium sylvaticum*)

Slender false brome, or *Brachypodium sylvaticum* (Huds.) P. Beauv., is a perennial bunchgrass native to Europe, northern Africa, and western Asia that inhabits forests, woodlands, and upland prairies (Invasive.org: Center for Invasive Species and Ecosystem Health, 2005; Piep, 2012). It reaches 30-120 cm in height, with blades 4-15 mm wide. Per stem, the inflorescence produces 3-12 spikelets with short to nonexistent pedicels. (*Brachypodium* is derived from the Greek "*brachys*" for "short," and "*podion,*" "foot;" "*sylvaticum*" references the Latin "*sylva,*" or "forest.") Slender false brome forms dense clumps and spreads by seed dispersal, not by rhizome production, with seeds that are transported away from parent plants by animals, humans, and waterways. Seedbanks appear not to persist much longer than one year, but seed viability over time has not been tested in North America (Invasive.org: Center for Invasive Species and Ecosystem Health, 2005). One plant may produce hundreds of seeds in a season. In the California Floristic Province, BRSY is documented at elevations of less than 600 m (Piep, 2012) in forest understories and open grasslands.

Slender false brome was introduced to North America as an ornamental grass, probably in the early 20th century. First records of the species are from 1939 near Eugene, Oregon, and by 1966, near Corvallis, Oregon, two large colonies were apparently thoroughly naturalized (Chambers, 1966). By 2001, BRSY covered thousands

of acres in Oregon, concentrated in the Willamette Valley (Invasive.org: Center for Invasive Species and Ecosystem Health, 2005), and in 2007, BRSY was sighted in Washington State, and has been found in Virginia and British Columbia (Piep, 2012). In December of 2003, BRSY was first identified in California from samples collected near Schilling Lake, near Woodside, in San Mateo County on the San Francisco Peninsula (False-brome Working Group, 2004). By 2009, the San Francisco Peninsula population was estimated to have reached as much as 300 acres (~120 hectares), occurring on private parcels and in public lands such as open space preserves and a county park, with nearly all California occurrences found in and adjacent to the Santa Cruz Mountains. A small BRSY population was recorded in 2009 about 500 km to the north, in Del Norte County, per Calflora records (Calflora, 2013).

Figure 1: Slender false brome forming monoculture on woodland hillside in San Mateo County, California.

The California population of BRSY may have been planted as an ornamental grass, as in Oregon. Seeds may also have been introduced in feed for horses (Institute for Applied Ecology, 2013), as the area around Woodside has a thriving equestrian community. The North American invasions of BRSY in the Pacific Northwest and in California have been documented as excluding most native plant species by completely dominating the understory and open habitats (King County Noxious Weed Control Program, 2012). As a non-native invader, BRSY has low palatability to wildlife (Kaye, 2001). Slender false brome has a very broad ecological amplitude (Kaye), with tolerance for shade, partial shade, and sun, a range of elevations and plant communities, and varying moisture regimes. Dispersal seems to follow roads and trails, as seeds are carried by human and animal activity, as well as waterways, including ephemeral streams. Slender false brome has received an A-level rating as a noxious weed in the state (California Department of Food and Agriculture, 2010), a designation indicating it is a species known to cause ecological and economic damage and/or be a human health risk, and is recommended for statewide eradication by the CDFA.

Santa Cruz Mountains Ecosystem

The ecosystem of the Santa Cruz Mountains comprises a wide variety of flora due to the complex topography and unique weather patterns of the area. The Pacific Coast Range of California is formed of low mountains oriented from the northwest to the southeast, divided into several ranges including the Santa Cruz Mountains. Spanning the San Francisco Peninsula southward, the Santa Cruz Mountains are bounded by the Golden Gate to the north, the Pajaro River to the south, the Pacific Ocean to the west, and

the San Francisco Bay to the east, with a north-south distance of about 120 km, and an east-west distance of about 45 km. The area includes San Francisco, San Mateo, Santa Clara, and Santa Cruz Counties, with San Mateo County, on the peninsula towards the north, lying completely inside the Santa Cruz Mountain region. Elevations in this northern stretch of the Santa Cruz Mountains do not exceed 900 m. The presence of the Pacific Ocean to the west contributes greater moisture to the coastal side of the range. The complex topography of valleys and ridges without consistent orientation, along with ocean and bay influences, create a wide range of microclimates within a relatively small geographic area (Beidleman & Kozloff, 2003). Summer fog is an important attribute of the area, and is formed when humid air meets the cold ocean water offshore, causing the moisture in the air to condense. Rising warm air on the land mass causes the ocean fog to be drawn inland, coating the coastal ridges with moisture. Vegetation in the Santa Cruz Mountains is adapted to this fog in its moisture requirements. The Mediterranean climate of Central California brings rainfall only during the winter months of the year, from about October to April, but there may be extended dry spells even during the rainy season and rainfall amounts can fluctuate greatly. Fog reduces the amount of water loss from plants and soil. Additionally, the mist of the fog condenses on foliage and drips to the soil, contributing a significant amount of moisture; in some cases more than the rainfall itself (Beidleman & Kozloff).

The "San Francisco Bay Area" region of the California Floristic Province includes vegetation community types ranging from wet redwood forest, to dry oak and pine woodland, to chaparral (Hickman, 1993). The vegetated areas of San Mateo County

have predominantly redwood forest, Douglas-fir forest, oak woodlands, annual grassland, and chaparral. Redwood (*Sequoia sempervirens*) groves thrive in areas with greater annual rainfall (Beidleman & Kozloff, 2003), and are characteristically associated with *Lithocarpus densiflorus* (tanbark oak), *Polystichum munitum* (western sword fern), *Oxalis oregana* (redwood sorrel), may include *Umbellularia californica* (California bay), and in drier areas are associated with *Pseudotsuga menziesii* (Douglas-fir) (Beidleman & Kozloff; Thomas, 1961). The understory of redwood groves supports relatively few species of vegetation because of the dense canopy and acidic duff produced by the redwood trees; the species that thrive here are well-adapted to those unique conditions. Because of the delicate balance of moisture and nutrients supporting the redwood plant community, this association of vegetation is particularly sensitive to the pressures of a non-native invasive plant. The effects of invasion happen at multiple scales, can influence community structure, ecosystem processes and function, and the nature and intensity of ecological interactions (Brown et al., 2008), possibly leading to disastrous effects in a sensitive system like the redwood plant community. Redwoods themselves are intolerant of prolonged drought (Ornduff et al., 2003) and would suffer if a plant invader competed for moisture.

Species Distribution Modeling and Maxent

Ecological researchers have begun over the past decade to use species distribution models (SDMs) to predict potential distribution of native and exotic species (Baldwin, 2009; Phillips et al., 2006). SDMs are tools that can predict where a species is likely to occur given the environmental conditions related to that species' success. Predictive

7

modeling of a species' geographic distribution is an aid to land managers who seek to preserve declining native species, or to eliminate invasive exotic species. SDMs may be used to guide field surveys for species with limited distribution (Phillips et al.) by determining the potential distribution of an invasive species, thereby allowing land managers to focus control efforts. The Maxent algorithm estimates a target probability distribution by finding the probability distribution of maximum entropy (the distribution closest to uniform) and identifies suitable habitat outside of areas where the species has been identified.

Presence-only versus presence-absence modeling. Presence-only modeling, which uses occurrence data, has been shown to be the most accurate modeling method for non-native invasive plant species, providing more useful results than presence-absence modeling (Anderson et al., 2003; Elith et al., 2006; Jarnevich & Reynolds, 2011; Phillips et al., 2006). Because the non-native species has appeared in the location in recent time and is not in equilibrium with its environment, absence data may not accurately portray the inability of the species to thrive at a location. Absence at any location may indicate unsuitability of the environmental features, or may simply indicate that the species has not yet invaded

Maximum entropy (Maxent). Maxent is a species distribution modeling technique that uses machine learning method to analyze presence-only species data matched with a suite of environmental variables to predict suitability of habitat. Maximum entropy theory applied to species dynamics specifies that without outside influences, processes move towards maximum entropy; that is, towards the most random

distribution. In applying a model with environmental constraints, the assumption is that in the absence of those constraints, the species will move to the geographical distribution of maximum entropy (Phillips et al., 2006). By employing a model that analyzes the environmental constraints, also known as covariates, it is possible to identify geographic locations where the chances are significantly better than random that the modeled species could occur.

