

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Utility of temporally-biased invasive species distribution models in the detection of *Euwallacea* sp. nr. *forficatus* in California

**Permalink**

<https://escholarship.org/uc/item/1557w99t>

**Author**

Dimson, Monica

**Publication Date**

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Utility of temporally-biased invasive species distribution models in the detection of  
*Euwallacea* sp. nr. *forficatus* in California

A thesis submitted in partial satisfaction  
of the requirements for the degree Master of Arts  
in Geography

by

Monica Dimson

2017

© Copyright by

Monica Dimson

2017

## ABSTRACT OF THE THESIS

Utility of temporally-biased invasive species distribution models in the detection of  
*Euwallacea* sp. nr. *fornicatus* in California

by

Monica Dimson

Master of Arts in Geography

University of California, Los Angeles, 2017

Professor Thomas Welch Gillespie, Chair

Species distribution models (SDMs) are valuable risk assessment tools in the management of invasive species, for which early detection and containment are critical. Few studies have evaluated the utility of invasive SDMs trained on temporally-biased presence data. While the abundance and range of occurrence records may increase with time after invasion, management objectives become more difficult to achieve as a destructive species nears establishment. This research assesses the relative predictive ability of models for the invasive shot hole borer (Coleoptera: Curculionidae: Scolytinae: *Euwallacea* sp. nr. *fornicatus*), which was first detected in Southern California in 2003. A series of 100-meter resolution models were developed in Maxent, selected for its ability to produce reliable models with relatively few occurrence records. Models were trained using data from five chronologically cumulative sampling periods, which

simulate stages of invasion. The effects of spatial extent and spatial filtering were also examined. All models achieved high AUC (area under the receiver operating curve) values  $> 0.93$  and correctly classified  $87.7 \pm 18.8\%$  of independent test records, indicating high model performance regardless of the degree of temporal bias. The leading contributing variables were minimum temperature of the coldest month (for sixteen models) or percent impervious surface (for four models). Sensitivity was consistently higher for models that used the larger spatial extent, which suggests that for an emerging species, larger backgrounds may be less restrictive on model outcomes. Spatial filtering produced more discriminating results without compromising model sensitivity. The study finds that invasive SDMs can be useful in identifying areas vulnerable to invasion, particularly if they are integrated into adaptive management strategies.

The thesis of Monica Dimson is approved.

Glen Michael MacDonald

Yongwei Sheng

Thomas Welch Gillespie, Committee Chair

University of California, Los Angeles

2017

## TABLE OF CONTENTS

1. Introduction
2. Background
  - 2.1. Species Impact
  - 2.2. Study Area
  - 2.3. Modeling Method
3. Methods
  - 3.1. Species Data
  - 3.2. Environmental Variables
  - 3.3. Model Parameters
  - 3.4. Model Evaluation
4. Results
  - 4.1. Suitable Area
  - 4.2. Model Sensitivity
  - 4.3. Area Under the Receiver Operating Curve
  - 4.4. Variable Contribution
5. Discussion
  - 5.1. Utility of Temporally-Biased Models
  - 5.2. Effects of Spatial Filtering and Spatial Extent
  - 5.3. Limitations of the AUC
  - 5.4. Important Predictors of ISHB
6. Conclusion
7. Tables and Figures
8. Appendix
9. References

## TABLES AND FIGURES

Table 1	Environmental variables
Table 2	Description of model series
Table 3	Minimum training presence thresholds
Table 4	Area predicted suitable
Table 5	Model sensitivity results
Table 6	Area under the receiver operating curve results
Table 7	Mean variable contribution
Table 8	Mean, minimum, and maximum values of leading contributing variables
Figure 1	Spatial extents
Figures 2a-2e	California unfiltered model outputs
Figures 3a-3e	California filtered model outputs
Figures 4a-4e	Ecoregion unfiltered model outputs
Figures 5a-5e	Ecoregion filtered model outputs
Figure 6	One-way ANOVA results
Figure 7	Two-way ANOVA results

## ACKNOWLEDGEMENTS

I would like to thank the Eskalen Lab at University of California, Riverside; University of California Cooperative Extension, Orange County; and USDA Forest Service, Forest Health Protection for contributing their invasive shot-hole borer survey data. Funding for the UC Cooperative Extension and UC Riverside surveys was provided by Orange County Parks and the California Avocado Commission, respectively.

## 1. INTRODUCTION

Species distribution models (SDMs) describe the relationship between species occurrence and environmental conditions at a given location. SDMs have a wide variety of applications, including risk assessment for biological invasion (Peterson 2003; Elith et al. 2010; Jiménez-Valverde et al. 2011). Invasive species can have a significant economic and ecological impact outside of their native range. In the United States, the annual cost of invasive species management is estimated at 219 billion USD (Pimentel 2011). Invasive species also threaten native biodiversity through interspecific interactions (e.g. competition, predation, parasitism) and habitat modification (Pimentel et al. 2005). Modeling the distribution of invasive species can help to identify areas that are most vulnerable to invasion, particularly when resources for full systematic surveys are limited.

Invasive species are typically in disequilibrium with a novel environment, which presents a challenge for correlative models like Maxent (Elith et al. 2010; Uden et al. 2015). A modeling study of incipient *Phytophthora ramorum* found that invasive SDMs are less accurate in the early stages of invasion and more likely to under-predict ranges (Vaclavik and Meentemeyer 2012). It takes time for a species to establish and realize its full potential range, and models trained on records from the onset of invasion will likely be affected by spatial (e.g. from uneven survey effort) and temporal (e.g. surveying during ongoing species expansion or contraction) sampling bias (Bean et al. 2012).

Several alternative approaches have been proposed to address the issue of disequilibrium. These include correlative models that project predictions from the native range onto introduced sites, mechanistic models based on species physiology, hybrid models that select environmental predictors according to species physiology, or collection of input data via planned, transect

sampling (Peterson 2003; Elith et al. 2011; Jiménez-Valverde et al. 2011; Vaclavik and Meentemeyer 2012; Brummer et al. 2013). Unfortunately, these data and methods are rarely readily available, and are often costly and time-consuming to acquire or implement.

There is a need to further examine the utility of correlative SDMs in the early stages of biological invasion, despite the implicit temporal bias. Many invasive SDM studies have focused on species that are well-established at the time of study (Welk et al. 2004; Lippitt et al. 2008; Elith et al. 2010; West et al. 2016). Though predictive ability has been found to increase as species near establishment, these studies do not address early detection, prevention, and containment objectives critical to successful invasive species management (Epanchin-Niell and Liebhold 2015). It is impractical to wait until a destructive species has spread and established before attempting to predict its potential extent. As newly introduced species approach equilibrium in the invaded range—itsself a difficult threshold to identify—efficient control and total eradication become near impossible to achieve (Venette et al. 2010; Uden et al. 2015).

In this research, I modeled the distribution of the invasive shot hole borer (Coleoptera: Curculionidae: Scolytinae: *Euwallacea* sp. nr. *fornicatus*) (ISHB), an emerging invasive species in California that threatens commercial avocado industries, native forests, and urban landscapes (O'Donnell et al. 2016). A series of SDMs were produced using Maxent, a common correlative distribution model that consistently ranks among the most reliable species distribution modeling methods (Elith et al. 2006; Phillips et al. 2006; Wisz et al. 2008). Using ISHB survey data collected from 2012 to 2016, five chronologically cumulative sampling periods were created to simulate stages of invasion. Each sampling period was modeled 1) with and without spatial filtering and 2) using two different spatial extents. These additional parameters were examined

because the Maxent model is especially sensitive to sampling bias and can be greatly impacted by the size of the background area (VanDerWal et al. 2009; Elith et al. 2011).

The objectives of this study were to 1) determine the relative predictive ability of models trained on temporally-biased ISHB data, 2) examine the effects of spatial extent and spatial filtering on invasive SDM outcomes, and 3) determine how SDMs for emerging invasive species can serve early detection and prevention management objectives.

## **2. BACKGROUND**

### **2.1. Species Impact**

The common name invasive shot hole borer (ISHB) refers to two closely related ambrosia beetles that belong to genetically distinct clades of the *Euwallacea fornicatus* species complex (Stouthamer et al. 2017). The beetles are native to separate but overlapping regions of Asia. Each vectors a fungal symbiont (*Fusarium* spp.) that infects the xylem of the host plant and also serves as the insect's food source (Eskalen et al. 2013). The mutualistic beetle/disease complex causes limb failure, branch dieback, and in severe cases, tree mortality.

ISHB was first detected in 2003 in Los Angeles County, but its potential impact was not recognized until 2012 (Lynch et al. 2016). By 2013, ISHB/*Fusarium* had been observed on 112 tree species (Eskalen et al. 2013). Over fifty host species support both fungal growth and beetle reproduction. These reproductive hosts include nineteen California native species, six invasive plants, and many trees common in the urban forest. ISHB is also a pest of avocado (*Persea americana*) (Stouthamer et al. 2017), a major agricultural crop in California valued at 274 million USD (California Department of Food and Agriculture 2016). Avocado groves have been affected in Israel as well, where ISHB was first recorded in 2005 (Mendel et al. 2012).

Control tactics for ambrosia beetles are limited. ISHB feeds and mates within the tree galleries and does not appear to use an aggregation pheromone (Umeda et al. 2016). Infection by *Fusarium* inhibits the host plant's ability to respond to chemical treatments (Eatough Jones et al. 2017). Early detection and containment are therefore critical management strategies for ISHB.

## **2.2. Study Area**

The current range of ISHB in California now includes Los Angeles, Orange, and San Diego Counties, as well as limited regions of Riverside, San Bernardino, Ventura, Santa Barbara, and San Luis Obispo Counties. ISHB models were created at two spatial extents: 1) the state of California and 2) an ecological sub-set defined by Level III Ecoregions (Fig. 1). Ecoregions describe areas of similar ecosystem type, resource quality and availability, and probable response to disturbance (Griffith et al. 2016). The Ecoregion hierarchy classifies ecosystems at different spatial scales, providing an adaptable framework for ecosystem research and management.

The geographical and biological diversity of California divides the state into 13 Level III ecoregions. In Southern California, the Mediterranean climate produces cool, wet winters and warm, dry summers (Franklin 1998). This climate regime supports some of the highest levels of regional biodiversity in the world; Mediterranean California is home to 4,300 native plant species, 35% of which are endemic to the region (Cowling et al. 1996). Urbanization and agriculture are the main threats to regional diversity.

The urban forest of Los Angeles is also extremely diverse and supports a higher number of tree species (both native and exotic) than some native forests in the U.S. (Gillespie et al. 2016). Many of these species are susceptible to ISHB/*Fusarium*.

The modeled subset includes Ecoregions 8 and 85. Ecoregion 85, the Southern Californian/Northern Baja Coast, is an area characterized by coastal sage scrub and chaparral, as well as major urban areas and some agricultural land (Griffith et al. 2016). Ecoregion 8 describes the Southern California Mountains, an area of higher elevations, cooler summers, and higher precipitation than Ecoregion 85. It is also less developed, as much of the area is National Forest public land, and dominated by chaparral, oak woodlands, and coniferous woodlands.

### **2.3. Modeling Method**

Maxent is a machine-learning program that predicts species geographic distributions using species point data and environmental variables (Phillips et al. 2006). It performs well with limited occurrence records and produces robust results with small sample sizes, compared to other correlative SDM methods (Hernandez et al. 2006; Anderson and Gonzalez 2011). Maxent is a presence-only model and does not require field records of species absence. Instead, pseudo-absences are drawn from the background data (Phillips and Dudík 2008). This makes Maxent a valuable tool for modeling incipient species like ISHB, for which absence data in the novel range is often unavailable or unreliable (Phillips et al. 2006; Jiménez-Valverde et al. 2008).