Maxent software for species habitat modeling (v. 3.3.3k, available for free from http://www.cs.princeton.edu/~schapire/maxent/) (Phillips et al., 2013) provides a graphic user interface (GUI) to allow easy manipulation of inputs. Running Maxent incorporates metrics for evaluating model fit and contribution of each environmental layer. Maxent is a presence-only modeling technique that represents a species' geographic distribution by relating recorded observances of the species to a host of environmental variables at those locations. Maxent provides niche-based modeling, representing an approximation of the species' fundamental environmental niche. The fundamental niche that a species could occupy is a theoretical construct of potential distribution. A species will most likely not occupy its entire fundamental niche; that is, there will be areas that a species could thrive, but the species is not found there. The species occupies only the area known as its realized niche, because of human influence, biotic interactions, or geographic barriers (Pulliam, 2000), not the entire fundamental niche, shown by modeling, where long-term survival would be possible (Phillips et al., 2006). Areas with environmental variables that satisfy the requirements of BRSY's fundamental niche represent the grass's potential distribution; areas occupied by BRSY constitute BRSY's realized distribution. Surveys

of potential distribution sites have provided BRSY presence data to use in the Maxent model. Environmental conditions at these recorded areas of occurrence are samples from the realized niche (Phillips et al., 2006), and can be used to model BRSY's fundamental niche.

Model Structure and Evaluation

Three components are needed for statistical modeling of species distribution (Austin, 2002): an ecological model, a data model, and a statistical model. The ecological model provides constraints that represent ecological assumptions about what is necessary for the species to thrive. An ideal ecological model would represent all environmental features that constrain the geographical distribution of the species, with the assumption that the absence of those features would tend towards the distribution of maximum entropy (Phillips et al., 2006). The data model describes the methods by which the presence locations were collected. The statistical model provides a scientific method for analyzing the environmental attributes and species presence data; Maxent is a statistical model. Maxent uses what is known as presence-only data, where coordinates of species presence have been recorded, and evaluates the study area for potential distribution by classifying the likelihood of presence versus random, rather than presence versus absence. Maxent models a probability distribution for a species over the pixels in the study region without interpreting pixels with no species records as absences (Phillips et al.). Use of presence-only data is an appropriate modeling method for an invasive species where distribution is unstable, and absence data do not necessarily indicate unsuitable habitat (Jarnevich & Reynolds, 2011).

Maxent evaluates background data by assigning a probability that a randomly chosen presence site is ranked above a random background site. The model uses the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) to evaluate results. An AUC of 0.5 is considered a random ranking; an AUC of 1.0 is a perfect ranking. This means that a model with an AUC of 0.5 is just as likely to predict a presence point as a random background point; a model with an AUC of 1.0 will always predict a presence point over a random point. Researchers have evaluated the default settings of Maxent by comparing Maxent's "presence from random" classification using presence-only data, with the results of "presence from absence" models, testing the same geographic area. Maxent consistently produced a high AUC, with a model generating an AUC of >0.75 considered to be potentially useful (Elith, 2002).

Species distribution models identify areas of suitable habitat by relating environmental variables to locations of known occurrence where the species has been recorded in field-collected data. The best scenario is if the whole range of suitability for each variable is represented in the occurrence data, thereby capturing the entire possibility of the fundamental niche. A small geographic range represented in occurrence data may not provide this full distribution (Phillips et al., 2006). Environmental conditions are described in climatic variables (e.g., temperature and precipitation), topographic variables (elevation; slope; aspect), landcover variables (vegetation communities; canopy cover), and potential vegetation (soil; geology). The Maxent user provides a file of species occurrence points, and chooses which environmental variables to enter into the model. Maxent shows the AUC for each environmental variable entered,

as well as an AUC for the predictive accuracy of the model as a whole. Those environmental variables with a higher AUC contribute more information to Maxent's prediction across the study area; a model with a high AUC is much more likely to predict a presence location than a random location. For this study, eighteen environmental variables were incorporated for use in Maxent to detect potential slender false brome habitat in the northern stretch of the Santa Cruz Mountains of California. The purpose of this analysis was to identify the areas, at a 30 m resolution, that would best support slender false brome.

 Field Observations

The BRSY species occurrence data for this study were collected by staff of a local land management agency, Midpeninsula Regional Open Space District (MROSD, District) during surveys of private property, District Open Space Preserves (OSPs), and other public lands, for the purpose of finding and eliminating slender false brome infestations. We traversed land parcels on foot, using GPS (global positioning system) units to mark locations where the grass was observed, recorded in geographic coordinates. Subjective field observations of occurrence locations showed that slender false brome seemed most likely to thrive in redwood habitat, but was found less in nearby areas dominated by Douglas-fir trees. Infestation patterns seemed to show that BRSY seeds travel downhill with gravity or water flow, so that steeper areas showed dispersal over greater distances, whereas level terrain seemed to have denser infestations, as seeds were most likely to fall close to parent plants. Environmental attributes related to these observations include vegetation categories, canopy cover, slope, hydrologic unit, and

stream flow, and were therefore included among the layers used in the model for this study.

Figure 2: Study area shown with box outline. Inset map shows location in state of California.

Method

Study Area

The location for this study was in the northern portion of the Santa Cruz Mountains of coastal California, comprised of portions of San Mateo County, Santa Clara County, and a small area of Santa Cruz County (Figure 2), and consists of a box drawn in geographic space with all presence data located well inside the box. The bounds of the rectangle were set at -122.46 to -122.03 longitude, 37.24 to 37.50 latitude, forming a box approximately 38 km west-to-east, and 28 km north-to-south. These study area limits were chosen by visual map inspection considering 1) all BRSY presence data recorded in the Santa Cruz Mountains, California, 2) the hydrologic units of the area, and 3) the northern expanse of MROSD's Open Space Preserves (see Figures 3 and 4).

Most of the known BRSY presence data in California are concentrated in a small geographic location, allowing this study to focus on a single area radiating out from those points. Hydrologic units were considered important because the seeds of BRSY appear to be transported by waterways and gravity. We noted in subjective field observations that small infestations of BRSY could be present in isolated locations great distances from the main infestation, but that these outlier locations were immediately adjacent to waterways downstream from the heavy infestations, leading us to conclude that the seeds had been transported in the streams and that waterways could be important vectors for long-distance seed dispersal. The nationally recognized Hydrologic Unit Codes (HUCs) used to designate watersheds increase numerically with a decrease in basin size. The most detailed level, HUC 12, which indicates watersheds of second and third order

streams, was used for this study because of the limited geographic extent of BRSY

occurrence points, and the presence of small waterways in this region.

The presence data used in this study were collected by staff working for

Midpeninsula Regional Open Space District, an agency that manages open space

preserves for human enjoyment and for the sake of preserving the land itself. Part of the

mission of MROSD is to protect native flora and fauna species and minimize exotic

invasive ones. Providing BRSY distribution modeling data to a public agency able to

address infestations on their lands and nearby parcels assists the effort in decreasing or

eradicating slender false brome in California, so considering MROSD preserves in

planning the study area was desirable.

Figure 3: Study area outline showing presence data as dots and local hydrologic units (HUC 12) as polygons.

Figure 4: Study area outline showing presence data as dots, MROSD open space preserves as darker polygons, and other public lands as lighter polygons.

Presence Data

Presence data consisted of 1,321 geographic point locations recorded with hand-held GPS (global positioning system) units where slender false brome was observed growing during 2009 to 2012 field surveys conducted by staff of Midpeninsula Regional Open Space District. We recorded only one point at locations where BRSY occurred in a large patch. Although some vegetation monitoring efforts use polygons recorded with GPS units to show species presence, we did not employ this method because steep and challenging terrain sometimes prevented on-foot access to the entire infestation perimeter, and often tree canopy prevented consistent satellite reception. Most surveys occurred on privately owned parcels in which the entire parcel was inspected for the presence of slender false brome. For the purpose of treating BRSY infestations, we drew

maps by hand in the field to show the extent of an infestation, using a simple "x,y" coordinate pair recorded with GPS to locate that infestation on the surveyed parcel. For the purpose of this study, the GPS coordinates alone provide an accurate record of observed presence of BRSY. Critiques of presence-only data modeling specify that sampling must be either random or representative, with important covariates not missing due to inaccessibility or infrequency (Franklin, 2009; Royle et al., 2012; Syfert et al., 2013; Yackulic et al., 2013). The presence data for this study are representative in that entire multi-acre parcels were surveyed and all observed BRSY presence was recorded regardless of proximity to roads or trails, and unrelated to environmental conditions. Areas that were inaccessible on foot were still visually inspected from the nearest vantage point, and presence of BRSY was recorded.

The GPS units used in field surveys were Garmin GPSMAP 60CSx handheld devices, formatted to collect location data in longitude/latitude decimal degrees using the WGS 84 datum. One point in the dataset fell in the urban region considerably to the east of the survey area, was deemed to have an erroneous x value, and was excluded from this study. The remaining 1,320 points were reprojected in ArcGIS v.10.0 (Esri, 2010) to UTM coordinates of an easting and northing (NAD 1983, UTM Zone 10N). These coordinates were exported to a text file and converted to a .CSV file with the format "Species,x,y", (e.g., "BRSY,565971.2849,4138121.5010") with one coordinate pair per presence record, for use in Maxent.

Environmental Variables

        Environmental variables ("layers", "features", or "covariates") used as inputs for Maxent should be representative of characteristics that allow the modeled species to exist and increase. The choice of variables has a subjective component, but if the modeled species is a plant, concepts of botany lead towards choosing features that influence vegetation success in any environment: moisture, light, temperature, and nutrients. Choosing environmental variables asserts that they represent all environmental features that constrain geographical distribution of the species (Phillips et al., 2006); however, Maxent allows determination of which features are most influential for the species modeled. For this study, environmental variables used can be classified in four categories: climate, topography, hydrology, and potential vegetation/vegetation. Climate variables used (USDA NRCS 30-year normal, averaged over 1981 to 2010) were annual precipitation in inches, and annual maximum temperature and annual minimum temperature both in degrees Fahrenheit. Topography variables were in meters (referenced to the North American Vertical Datum of 1988), at the 30 m by 30 m scale, derived from a Digital Elevation Model (DEM), and included elevation in meters, slope in degrees, and aspect in compass degrees. Hydrologic units, at the sub-watershed level ("HUC 12", referring to the twelve-digit code used nationally), and National Hydrography Dataset Flowlines (unique stream segments) were used. Potential vegetation features included two sets of soil data and two sets of geology data (covering the same geographic space, but comprised of differing edaphic and geologic information respectively). Vegetation features included two landcover classification layers (National

Landcover Database, with 16 classes, and California Digital Conservation Atlas, with 12 classes), four vegetation type classifications (from various sources, with 108, 43, 22, and 22 categories respectively), and total cover from above (non-overlapping vegetation cover, from the U.S. Department of Agriculture Forest Service, in 11 categories by percentage).

Environmental variables were derived from various sources in both vector (shapefile) and raster format. All were processed in ArcGIS: reprojected to NAD 83 UTM Zone 10N, clipped to a bounding box larger than the study area, and converted to rasters with a 30 m by 30 m cell size aligned to a registration point, so that every layer matched exactly in grid size and location for direct geographical comparison. The 30 m size creates a grid with cells small enough to likely capture the environmental factors associated with BRSY (Franklin, 2009). Although resampling of larger grid cells to smaller grid cells does not increase the resolution of the data, the matched cell sizes allow all data to be spatially linked. For use in Maxent, data is categorized as either Continuous (e.g., elevation) or Categorical (e.g., soil type) in format. See Table 1 for environmental variables used.

Table 1: Environmental Variables

FEATURE	LAYER NAME	RANGE OR TYPES	UNIT OR FIELD	FORMAT	SOURCE
CLIMATE					
temperature maximum	tempmax_clip	56.46 to 73.3	DEGREES F	CONT	USDA NRCS
temperature minimum	tempmin_clip	43.8 to 51.3	DEGREES F	CONT	USDA NRCS
precipitation annual	precipann_clip	14. 54 to 61.30	INCHES	CONT	USDA NRCS
TOPOGRAPHY					
elevation	nedutm_clip	-2.27 to 1156.7	METERS	CONT	CaSIL
slope	nedutm_slope	0 to 61.24	DEGREES	CONT	CaSIL
aspect	nedutm_aspect	8 categories	DEGREES	CONT	CaSIL
HYDROLOGY					
hydrologic units	hu12_polygon	[~13 HUs in study area]	HUC_12	CAT	CaSIL
flowlines	nhdflowline_R	[1 or nodata]	[present / absent]	CAT	NHD
POTENTIAL VEGETATION / VEGETATION					
soil	bay_area_soils	29 types	MUID_1	CAT	MROSD
soil	gsmsoilmu_a_ca	29 types	MU_SYM	CAT	USDA NRCS
geology	geology_a_ca	20 types	ORIG_LABEL	CAT	USDA NRCS
geology	bay_area_geology	101 types	P_TYPE	CAT	MROSD
landcover, 12 types	casil_landcover12	[10 types in study area]	GRID_CODE	CAT	CaSIL
landcover, 16 types	nlcd06utm_clip	16 types	NLCD TYPES	CAT	USGS
vegetation, 108 types	vegclass4_utm	108 types	NUM_NAME	CAT	MROSD
vegetation, 22 types	usfs_veg_types	22 types	TYPE	CAT	MROSD
LCMMP regional dominance	evegregdom	43 types	REGIONAL_D	CAT	USDA FS
LCMMP wildlife-habitat relationship	Evegwhrtyp	22 types	WHRTYPE	CAT	USDA FS
LCMMP total cover from above	Evegtotcfa	11 categories + nodata	TOTAL_TREE_CFA	CAT	USDA FS

CaSIL – California Spatial Information Library
http://atlas.ca.gov/
MROSD – Midpeninsula Regional Open Space District
http://www.openspace.org/
USDA FS – United States Department of Agriculture Forest Service
http://www.fs.fed.us/r5/rsl/projects/mapping/
USDA NRCS – United States Department of Agriculture Natural Resources Conservation Service
http://datagateway.nrcs.usda.gov/
USGS – United States Geological Survey
http://viewer.nationalmap.gov/viewer/

Maxent Modeling

ASCII format and mask raster. Maxent input requires all environmental layers to be converted to ASCII file format with identical headers:

```
ncols         2134
nrows         2000
xllcorner     541000
yllcorner     4103000
cellsize      30
NODATA_value  -9999
52 42 43 90 11 11 90 90 43 71 -9999 -9999...
```

Environmental rasters were converted to ASCII with the "Raster to ASCII" tool in ArcGIS. Because these rasters were created (using "Polygon to Raster" tool in ArcGIS) or formatted with an identical bounding box, cellsize, and registration point, ASCII conversion resulted in matching headers and coverage. The "NODATA_value" refers to the listing in the ASCII for any area covered by the bounding box but where the original raster had no coverage, and therefore had no data to display. For this study's dataset, the "nodata" areas were present in two layers: *vegclass4*, a polygon layer with fine-scale vegetation detail, but limited geographical coverage, and *nhdflowline*, a polygon layer that covered the entire bounding box but only provided data as lines in areas of watercourses. Maxent will not process a cell if a "nodata" value of -9999 occurs in any layer in that cell. This means that using these limited rasters may add important information to the species distribution analysis (e.g., more detailed vegetation classes in the case of *vegclass4*), but may also confine Maxent to an unacceptably small geographical area and result in overfitting the model. A useful function of the "nodata" limiting feature is that a mask raster can be created to set the region that Maxent analyzes once a study area has been determined. For this study, environmental layer rasters were originally clipped to an area larger than the study area, so once the study area was

determined, a raster was created with a value of 1 inside the study area, and a value of -9999 for the rest of the area of the bounding box. Including this mask raster with Maxent inputs means that Maxent will analyze only the cells that fall within the study area.

Figure 5: Maxent graphical user interface, with species and environmental layer inputs, and parameters window.

Running Maxent. Analysis of species distribution with Maxent modeling starts with opening the graphical user interface, pointing to the directories of species data, environmental layers (set to Categorical or Continuous), and an output file, and choosing the model's settings (see Figure 5). Maxent allows specifying that part of the presence data be used for training the model, and part set aside for testing the model. Maxent was

22

set to remove duplicate occurrence points, or records occurring in the same grid cell, so that each cell's environmental attributes would only contribute once to the model in a run. In this study, all Maxent runs were completed with 75% (n = 754) training data and 25% (n = 251) testing data out of all non-duplicate BRSY occurrence points (n = 1,005) used in the model.