Presence-only models have a few disadvantages. True-absences can be valuable in providing a more complete picture of sampling bias, and may also help to produce more robust predictions if they are truly representative of unsuitable habitat (Elith and Leathwick 2009). Inclusion of true-absences, together with dispersal constraints, has been shown to improve invasive SDM performance (Vaclavik and Meentemeyer 2009). However, absence from a given area is difficult to verify for mobile species (Jiménez-Valverde et al. 2008). Observed absences could equally signify that habitat is unsuitable, or that it is suitable but currently unoccupied

(Elith and Leathwick 2009; Jiménez-Valverde et al. 2011). For an emerging species like ISHB, it is difficult to confirm the exact significance of observed absence.

### **3. METHODS**

#### **3.1. Species Data**

Presence-only occurrence data for ISHB were obtained from University of California (UC) Riverside; UC Cooperative Extension, Orange County; and USDA Forest Service, Forest Health Protection. The dataset was compiled from reports and informal surveys that recorded ISHB observations in Southern California from 2012 to 2016. Point locations were collected using hand-held GPS devices and represent individual ISHB-infested trees. Infestation was confirmed by trained experts using direct beetle observations, beetle signs and/or symptoms of attack, and/or *Fusarium* infection.

The original dataset is the most spatially and temporally comprehensive dataset currently available and includes 9,530 ISHB presence records. By default, Maxent removes duplicate points that occur within a single grid cell, the size of which is determined by the spatial resolution of the environmental layers, in this case 100-meters (Fourcade et al. 2014). This reduced the dataset to 1,052 unique records. Though the ISHB point data are highly accurate and precise, they are affected by an unquantified degree of spatial sampling bias, which is primarily the result of uneven sampling effort. Much of the data was collected not during transect surveys, but during visual tree assessments within artificial boundaries where researchers have had permission to survey (e.g. county parks, riparian corridors, university campuses) (Umeda et al. 2016). This resulted in clusters of dozens of points within a single 100-meter grid cell. Spatial sampling bias is a common obstacle in species distribution modeling, and the utility of data

despite these limitations is a point of investigation in this study (Rushton et al. 2004; Newbold 2010).

### **3.2. Environmental Variables**

Bioclimatic variables were downloaded from WorldClim 1.4, a popular dataset in species distribution modeling because it provides publicly available, global coverage at a relatively fine spatial resolution of 1-kilometer (Hijmans et al. 2005). Rasters were resampled to 100-meters using bilinear interpolation, which has been used, with robust results, to resample 250-meter climate data to as fine a scale as 10-meters (Rovzar et al. 2016). Resampling from this relatively coarser resolution may obscure microclimates in areas of high topographic relief, however, the topographic data have a finer resolution and may help to preserve this variability.

Topographic, impervious surface, tree canopy cover, and NDVI (normalized difference vegetation index) data were acquired at a 30-meter resolution. These layers were aggregated by mean and resampled to 100-meter resolution using bilinear interpolation. Elevation, percent impervious surface, and percent tree canopy cover rasters were downloaded from the U.S. Geological Survey National Elevation Dataset and National Land Cover Database, respectively (Xian et al. 2011; Homer et al. 2015; USGS 2016). Aspect and slope were calculated from a mosaic of elevation rasters. Dry-season NDVI was calculated from August through September 2016 Landsat 8 scenes (data obtained from the U.S. Geological Survey's Earth Resources Observation and Science Center). NDVI is a remotely-sensed measure of greenness often used to quantify vegetation distribution and density.

Lastly, a categorical shapefile for geologic substrate was acquired from the U.S. Geological Survey (Ludington et al. 2005). Vector to raster conversion was performed in

ArcMap 10.4.1 using a 100-meter cell size and the Maximum Area setting, which assigns values based on the majority area of the destination cell. This technique has been shown to be more accurate than alternative methods (Congalton 1997).

ArcMap 10.4.1 was used to process all environmental layers to identical spatial resolution and extent, as required by Maxent. Rasters were projected to an equal area coordinate system (NAD 1983 California Teale Albers) before resampling. Outputs were then re-projected to WGS 1984 and extracted by the two spatial extents used in the study (Fig. 1). A resolution of 100 meters was selected so that results would be more applicable to management and conservation decision-making (Brummer et al. 2013).

Correlation analysis was performed for all continuous variables using the principal component analysis tool in ArcMap 10.4.1. Highly correlated variables ( $\geq|0.75|$ ) in the correlation matrix were removed so that only one of the variables remained. In selecting and refining the predictors, I prioritized variables of ecological relevance to insect development, host availability and conditions, and disturbance (Elith et al. 2011). Twelve variables were used as predictors for all ISHB models (Table 1).

The predictor set is limited in that it does not include biotic interactions that may influence beetle distribution. This is a common restriction on correlative SDMs, particularly for an invasive species that is relatively new to science. Little is known about ISHB's life history in terms of interspecific competition or predation, especially in California (Umeda et al. 2016), or about specific host characteristics that influence vulnerability.

Dry-season NDVI and land cover variables were intended to serve as proxies for host plant material, but host availability does not guarantee susceptibility. Likelihood of infestation is also influenced by host vigor and stand conditions (Paine et al. 1997), which can vary from year

to year. Because the bioclimatic variables are averaged over a thirty-year period, conditions that induce drought or moisture stress are not represented by these predictors.

### **3.3. Model Parameters**

Maxent version 3.3.3 k was used to create twenty ISHB models at a 100-meter spatial resolution (Phillips et al. 2004). I tested models using three parameters: length of sampling period, spatial extent, and spatial filtering (Table 2). Default settings and 10,000 background points were used for all models. Each model used in the analyses was the average of ten cross-validation replicates. This form of replication randomly splits the records into equal-size “folds,” each of which was used for model evaluation in one of the replicates. In this case, each run used nine of the folds for model training and the remaining fold—approximately 11% of the total sample size—for model testing.

#### **3.3.1. Length of Sampling Period**

In this research, temporal sampling bias was represented by the length of the sampling period from which ISHB records were taken. Five chronologically cumulative datasets, referred hereto after as sampling periods, were derived from the occurrence records. I refer to the models by the last year of data in the sampling period (e.g. models based on 2012-2015 records are called “2015 models”). There were four iterations of each sampling period, using different combinations of spatial extent and spatial filtering (described below).