Maxent output includes graphic display of the area analyzed showing AUC, with a number closer to 1 indicating a greater probability that a randomly chosen presence site will be ranked above a random background site. Environmental layers used to run the model influence output. Visual examination of the graphic display and statistical output can be used to choose which layers to include or exclude for subsequent runs, based on their ecological relevance as shown by the model. See Figure 6, showing Maxent's map output of suitable habitat.

Maxent was run using all environmental layers ("Output_0"), then with the geographically limited vegetation layer, *vegclass4*, removed ("Output_1"). These outputs were evaluated for the jackknife AUC (see Figure 7), all layers with AUC < 0.80 were removed, and two subsequent runs were performed, one with *vegclass4* ("Output_4"), and one without ("Output_2"). An additional run ("Output_3") was performed using a separate version of the *nhdflowline* layer, where all non-stream locations have the nodata value, meaning only linear stream areas were analyzed; other runs used a version of the *nhdflowline* layer with all non-stream locations having a value of 0, so that stream data was analyzed without excluding non-stream areas.

Choice of the AUC < 0.80 limit for removing environmental variables was made based on Elith's assessment that a model AUC of > 0.75 is considered useful (Phillips & Dudik, 2008), and with the objective of not missing habitat (Jarnevich & Reynolds, 2011), since the modeling purpose is to eliminate an invasive species. Because an invasive species is not in equilibrium with its environment, full knowledge of the most influential covariates is not initially possible, and distribution modeling is ideally an iterative process.

Output_0
Area limited
geographically by
vegclass4 raster
extent.

Output_1
Includes full
study area,
and a larger
region is
predicted for
potential
distribution.

Scale of AUC

Output_3
Limited by
nhdflowline raster
extent. Streams
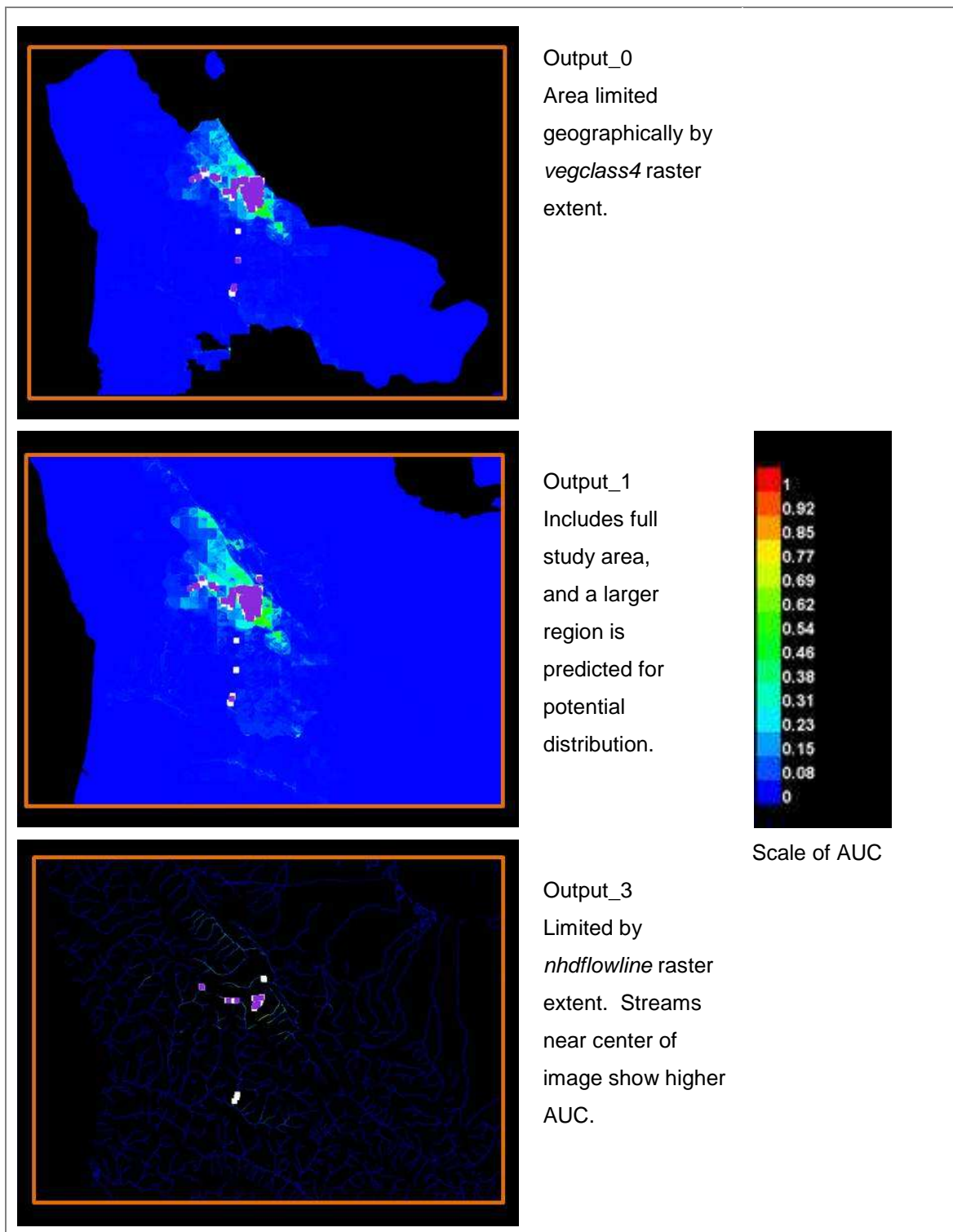near center of
image show higher
AUC.

Figure 6: Maxent output graphics, with study area box superimposed on results. Warmer colors show better environmental conditions; purple areas show presence data locations.
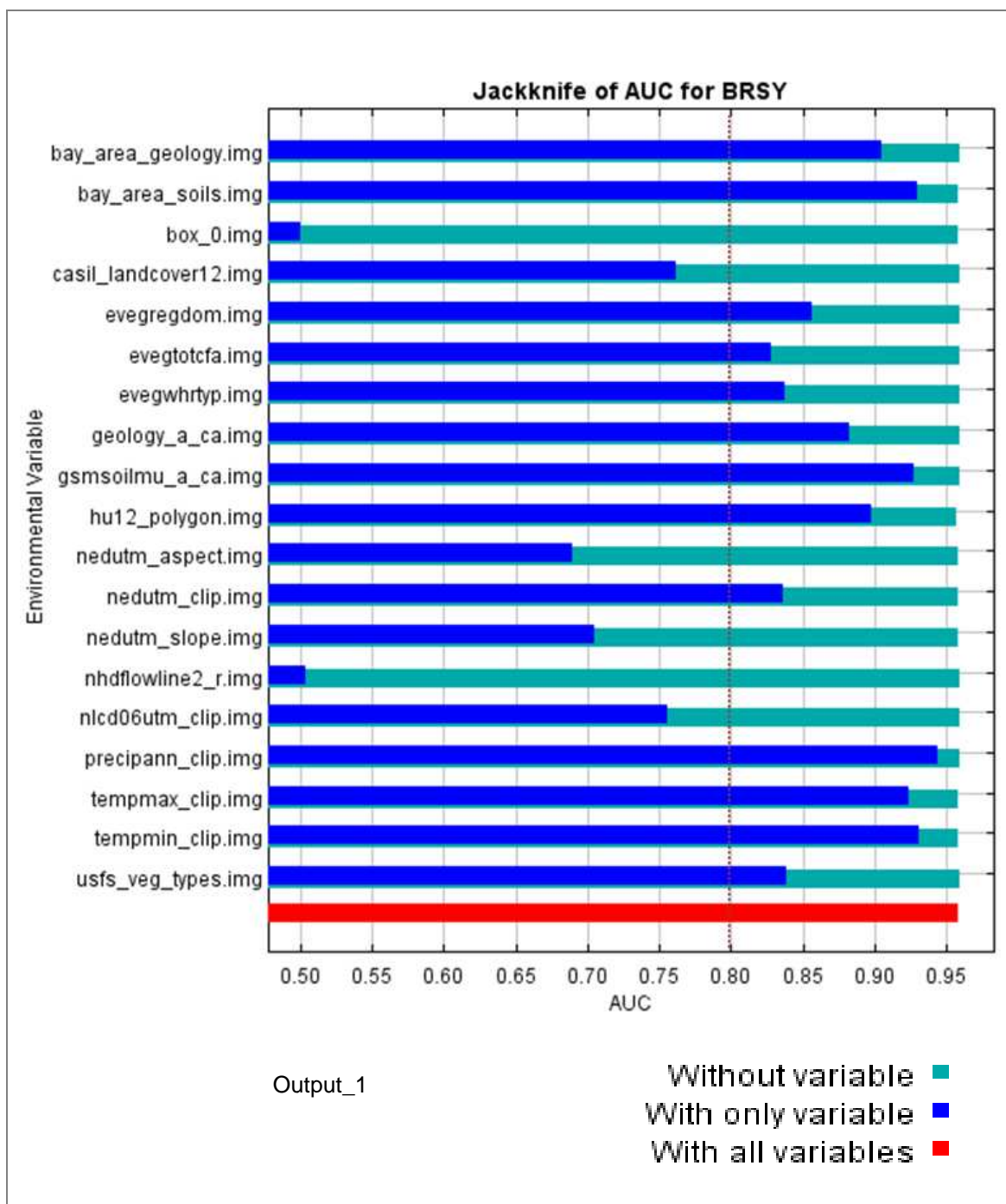
Figure 7: Maxent output graphics, showing AUC on test data, and dashed line at 0.80 added to select environmental variables with lower AUC and therefore less model contribution.

Output Processing

Maxent produces an ASCII grid (.ASC) for each run showing the presence probabilities (gain) for each grid cell of the study area. These files were converted to Maxent output raster files in ArcGIS. To meaningfully link the Maxent gain output to the species presence data, the raster values were extracted for each cell that contained an original presence location using the "Extract Values to Points" tool. This was completed for each .ASC-to-raster output resulting from the Maxent Output_2 (all layers with >0.80 AUC, except *vegclass4*), and Output_4 (all layers with >0.80 AUC, including *vegclass4*). The new point files, now with an appended field showing the Maxent gain in the cell containing each point, were exported and converted to Excel files. The table was sorted according to gain value, and the entire list of points was divided into cutoff percentages, with associated gain, for use in creating new rasters. Cutoff percentages selected were 100, 98, 95, 90, 80, 70, and 50, out of the total number of presence points. At each cutoff percentage the associated gain value was noted. For example, with the cutoff gain value at 98%, ninety-eight percent of the presence points had an equal or greater gain value than the noted value. Tying Maxent output back to the presence points means we can say that of the cells where Maxent predicts more likely occurrence, we are choosing the *most likely* cells, with environmental conditions similar to cells where 98% (in the previous example) of the presence points occurred. Indeed, visual inspection showed that the presence points associated with the lowest gain values were farthest from the main infestation, where environmental conditions were likely to be less similar. This sorting method was employed with the rationale that invasive species monitoring is most likely

not possible at 100% of the area shown by the model to have a chance of supporting the

species. At the same time, a very limited survey approach, with a 50% match, will likely

exclude too many actual presence locations. Choosing a gain value associated with

presence data location cells means that the model output interpretation is tied to

conditions where the species is present. Choosing various cutoff percentages, with a

proportionally sized presence-likely geography associated, means that monitoring

surveys able to target limited area can focus on the most suitable habitat found in that size

area. Table 2 shows gain values associated with Maxent output at each chosen cutoff

percentage.

Table 2: Gain Values at Cutoff Percentages

GAIN CUTOFF %	POINT NUMBER	≥ GAIN VALUE	
		Output_2	Output_4
100%	1320	0.005985*	0.006041
98%	1294	0.115783*	0.059212
95%	1254	0.290746*	0.023659
90%	1188	0.424793	0.385971
80%	1056	0.490851*	0.481561
70%	924	0.551750	0.532657
50%	660	0.582084	0.582084

 * Value also used for Output_3

Using the noted gain value, a new raster was created for each cutoff percentage

for Output_2 and Output_4. This was done by starting with the Maxent raster results,

added to the ArcGIS Raster Calculator with, for example, [Output_2 raster] >= [gain value]. The resulting raster showed a "1" value in each cell with a gain greater than or equal to the cutoff gain value chosen, and a "0" in all other cells. With a lower cutoff percentage, fewer cells received the value of 1, but those cells had the greatest likelihood of supporting BRSY. With each increase in cutoff percentage (50% to 100%), more cells were given the value of 1 in the resulting raster, but the added cells had decreasing likelihood of providing a suitable environment for BRSY. Through this process, a bull's-eye sort of mapping of suitable BRSY habitat is possible, with a choice of geographic size for the survey target, and a means of identifying the most likely areas if resources are limited.

Output_3 from the Maxent run incorporating the *nhdflowline* layer version limited to stream areas was processed using the Output_2 gain values at the percentages of 100, 98, 95, and 80, with rasters created for each cutoff. Output_3 cells could not be matched directly to presence point cells because these streamline-only cells had such a limited extent. Because Output_3 geographical coverage was the same as Output_2, the gains of Output_2 were substituted.

Measure of Model Performance

To measure the performance of the Maxent species distribution model created in this study, a separate BRSY dataset was used. These data were recorded observations of *B. sylvaticum* from the online database maintained by Calflora, a non-profit organization that acts as a repository of information on California plants for education, research, and conservation (Calflora, 2013). The observations are from public agencies, non-profits,

scientists, and private donors. Per Calflora, accuracy of identification is not reviewed.

However, most BRSY observations in the Calflora database were made by known

biologists working in the Santa Cruz Mountains and familiar with *B. sylvaticum*. From

the "Search for Plant Occurrence Records", a "Brachypodium sylvaticum" search in the

"full scientific name" field with no other delimiters (choosing "plain text" for data

export) gave all the Calflora records for BRSY, a total of 149 observations between

January 2004 and March 2011. From these records, five observations had no geographic

coordinates listed, and so were removed. One observation was in Del Norte County,

about 500 km north of the study area, and so was not included because it falls outside the

environmental data assembled for this study. The remaining 143 observations were

added to ArcMap as x,y data, to compare to the original BRSY dataset. To ensure that

only unique Calflora observations were kept, all Calflora observations within 50 m of any

BRSY point in this study's original dataset were eliminated if observation dates matched.

(The distance of 50 m was chosen because satellite reception in the study area usually

provided GPS accuracy of 50 m or better, so adjacent points inside that range were more

likely to be the same occurrence location.) The remaining Calflora observations, thought

to be unique BRSY observations, numbered 37. These 37 Calflora observations were

compared with Maxent modeling results to see if the points fell in areas identified by

Maxent as likely BRSY habitat. To make the evaluation, the Calflora observations were

compared to the raster from Maxent Output_2. Using the "Extract Values to Points" tool

in ArcMap, each Calflora observation was linked to the gain value of the Output_2 raster

cell in which it fell, and those values were appended to the Calflora data attribute table

for each point. The gain values associated with each Calflora data point were used to

show where on the "bull's-eye" of successively more suitable BRSY habitat those

independent observations fell.

Results

All Maxent iterations had similarly high AUC values, and manipulation of environmental variables used in running Maxent did not substantially change predictive strength of the model according to AUC (see Table 3). For the first model iteration (Output_0) using the fine-scale vegetation layer *vegclass4*, and excluding the streams-only version of *nhdflowline*, results were far better than random prediction, with AUC = 0.960 for training data and AUC = 0.957 for the 25% of presence data partitioned for testing. Removing AUC < 0.80 environmental variables for the next iteration, Output_4, gave AUC = 0.958, with only a slight improvement to model performance.

Table 3: Maxent Output Statistics

Receiver Operating Characteristic Curve AUC

	AUC		
	Training Data	Test Data	
Output_0	0.960	0.957	excluding *nhdflowline3*
Output_1	0.961	0.959	excluding *vegclass4* and *nhdflowline3*
Output_2	0.961	0.960	excluding *vegclass4*, *nhdflowline3*, and environmental variables with AUC < 0.80
Output_3	0.992	0.984	excluding *vegclass4, using nhdflowline3* (limited to streams)
Output_4	0.960	0.958	excluding *nhdflowline3, using* vegclass4 (fine-scale vegetation detail)

Random prediction AUC = 0.5

The model iterations using the coarse-scale vegetation layer started with AUC = 0.959 for test data in Output_1, and removing environmental variables with less contribution marginally improved results, with AUC = 0.960 for test data in Output_2. The Maxent run for Output_3 was limited to examining only the raster cells included in streambeds and exhibited a very high AUC (0.984). Although Maxent is not usually used to model linear data (Elith et al., 2011), this iteration of the model did show a selection of stream areas that, according to an acceptable AUC, are more likely to provide suitable BRSY habitat than most of the streams in the study area.