Models with a longer sampling period had a larger training dataset (Table 2) that was potentially more environmentally representative of the ISHB niche. This research did not control for differences in sample size, the effects of which may be less important than those of sampling

bias (Bean et al. 2012; Boria et al. 2014). Improvements in model performance are more often due to spatial bias treatment, rather than larger sample size alone. Bean et al. (2012) observed that “well-sampled data with few records are better than biased data of any sample size” (p. 255). Sample size also has an inconsistent relationship with area under the receiver operating curve (AUC), a common SDM performance metric. AUC has been found to both increase with (Hernandez et al. 2006) and be relatively unaffected by changes in sample size (Wisiz et al. 2008; Anderson and Gonzalez 2011).

### **3.3.2. Spatial Extent**

Ideally, areas that have not been surveyed should be excluded from the background area in Maxent (Elith et al. 2011). In a presence-only model, selecting too large a background may produce misleading model results, including inflated performance statistics (VanDerWal et al. 2009). ISHB models were thus created using two spatial extents: 1) the state of California and 2) the external boundary of Ecoregions 8 and 85 (Environmental Protection Agency, Level III) (Fig. 1). The outputs are referred to as California and Ecoregion models, respectively. Ecoregions 8 and 85 were merged in ArcMap 10.4.1 to obtain a single outer boundary for the area. The small portion of Ecoregion 85 that extends into Mexico was clipped from the extent. The vast majority of ISHB records have been collected, coincidentally, from Ecoregion 85. Few areas outside of this extent have been surveyed, and this may have inflated some performance metrics for the California models. However, the political boundary of California was tested because it reflects the likely jurisdiction of policymakers. The future distribution of ISHB may concern resource managers at the state level, particularly if preventative strategies outside of Southern California are to be implemented.

### **3.3.3. Spatial Filtering**

A study by Yackulic et al. (2013) found that 87% of Maxent modeling studies are affected by spatial sampling bias. A spatially biased model is often fitted more closely to survey effort, rather than the actual distribution of a species (Phillips et al. 2009). Spatial filtering can help mitigate this bias by reducing clusters of geographic records. It has been shown to improve model predictive accuracy and improve representation of non- or less-surveyed areas in the modeled extent (Kramer-Schadt et al. 2013).

In order to examine the impact of spatial sampling bias on emerging, invasive SDMs, ISHB models were created both with and without spatial filtering. I used a method similar to the rarefied unfiltered treatment in Boria et al. (2014) and the systematic sampling treatment in Fourcade et al. (2014), in which records are subsampled at a coarser resolution than that of the model. To create the spatially filtered dataset, NOAA Biogeography Branch's Sampling Design Tool was used to randomly sample records from a 500-meter fishnet generated in ArcMap 10.4.1. This further reduced the number of ISHB observations to 372. Models that have been treated with this method are hereafter referred to as "filtered," versus "unfiltered."

### **3.4. Model Evaluation**

The primary metric of model performance in this study was sensitivity, or the proportion of true-presences correctly classified by each model. Sensitivity was calculated using 150 ISHB records that were excluded from model training (Appendix A). Thirty records were randomly selected from each survey year using NOAA Biogeography Branch's Sampling Design Tool.

The predicted values at each of the model-independent test points were extracted from the Maxent outputs using ArcMap 10.4.1. Correctly classified records were those points with

values greater than or equal to the minimum training presence (MTP) threshold of each respective model (Table 3).

Thresholds are often used to create binary maps of habitat suitability, and should be selected based on study objectives and restrictions (Hernandez et al. 2006; Holcombe et al. 2010). Fixed thresholds (e.g. threshold=0.5, 10 percentile training presence) omit a defined percentage of training records, which can help to account for varying quality of input data (Ward and Morgan 2014). The MTP, which sets the threshold to the lowest probability predicted for any of the true positives that trained the model, was selected because confidence in the precision and accuracy of the ISHB records is high.

In addition to sensitivity, MTP was used to calculate suitable area. The range of suitable values, based on MTP, were also calculated for the leading contributing variables.

Metrics based on specificity, the proportion of correctly classified true-absences, should be avoided for presence-only models (Hernandez et al. 2006; Merow et al. 2013). Errors of commission (true-absences classified as presences) are difficult to interpret, not only because true-absences are lacking, but because they could indicate sites where the species could exist and does not yet. For such species (e.g. rare, endangered, or invasive), which are known to presently occupy a limited part of their potential range, omission (true-presences classified as absences) is of greater concern than commission (Uden et al. 2015; Rovzar et al. 2016). Commission errors could also be caused by factors not included in the model, e.g. dispersal limitations and biotic interactions (Jiménez-Valverde et al. 2011).

One of the most common threshold-independent metrics used in model assessment is the area under the receiver operating curve (AUC), which relies on measures of both sensitivity and specificity (Elith and Leathwick 2009; Merow et al. 2013). AUC describes the probability that

the model can randomly predict a higher probability for a presence cell versus an absence (or pseudo-absence) cell (Phillips and Dudík 2008). Values range from 0 to 1.0, where  $AUC = 0.5$  indicates model predictive ability equal to that of random prediction. Models with  $AUC \geq 0.75$  are generally considered suitable for use in conservation planning (Elith et al. 2006).

Though AUC is the only measure of predictive accuracy in some SDM studies (Broennimann et al. 2007; Wisz et al. 2008; Brummer et al. 2013), it is not an ideal performance metric of presence-only models due to the lack of true-absence data (Lobo et al. 2008; Pearson 2010). In Maxent, AUC has been found to increase when using too large a background area (VanDerWal et al. 2009; Merow et al. 2013). AUC values remain high even when spatial sampling bias is deliberately added to a model, but also decrease when that same bias is corrected through spatial filtering (Kramer-Schadt et al. 2013; Syfert et al. 2013; Fourcade et al. 2014). Because alternative metrics for presence-only models are few, AUC continues to be used in model evaluation. However, it is recommended that a secondary metric be used to corroborate high values and justify confidence in the model (Warren and Seifert 2011; Merow et al. 2013).

This study used mean AUC as a comparative metric between models, rather than as an absolute measure of model performance. Four one-way analysis of variance (ANOVA) tests were used to compare the effect of different sampling periods on AUC within each model series. I refer to the model series by spatial extent and spatial filtering, where +F indicates filtered models (i.e. California, California+F, Ecoregion, Ecoregion+F). Two-way ANOVA was also used to compare models within the same sampling period (i.e. all 2012 models, all 2013 models, etc.). All ANOVA and post hoc Tukey tests were conducted in RStudio (RStudio Team 2016).

## 4. RESULTS

### 4.1. Suitable Area

Area predicted suitable increased with length of sampling period, all other parameters being equal (Table 4). Filtered models were generally more conservative than their unfiltered counterparts (Fig. 2a-2e, 3a-3e, 4a-4e, 5a-5e). Spatial filtering decreased area predicted suitable by  $35.2 \pm 10.3\%$  in the California extent and  $22.7 \pm 10.8\%$  in the Ecoregion extent (excluding 2012 Ecoregion models). Area predicted suitable was 5.8% greater for the filtered versus unfiltered 2012 Ecoregion model. Due to the areal difference of the spatial extents, the suitable area of the California models cannot be directly compared to that of the Ecoregion models.

### 4.2. Model Sensitivity

Mean sensitivity increased with length of sampling period, all other parameters being equal (Table 5). The largest increase in sensitivity occurred between the 2012 and 2013 models. On average, sensitivity of each 2013 model was  $42.9 \pm 6.9$  percentage points higher than that of its 2012 counterpart. Average sensitivity of the 2014, 2015, and 2016 models was consistently high at  $99.3 \pm 0.6\%$ .

Sensitivity was higher to test records from the modeled sampling period (e.g. 2013 models were more sensitive to 2012-2013 records than records after 2013). All models correctly predicted  $97.9 \pm 1.6\%$  of the 2012 test records. However, average sensitivity to the 2015 test records ( $77.5 \pm 29.8\%$ ) was lower than that of the 2016 test records ( $82.0 \pm 28.5\%$ ).

Spatial filtering did not have a consistent effect on sensitivity. The largest increase in sensitivity due to spatial filtering occurred among the 2012 models. Sensitivity of the 2012 California and 2012 Ecoregion models both increased by eight percentage points when training

records were filtered. The largest decrease occurred in the sensitivity of the 2013 Ecoregion models, which decreased by ten percentage points when records were filtered.

The effect of spatial extent on sensitivity depended on sampling period. Average sensitivity of the 2012, 2013, and 2014 Ecoregion models was  $5.7 \pm 4.3$  percentage points lower than that of their California counterparts. Among 2015 and 2016 models, however, the difference in sensitivity between California and Ecoregion models was less than one percentage point.

### **4.3. Area Under the Receiver Operating Curve**

All twenty ISHB models performed better than random prediction and obtained AUC greater than 0.93 (Table 6). Within each model series, AUC decreased with length of sampling period, which had a statistically significant effect on AUC at the  $p < 0.001$  level as determined by one-way ANOVA. Post hoc Tukey tests showed multiple pairwise differences among the AUC values of each model series (Fig. 6). The AUC of each 2012 model was statistically significantly higher than that of its 2016 counterpart ( $p < 0.001$ ).

In each sampling period, spatial extent had a statistically significant effect on AUC as determined by two-way ANOVA. The AUC of each California model was statistically significantly higher than that of its Ecoregion counterpart (e.g. 2015 California+F > 2015 Ecoregion+F) ( $p < 0.001$ ;  $p < 0.01$  for 2012 CA/Ecoregion pair) (Fig. 7).

All else being equal, there was no statistically significant effect of spatial filtering alone on AUC, as determined by two-way ANOVA, except in the 2016 model group. Post hoc Tukey tests showed no statistically significant difference between the AUC of an Ecoregion model and its filtered Ecoregion counterpart (Fig. 7). Among California models, spatial filtering increased AUC for the 2014 ( $p < 0.01$ ), 2015 ( $p < 0.01$ ), and 2016 ( $p < 0.001$ ) models only.

#### **4.4. Variable Contribution**

The leading contributing variables were minimum temperature of the coldest month (for sixteen models) or percent impervious surface (for four models) (Table 7). Both were among the top three contributing variables for nineteen out of twenty models. Also among the top three contributing variables were isothermality (more important in California models) and dry-season NDVI (more important in Ecoregion models). Slope and aspect contributed least to the models. Precipitation metrics, percent tree canopy cover, and geologic substrate contributed less than 4% to the models. Variable contribution differed somewhat between the two spatial extents.

Within the suitable area determined by MTP, mean minimum temperature of the coldest month was  $3.97 \pm 1.13^{\circ}\text{C}$  (Table 8). Mean suitable surface imperviousness was  $14.0 \pm 10.1\%$ , mean isothermality was  $52.9 \pm 2.5\%$ , and mean dry-season NDVI was  $0.19 \pm 0.04$ .

### **5. DISCUSSION**

The results of this study show that temporally-biased sampling data can be used to train high-performing SDMs for an incipient invasive species like ISHB, albeit with a limited shelf-life. All models achieved  $\text{AUC} > 0.93$  and correctly classified  $98.8 \pm 1.1\%$  of the independent test records from their sampling period. In discussing model predictive ability, I focus primarily on sensitivity and use AUC only as a relative measure of performance. Limitations of AUC are discussed below.

#### **5.1. Utility of Temporally-Biased Models**

Unsurprisingly, predictive ability was more robust for the time period on which the model was trained, then deteriorated with time after sampling period. This suggests that invasive

SDMs will more effectively meet detection and prevention aims if incorporated into an adaptive ecological management system, in which the results of management actions continuously inform subsequent decisions and strategies (Uden et al. 2015). There is no “best” or “final” model for an emerging invasive species. Model-informed surveys will yield additional species records, which should be added to training datasets as they become available. It is important to acknowledge that models have a temporal limit, but these actions may prolong the utility of the model.