Using gain cutoff percentages to create new rasters enabled mapping of the areas predicted by Maxent to be most likely habitat. Lower percentages indicated smaller areas with the most ideal conditions. Increasing percentages added area, but with an inverse relationship to suitability: these additional areas are less matched to ideal environmental conditions. Figure 8 and Figure 9 show maps of this comparison of using percentage cutoffs for Output_2 and Output_4, respectively. In Figure 8, the cutoffs of 70, 90, and 100% were mapped for Output_2. The area indicated as suitable with a 70% gain cutoff was ~650 acres (263 hectares); at 90%, the area indicated was 2.5 times larger, at 1,638 acres (663 hectares), and at 100%, the area was 25 times larger than the 90% area, at 42,370 acres (17,147 hectares). Table 4 shows the size of the modeled suitable area indicated by Maxent, using the percentage cutoffs. The shaded lines indicate the percentages 70, 90, and 100, mapped in Figures 8 and 9. Using gain cutoffs greatly changes the size of the mapped area, with a narrow or broad suitability focus.
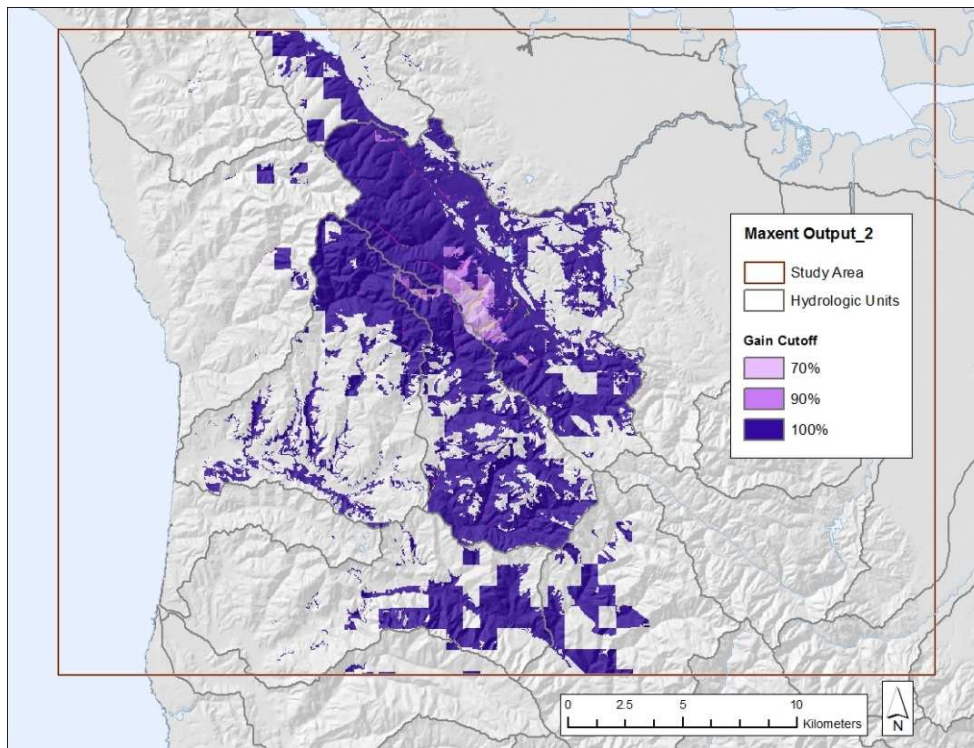
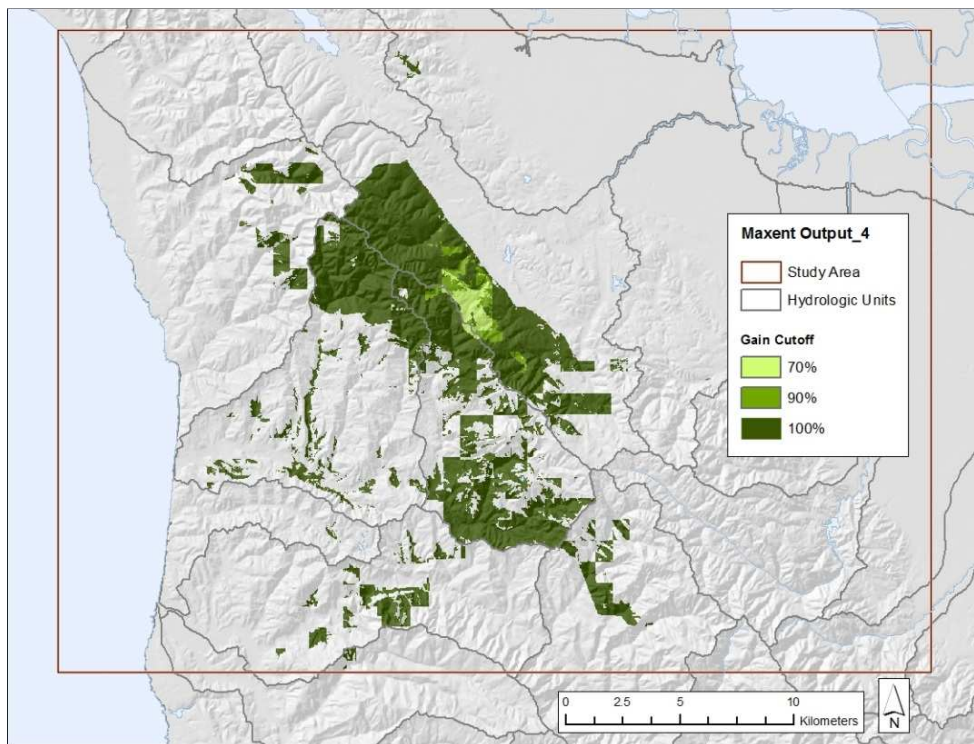**Figure 8: Maxent suitability results for Output_2.**

Figure 9: Maxent suitability results for Output_4.

Table 4: Suitable Area by Maxent Run and Gain Cutoff Percentage

	GAIN CUTOFF %	TOTAL NUMBER CELLS ANALYZED	NUMBER OF CELLS VALUE =1	% OF AREA	SUITABLE AREA IN HECTARES	SUITABLE AREA IN ACRES
Output_2		1047416				
	100%		190,517	18.2%	17,146.53	42,369.93
	98%		36,078	3.4%	3,247.02	8,023.55
	95%		15,114	1.4%	1,360.26	3,361.27
	90%		7,367	0.7%	663.03	1,638.38
	80%		4,366	0.4%	392.94	970.97
	70%		2,917	0.3%	262.53	648.72
	50%		1,558	0.1%	140.22	346.49
Output_4		555327				
	100%		113,967	20.5%	10,257.03	25,345.63
	98%		36,424	6.6%	3,278.16	8,100.50
	95%		13,364	2.4%	1,202.76	2,972.08
	90%		6,515	1.2%	586.35	1,448.90
	80%		3,609	0.6%	324.81	802.62
	70%		2,819	0.5%	253.71	626.93
	50%		2,166	0.4%	194.94	481.71

The coarse-scale vegetation analysis of Output_2 indicated 18% of the analyzed area (the entire study area, about 948 km²) could be suitable BRSY habitat, using the 100% gain cutoff. Output_4, with the fine-scale vegetation layer, analyzed an area about

half the size (~500 km²) of the entire study area, and indicated at the 100% cutoff that a similar percentage of the analyzed area, 20.5%, would be suitable BRSY habitat.