Despite these limits, the ISHB models were also able to predict records from the year subsequent to the training sampling period, i.e. the 2012 models predicted  $60.0 \pm 14.7\%$  of the 2013 records; the 2013 models predicted  $94.6 \pm 7.9\%$  of the 2014 records; the 2014 models predicted  $87.5 \pm 7.4\%$  of the 2015 records; and the 2015 models predicted  $97.5 \pm 1.7\%$  of the 2016 records (Table 5). This sensitivity suggests that Maxent models for emerging invasive species are able to anticipate new occurrences up to a certain threshold. That threshold is likely species-dependent and requires further investigation.

The 2016 models performed best in terms of sensitivity. These models were trained on a larger and presumably more diverse dataset that covered a broader geographic extent. However, the 2015 test records were the most difficult to classify among all models, and yielded higher omission rates than the 2016 test records (Table 5). This suggests that fewer novel environments were sampled in 2016 than in 2015, despite the discovery of ISHB in new counties. SDMs are often fitted to environmental rather than geographic space (Phillips et al. 2009), which is why the 2012 and 2013 (Fig. 2a-b, 3a-b, 4a-b, 5a-b) models could identify suitable habitat in San Diego County before ISHB was observed there in 2014 (Fig. 2c, 3c, 4c, 5c). Sampling across geographic space is thus less important than sampling across environmental space (Vaughan and Ormerod 2003).

It can be difficult to anticipate monitoring needs for a new invasive species before its full, potential threat has been recognized. But the sensitivity results in this study indicate that biases in species occurrence data can be minimized by balancing survey efforts across environmental space as well as geographic space. If this results in inconsistent geographic coverage across the landscape, correlative distribution models may help to fill in the gaps.

ISHB continues to spread, but it is already well-established in many regions of Southern California. Detection and monitoring efforts are still limited to visual surveys and, more recently, trapping using the lure quercivorol, which is attractive to several other species of ambrosia beetle (Carrillo et al. 2015; Umeda et al. 2016). SDMs can contribute to management efforts for ISHB and other incipient invasive species in a cost-effective, adaptive way. The results of this study show that these models were able to reliably detect new occurrences, and that this ability would be complemented by concurrent field surveys in novel environmental conditions. The average 2012 model was trained on the most biased data, but predicted over half of the 2013 test records (Table 5). Such predictions may have aided early containment of ISHB and identified vulnerable areas in and immediately surrounding Los Angeles.

## **5.2. Effects of Spatial Filtering and Spatial Extent**

Spatial filtering is recommended if geographic clustering attributable to sampling bias occurs in the species presence data. Filtered models were generally more discriminating in terms of suitable area (Table 4), but did not inhibit sensitivity. Sensitivity of the 2014, 2015, and 2016 models was high ( $\geq 94\%$ ) for both the filtered and unfiltered series (Table 5). Spatial filtering also increased sensitivity of the 2012 models, for which sampling bias was highest, by 15.6%

(California extents) and 17.9% (Ecoregion extents). This suggests that filtered models can guide the identification of priority conservation areas without compromising model sensitivity.

It is generally recommended that presence-only SDMs utilize a more conservative spatial extent, so that pseudo-absences drawn from the background area do not inflate performance metrics (Lobo et al. 2008). In this study, however, each California model achieved higher sensitivity and AUC than that of its Ecoregion counterpart. The surveyed Ecoregions may have been too narrow an extent (VanDerWal et al. 2009) that fitted the model too closely to the training data. On the other hand, California may be an unnecessarily broad extent for a species like ISHB. There is likely an intermediate extent that would better serve ISHB detection. Including a buffer around the surveyed area may help to produce less restrictive predictions for invasive species that have yet to establish.

While it is clear that the 2016 models were the most successful (within each series), it is more difficult to identify the spatial extent and spatial filtering settings that optimize model performance. Among the series in this study, the filtered 2016 California model provided the most accurate predictions of ISHB presence. But the “best” model in any comparative assessment will be constrained to the parameters of the study. In other words, there may be an in-between spatial extent that better serves ISHB prediction. Selection of spatial extent, and the level of filtering, should be species- and data-oriented.

### **5.3. Limitations of the AUC**

Though I discuss AUC as a comparative metric within and among model series, I attribute no ecological significance to the absolute AUC values, especially as a standalone performance metric.

AUC should not be used to compare invasive SDMs affected by varying degrees of bias. In the ISHB models, sensitivity increased as length of sampling period expanded, but AUC values decreased (Table 6). The 2012 models had the highest AUC values, but the lowest sensitivity to all records after 2012. The 2016 Ecoregion and 2016 Ecoregion+F models, despite achieving 100% sensitivity to the independent test records, achieved the lowest AUC values.

Spatial filtering is known to reduce geographic bias in model results (Kramer-Schadt et al. 2013; Syfert et al. 2013), but the effect of spatial filtering on AUC was inconsistent (Fig. 7). AUC also failed to reflect the improved sensitivity of the 2012 filtered models. Selection of spatial extent should not be based on AUC alone, as large background areas often inflate AUC values (Lobo et al. 2008; Wisz et al. 2008; VanDerWal et al. 2009).

#### **5.4. Important Predictors of ISHB**

The two most important predictors of ISHB were minimum temperature of the coldest month and percent impervious surface. Relative variable contribution differed by spatial extent, so that the third leading contributor was isothermality for the California models and dry-season NDVI for the Ecoregion models. The climate variables are likely linked to beetle physiology, while the land cover variables help to describe host availability.

The importance of the minimum temperature of the coldest month variable suggests that ISHB, like many species in the subfamily Scolytinae, is freeze-susceptible (Lombardero et al. 2000). Temperature minimums are often an important determinant of insect distributions (Ungerer et al. 1999). Cold temperatures may induce mortality and constrain reproduction, development, flight, and overwintering survival (Lombardero et al. 2000; Sidder et al. 2016). For ISHB, the lower limit of suitable minimum temperature was  $0.36 \pm 2.56^\circ\text{C}$  (Table 8), but care

should be taken in projecting this lower limit to other regions, as cold tolerance of other Scolytinae has been shown to vary geographically (e.g. due to microclimates) (Régnière and Bentz 2007; Sidder et al. 2016).