Mapping the results of Output_3, using the linear stream data, showed streambeds with higher probability of suitable BRSY habitat. At the 100% gain cutoff for Output_3, nearly all the streams falling within the largest Output_2 extent (Output_2 at 100%) were indicated. At the 98% gain cutoff for Output_3, considerably fewer streambeds were indicated. Figure 10 shows Output_3 streams at the 98% cutoff on top of the area mapped for Output_2 at 100%.
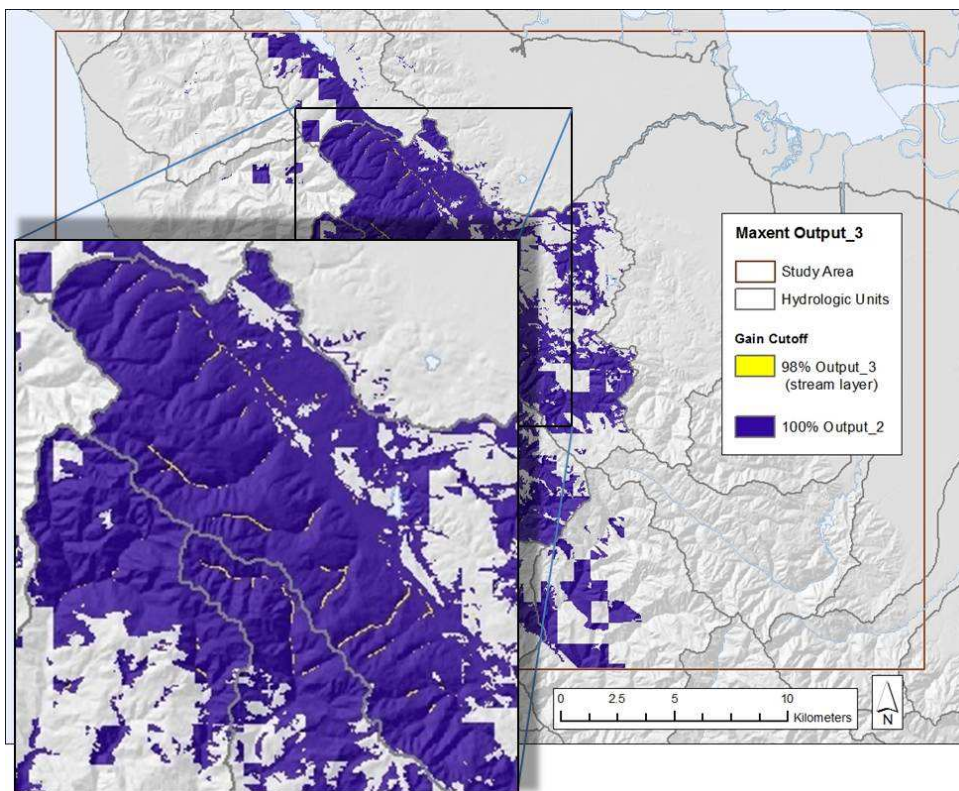

Figure 10: Mapped Maxent results for Output_3, linear stream data, shown on top of Output_2 results at the 100% cutoff. Inset is expanded area.

Model Performance Results

The separate Calflora dataset of BRSY points was matched with gain values for the Maxent Output_2 raster, showing the model gain value in each raster cell where a Calflora point resides. This scenario gives a limited evaluation of what were to happen if land managers used this study's Maxent modeling to find *B. sylvaticum* in the Santa Cruz Mountains. Using the percentage cutoffs listed in Table 2 for Output_2 gain levels, land managers using the 90% cutoff would have searched the area where 14 out of 37 (38%) of the Calflora observations were located. Increasing the cutoff to 95% would have meant that surveys for infestations would have included the locations of 28 out of 37 (76%) of the Calflora observations. See Table 5.

Table 5: Percent of Calflora Points Included in Survey Area by Gain Cutoff Percentage

	GAIN CUTOFF %	SUITABLE AREA IN HECTARES	SUITABLE AREA IN ACRES	% OF CALFLORA POINTS INCLUDED
Output_2				
	100%	17,146.53	42,369.93	92%
	98%	3,247.02	8,023.55	81%
	95%	1,360.26	3,361.27	76%
	90%	663.03	1,638.38	38%
	80%	392.94	970.97	30%
	70%	262.53	648.72	27%
	50%	140.22	346.49	14%

Discussion

Performing species distribution modeling with presence-only data entails the assumption that species sampling is random or is representative throughout a landscape. That is, important covariates cannot be missing from the sampled area and must be similar to the range of covariate values in the overall landscape. Another assumption is that detectability probability is constant across sites (Yackulic et al., 2013). In the case of the BRSY data used for this study, the sampling was not random but was representative of the target area in the Santa Cruz Mountains. The surveys that produced the presence data were an exhaustive inspection of entire parcels for the purpose of finding all BRSY for elimination, as opposed to point surveys limited to confined locations. If a 30 m grid cell in this study was split between an inspected parcel and an unvisited parcel, it would be possible for that cell not to register as a presence cell if the only BRSY occurred in the unvisited portion of the cell. Even if this were the case, the condition that detectability during sampling not vary with covariates that determine occurrence probability would still be met by the exhaustive surveys of all neighboring parcels. The survey sweep gave the BRSY a very likely chance of being recorded at all surveyed sites where it exists. Additionally, Phillips et al. (2009) stated that presence-only modeling inferences are fairly robust to sample selection bias as long as the bias is constant from background points to presence points, which was true of this study.

The environmental layers that most influenced model output were annual precipitation, average annual minimum temperature, average annual maximum temperature, and soils, as was indicated by higher AUC values (Figure 7). Removal of

layers with the least influence on model output, as shown by the jackknife AUC < 0.80, did not substantially change model fit. This may be because the environmental layers chosen accurately represent the requirements of BRSY in its habitat. Alternatively, high AUC values could indicate overfitting of the model, the result of an invasive species not being in equilibrium with its environment (Gallien et al., 2012), or of analyzing a geographical area that is too small and too homogeneous (Thuiller et al., 2004).

Use of a coarse vegetation layer covering the entire study area, or a fine-scale vegetation layer with about half the extent, did not greatly change the statistics of model output, with both iterations indicating a similar percentage of the analyzed area would be suitable habitat (18.2 and 20.5% respectively). The two vegetation layers did indicate different areas of suitability, showing that the number of vegetation categories in a layer influences Maxent in predicting habitat areas. Connecting data shown in the layer to ground truthing could indicate whether the fine-scale vegetation layer provides more accurate output. For example, the coarse vegetation layer identifies one presence data region as "Montane Hardwood-Conifer" that the fine-scale layer separates into "Douglas-Fir/Mixed Hardwoods," "Douglas-Fir/Coast Redwoods," and "Coast Live Oak." Subjective field observations indicated that BRSY was more likely to be found under redwoods than under Douglas-fir, so having the fine-scale vegetation distinction is most likely useful.

Mapped outputs showed a strong connection between hydrologic unit (HU) and Maxent results, as indicated by suitable areas closely following HU lines. While BRSY presence is likely to influence the downstream areas within that HU, there is likely less

real connection between upslope areas that happen to be contained within one HU. Although the HU_12 environmental layer had a high jackknife AUC value of almost 0.90, showing a high contribution to Maxent modeling, removal of this variable for subsequent iterations could possibly improve the model.

Mapped outputs in some areas showed a blocky grid pattern of "most suitable" versus "less suitable" regions (Figures 8 and 9), at a scale matching the resolution of the precipitation and temperature layers (800 x 800 m cells). Conditions in nature would not follow this arbitrary designation, thus showing the limitations of environmental data formats. Because these layers of precipitation and temperature were important in model creation, access to finer-scale climate data could be quite advantageous and eliminate some of the grid effect in suitable habit indications.

The linear data provided by the *nhdflowline* layer did not contribute to the model when used in a full study area format (jackknife AUC = 0.50). When the *nhdflowline* layer was used in a streams-only format, model contribution was similarly low, but mapped output showed interesting results. Streambeds indicated as suitable habitat fell within areas indicated by Output_2 model results, but did not match completely, and may therefore constitute a unique set of information. Monitoring of an invasive species such as BRSY that widely disperses seeds via watercourses would benefit from analyzing flowlines as part of the hydrological data.

Use of cutoff percentages for gain values, with the subsequent rasters that can be mapped, is a very useful methodology for turning Maxent output into target zones for field surveys. Land managers can make choices for invasive species monitoring by using

a higher percentage cutoff to ensure surveys cover the most potential habitat, or, if resources are limited, a lower percentage that gives less area but with the greatest match as suitable habitat. Use of a higher percentage may be most advantageous for open space preserves with limited acreage, where invasive species must be kept to a minimum for best land preservation. Use of a lower percentage may be most suitable for a county-wide effort to target areas of greatest probability of providing suitable BRSY habitat. In this model, the trade-off of identifying more habitat meant adding a significantly greater burden of area to survey, particularly with gain levels over 95%. A more finely-tuned model may improve the returns, so that amount of area to survey does not increase so dramatically with higher cutoff percentages.

Model evaluation with the separate Calflora dataset of BRSY observations showed that the Maxent model created with the coarse-scale vegetation layer and refined to include only layers with AUC >0.80 (resulting in the Output_2 gain values) was able to discriminate likely *B. sylvaticum* habitat in the Santa Cruz Mountains, and could be used to define survey size and location. Maxent modeling gave meaningful thresholds that land managers could use to target BRSY surveys. Lower thresholds showed the most suitable BRSY habitat, and higher thresholds also included somewhat suitable habitat, increasing the chance of finding BRSY locations but requiring surveys of greater acreage. If the Maxent model over-predicted BRSY habitat, a much greater percentage of the Calflora observations would have been included at lower cutoff thresholds. If the model under-predicted BRSY habitat, using even the highest cutoff percentage would not have indicated all the areas necessary to contain most of the Calflora observations.

Because each increase in model cutoff percentage added the locations of more Calflora points, model output appears to be successful at determining the most favorable habitat for BRSY.

The model did not predict habitat at 100% of the Calflora dataset locations, however. Three points, or 8% of the sample, fell in areas not predicted as habitat by the model. One of these points was in a location to the south, near the edge of the study area and about 15 kilometers from the closest BRSY sample from either dataset. This southernmost point was an isolated plant that MROSD staff connected to accidental spread by OSD vehicles (MROSD staff personal communication, January 2013), emphasizing that human influence can greatly change the invasion potential of a noxious weed. The other two points that were not in areas predicted as habitat by Maxent at the 100% gain level were 23 m and 170 m, respectively, from the edge of a grid cell that *was* predicted as habitat. In the case of the 23 m difference, the point fell in the next hydrologic unit from the majority of samples, so the HU layer in the model most likely had the greatest influence in excluding the habitat where this point fell. The 170-meter point is surrounded by areas indicated by Maxent as suitable habitat, inviting closer examination of which environmental layers influenced this point's exclusion. The Maxent model could be fine-tuned for most effective use by incorporating habitat observations about outlier points such as these three.

Conclusion

Maximum entropy (Maxent) modeling was an effective tool in predicting suitable habitat for the invasive grass *Brachypodium sylvaticum* in the Santa Cruz Mountains by

assessing environmental attributes of known presence locations. Model outputs

delineated geographical locations that are likely to support BRSY as an invasive species,

and an independent dataset corroborated the model's effectiveness. Modeling results can

be tailored to show the most suitable habitat, and therefore the most likely areas of

slender false brome presence, given an acreage limit for survey effort. This quantifiable

target output is valuable to land managers seeking to control invasions while needing to

maximize field investigations.

References

Anderson, R. P., Lew, D., & Peterson, A. T. (2003). Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling, 162*, 211-232. doi: http://dx.doi.org/10.1016/S0304-3800(02)00349-6

Austin, M. P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling, 157*, 101-118. doi: 10.1016/s0304-3800(02)00205-3

Baldwin, R. A. (2009). Use of Maximum Entropy Modeling in Wildlife Research. *Entropy, 11*, 854-866. doi: 10.3390/e11040854

Beidleman, L. H., & Kozloff, E. N. (2003). *Plants of the San Francisco Bay region: Mendocino to Monterey* (2 ed.). Berkeley and Los Angeles, California: University of California Press, Ltd.

Brown, K. A., Spector, S., & Wu, W. (2008). Multi-scale analysis of species introductions: combining landscape and demographic models to improve management decisions about non-native species. *Journal of Applied Ecology, 45*, 1639-1648. doi: 10.1111/j.1365-2664.2008.01550.x

Calflora. (2013). Information on California plants for education, research and conservation [web application] Berkeley, California: The Calflora Database [a non-profit organization].   Retrieved June 11, 2013, from http://www.calflora.org/

California Department of Food and Agriculture. (2010). Notes on Identification, Biology, and Management of Plants Defined as Noxious Weeds by California Law: Encycloweedia Data Sheets.   Retrieved February 26, 2013, from http://www.cdfa.ca.gov/plant/ipc/weedinfo/winfo_list-pestrating.htm

Chambers, K. L. (1966). Notes on some grasses of the pacific coast. *Madroño, 18*, 250-251.

Elith, J. (2002). Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In S. Ferson & M. Burgman (Eds.), *Quantitative methods for conservation biology* (pp. 39-58): Springer.

Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., . . . Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography, 29*(2), 129-151. doi: 10.1111/j.2006.0906-7590.04596.x

Elith, J., Phillips, S. J., Hastie, T., Dudik, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions, 17*, 43-57. doi: 10.1111/j.1472-4642.2010.00725.x

Esri. (2010). ArcGIS 10.0 (Version 10.0).

False-brome Working Group. (2004). False-brome Working Group Newsletter. *False-brome Working Group Newsletter*.

Franklin, J. (2009). *Mapping Species Distributions: Spatial Inference and Prediction*. New York: Cambridge University Press.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N. E., & Thuiller, W. (2012). Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography, 21*, 1126-1136. doi: 10.1111/j.1466-8238.2012.00768.x

Hickman, J. C., (Ed.). (1993). *The Jepson manual: higher plants of California* (J. C. Hickman Ed.). Berkeley and Los Angeles, California: University of California Press.

Institute for Applied Ecology. (2013). False Brome Working Group. Retrieved February 7, 2013, from http://appliedeco.org/

Invasive.org: Center for Invasive Species and Ecosystem Health. (2005). Weed Alert! *Brachypodium sylvaticum* (Huds.) P. Beauv. (Slender false-brome, false-brome). Retrieved February 8, 2013, from http://www.invasive.org/gist/alert/alrtbrac.html

Jarnevich, C. S., & Reynolds, L. V. (2011). Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. *Biological Invasions, 13*, 153-163. doi: 10.1007/s10530-010-9798-4

Kaye, T. (2001). *Brachypodium sylvaticum* (Poaceae) in the Pacific Northwest. *BEN: Botanical Electronic News, 277*. http://www.ou.edu/cas/botany-micro/ben/ben277.html

King County Noxious Weed Control Program. (2012, February 9, 2012). Noxious weeds: King County, Washington. Retrieved February 7, 2013, from http://www.kingcounty.gov/environment/animalsAndPlants/noxious-weeds/weed-identification/false-brome.aspx

Oregon Department of Agriculture. ODA Plant Programs, Noxious Weed Control: false brome (*Brachypodium sylvaticum* (Huds.) Beav.).   Retrieved February 7, 2013, from http://www.oregon.gov/ODA/PLANT/WEEDS/pages/profile_falsebrome.aspx

Ornduff, R., Faber, P. M., & Keeler-Wolf, T. (2003). Introduction to California plant life. from University of California Press

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*, 231-259. doi: http://dx.doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., & Dudik, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography, 31*, 161-175. doi: 10.1111/j.0906-7590.2008.5203.x

Phillips, S. J., Dudik, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications, 19*, 181-197. doi: 10.1890/07-2153.1

Phillips, S. J., Dudik, M., & Schapire, R. E. (2013). Maxent software for species distribution modeling.   Retrieved January 27, 2013, from http://www.cs.princeton.edu/~schapire/maxent/

Piep, M. B., (Ed.). (2012). Brachypodium. *Jepson eFlora.* Retrieved February 8, 2013, from http://ucjeps.berkeley.edu/cgi-bin/get_IJM.pl?tid=16044

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters, 3*, 349-361. doi: 10.1046/j.1461-0248.2000.00143.x

Royle, J. A., Chandler, R. B., Yackulic, C., & Nichols, J. D. (2012). Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution, 3*, 545-554. doi: 10.1111/j.2041-210X.2011.00182.x

Syfert, M. M., Smith, M. J., & Coomes, D. A. (2013). The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *Plos One, 8*(2). doi: 10.1371/journal.pone.0055158

Thomas, J. H. (1961). *Flora of the Santa Cruz Mountains of California: A Manual of the Vascular Plants*. Stanford, California: Stanford University Press.

Thuiller, W., Brotons, L., Araujo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography, 27*, 165-172. doi: 10.1111/j.0906-7590.2004.03673.x

Yackulic, C. B., Chandler, R., Zipkin, E. F., Royle, J. A., Nichols, J. D., Grant, E. H. C., & Veran, S. (2013). Presence-only modelling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution, 4*, 236-243. doi: 10.1111/2041-210x.12004