Isothermality, a measure of diurnal temperature oscillations relative to the annual temperature range, was also an important climatic predictor. Suitable isothermality ranged from  $43.8 \pm 3.3$  to  $62.3 \pm 4.0\%$  (Table 8), suggesting a preference for milder climates with less short-term variation in temperature. This may help to explain the predicted distribution of ISHB along the coast of California (Fig. 2a-2e, 3a-3e, 4a-4e, 5a-5e).

Apart from percent impervious surface, the California models are dominated by climate variable contributions (Table 7). Isothermality (and to some extent, minimum temperature) may have been more prominent for the California extent because Ecoregions are, by definition, areas of similar climate. A more limited spatial extent may allow models to be more discriminating in terms of non-climatic variables, such as dry-season NDVI, which contributed significantly more to the Ecoregion models.

Dry-season NDVI and percent impervious surface were intended to describe host availability. Average suitable surface imperviousness ( $14.0 \pm 10.1\%$ ) is skewed towards the lower end of the range, where a higher density of host material could be expected (Table 8). The range maximum ( $85.0 \pm 22.3\%$ ) supports field observations that ISHB will readily infest hosts in the urban forest as well as natural areas. The high contribution of surface imperviousness also suggests that indicators of anthropogenic activity (disturbance/development) are highly relevant to invasive species SDMs, as they indicate the probability of introduction (Lippitt et al. 2008).

ISHB distribution cannot be explained solely by climate, especially if more discriminating predictions are required. Variables that describe host conditions are just as

influential, and potentially more so (DeRose et al. 2013), because ISHB occurrence is necessarily linked to host presence. Predictions of ISHB habitat suitability are effectively predictions of host availability, if not actual susceptibility (based on host vigor). It is therefore important to account for finer-scale temporal variations in local climate conditions that affect host vigor. This cannot be provided by the thirty-year averages of the WorldClim variables, but change in NDVI, rather than dry-season NDVI alone, may be serve as a useful predictor for future models.

## **6. CONCLUSION**

The results of this research indicate that SDMs can be used to predict new occurrences of an emerging invasive species, in spite of temporal sampling bias. Model performance can be improved, and temporal sampling bias mitigated, by approaching the modeling process through an adaptive management strategy that continuously incorporates new field data into updated models. Management objectives should be considered in the implementation of spatial filtering and selection of spatial extent, which can produce more or less discriminating predictions as well as influence relative contributions of environmental predictors.

It is recommended that AUC be used cautiously in the interpretation of presence-only SDMs. A sensitivity metric should be used to corroborate and provide context for AUC values. One limitation of the sensitivity results in this study is that the test records, while independent from model training, were randomly subsampled from the original ISHB dataset and thus potentially affected by the same biases as the training data (Boria et al. 2014). Field validation would provide completely independent test records, as well as temporally novel 2017 records with which to test the 2016 models. Ground-truthing would also yield both sensitivity and

specificity results. Collection of true-absence data is needed to determine whether models are over-predicting ISHB occurrence.

Whether the modeled species is emerging or established, invasive or native, SDMs are subject to bias in occurrence records. In a presence-only correlative model like Maxent, invasive SDMs will only be sensitive to environments represented by or similar to the training data. This may be inadequate if the goal is to describe the species' full, potential range or niche. If, however, the objective is to detect and protect environments with conditions similar to those that have already proven suitable, then the methods in this paper provide a framework for creating reliable invasive SDMs.

## TABLES AND FIGURES

**Table 1**

Descriptions of variables and rationale for inclusion in models.

<b>Variable</b>	<b>Rationale</b>	<b>Source</b>
Isothermality	Insect development	WorldClim 1.4
Max Temp. of Warmest Month	Insect development	WorldClim 1.4
Min Temp. of Coldest Month	Insect development	WorldClim 1.4
Temperature Annual Range	Insect development/ Host environmental conditions	WorldClim 1.4
Precipitation of Wettest Qtr.	Host environmental conditions	WorldClim 1.4
Precipitation of Driest Qtr.	Host environmental conditions	WorldClim 1.4
Slope	Host environmental conditions	USGS 3D Elevation Program
Aspect	Host environmental conditions	USGS 3D Elevation Program
Geologic substrate	Host environmental conditions	USGS Open-File Report 2005-1305
Percent Impervious Surface	Disturbance/Host availability	National Land Cover Database
Percent Tree Canopy Cover	Host availability	National Land Cover Database
Dry-season NDVI	Host availability	USGS EROS Center

**Table 2**

Descriptions of model outputs created in the study. Sampling period refers to the temporal range of occurrence records used in model training and testing. Ecoregion Spatial Extent refers to ecoregions 8 and 85 (EPA Level III).

<b>Model Series</b>	<b>Sampling Period</b>	<b>Spatial Extent</b>	<b>Spatial Filtering</b>	<b>Sample Size (Training records)</b>
<b>California unfiltered</b>	2012	California	No	120
	2012-2013	California	No	239
	2012-2014	California	No	530
	2012-2015	California	No	637
	2012-2016	California	No	1052
<b>California filtered</b>	2012	California	Yes	64
	2012-2013	California	Yes	116
	2012-2014	California	Yes	215
	2012-2015	California	Yes	267
	2012-2016	California	Yes	372
<b>Ecoregion unfiltered</b>	2012	Ecoregion	No	120
	2012-2013	Ecoregion	No	239
	2012-2014	Ecoregion	No	530
	2012-2015	Ecoregion	No	637
	2012-2016	Ecoregion	No	1052
<b>Ecoregion filtered</b>	2012	Ecoregion	Yes	64
	2012-2013	Ecoregion	Yes	116
	2012-2014	Ecoregion	Yes	215
	2012-2015	Ecoregion	Yes	267
	2012-2016	Ecoregion	Yes	372

**Table 3**

Minimum training presence (MTP) threshold of each model output.

<b>Model Series</b>	<b>Sampling Period</b>	<b>MTP</b>
<b>California unfiltered</b>	2012	0.1036
	2012-2013	0.048
	2012-2014	0.0209
	2012-2015	0.0287
	2012-2016	0.017
<b>California filtered</b>	2012	0.2534
	2012-2013	0.0571
	2012-2014	0.0501
	2012-2015	0.0686
	2012-2016	0.0296
<b>Ecoregion unfiltered</b>	2012	0.0532
	2012-2013	0.0293
	2012-2014	0.0159
	2012-2015	0.0158
	2012-2016	0.0073
<b>Ecoregion filtered</b>	2012	0.1386
	2012-2013	0.1181
	2012-2014	0.065
	2012-2015	0.0351
	2012-2016	0.0221

**Table 4**

Area predicted suitable by the minimum training presence threshold, expressed as a percentage of the total spatial extent. California models cannot be compared to Ecoregion models.

California Models			Ecoregion Models		
Sampling Period	Unfiltered	Filtered	Sampling Period	Unfiltered	Filtered
2012	2.1%	1.6%	2012	12.0%	12.8%
2013	4.7%	3.5%	2013	21.9%	13.3%
2014	8.1%	4.4%	2014	28.4%	21.1%
2015	9.1%	5.3%	2015	33.5%	29.7%
2016	12.9%	7.5%	2016	44.5%	37.9%
<b>Average</b>	7.4 ±3.7%	4.4 ±1.9%	<b>Average</b>	28.1 ±10.9%	23.0 ±9.7%

**Table 5**

Sensitivity of model outputs, calculated using the minimum training presence threshold. Thirty model-independent test records were randomly selected from each sampling year (columns indicate sensitivity to test records from that year only). Filtered models are indicated by +F.

Sampling Period	Model Series	% Correctly Classified Test Records						
		2012	2013	2014	2015	2016	All Records	In Sampling Period
2012	CA	96.7	53.3	50.0	20.0	36.7	51.3	96.7
	CA +F	96.7	76.7	53.3	40.0	30.0	59.3	96.7
	Ecoregion	96.7	43.3	50.0	13.3	20.0	44.7	96.7
	Ecoregion +F	100.0	66.7	50.0	23.3	23.3	52.7	100
2012 - 2013	CA	96.7	100.0	100.0	86.7	93.3	95.3	98.4
	CA +F	96.7	100.0	100.0	90.0	93.3	96.0	98.4
	Ecoregion	96.7	100.0	93.3	83.3	90.0	92.7	98.4
	Ecoregion +F	96.7	100.0	83.3	53.3	80.0	82.7	98.4
2012 - 2014	CA	96.7	100.0	100.0	93.3	100.0	98.0	98.9
	CA +F	96.7	100.0	100.0	90.0	96.7	96.7	98.9
	Ecoregion	96.7	100.0	100.0	90.0	93.3	96.0	98.9
	Ecoregion +F	100.0	100.0	100.0	76.7	93.3	94.0	100
2012 - 2015	CA	96.7	100.0	100.0	100.0	100.0	99.3	99.2
	CA +F	96.7	100.0	100.0	96.7	96.7	98.0	98.4
	Ecoregion	100.0	100.0	100.0	96.7	96.7	98.7	99.2
	Ecoregion +F	100.0	100.0	100.0	96.7	96.7	98.7	99.2
2012 - 2016	CA	96.7	100.0	100.0	100.0	100.0	99.3	99.3
	CA +F	100.0	100.0	100.0	100.0	100.0	100.0	100
	Ecoregion	100.0	100.0	100.0	100.0	100.0	100.0	100
	Ecoregion +F	100.0	100.0	100.0	100.0	100.0	100.0	100
<b>Average Sensitivity</b>		97.9	92.0	89.0	77.5	82.0	87.7	98.8
<b>Standard Deviation</b>		1.6	17.4	20.0	29.8	28.5	18.8	1.1

**Table 6**

Test area under the receiver operating curve (AUC) and standard deviation. Sampling periods are described by the last year of test data. Filtered models are indicated by +F.

Model Series	Sampling Period				
	2012	2013	2014	2015	2016
California	0.993 ±0.0015	0.986 ±0.0023	0.974 ±0.0026	0.970 ±0.0028	0.954 ±0.0032
California +F	0.994 ±0.0016	0.991 ±0.0022	0.987±0.0023	0.983 ±0.0028	0.978 ±0.0027
Ecoregions	0.977 ±0.0088	0.969 ±0.0084	0.958 ±0.0069	0.948 ±0.0078	0.933 ±0.0069
Ecoregions +F	0.966 ±0.0108	0.965 ±0.0111	0.954 ±0.0112	0.941 ±0.0135	0.935 ±0.0122

**Table 7**

Mean variable contribution to all twenty ISHB models (second column) and to individual model series.

Variable	Mean Contribution				
	All models	CA series	CA+F series	Ecoregion series	Ecoregion+F series
Min temp. coldest month	34.5 ±13.3	45.3 ±8.7	35.7 ±12	31.0 ±9.9	26.2 ±13.9
Percent impervious surface	24.9 ±9.5	19.1 ±7.1	28.0 ±11.8	21.8 ±5.2	30.6 ±7.8
Isothermality	13.7 ±6.3	18.0 ±4.5	18.9 ±4.0	10.9 ±3.3	7.2 ±4.1
Temperature annual range	5.9 ±3.3	8.4 ±3.3	6.8 ±3.2	3.2 ±1.3	5.4 ±2.7
Dry-season NDVI	5.4 ±5.4	0.6 ±0.1	0.9 ±0.5	12.0 ±3.0	7.9 ±3.7
Max temp. warmest month	4.9 ±2.6	5.2 ±2.4	5.1 ±0.6	5.6 ±4.0	3.6 ±1.6
Percent tree canopy cover	3.3 ±2.3	1.2 ±0.5	2.4 ±1.4	3.3 ±1.3	6.2 ±1.8
Geologic substrate	2.4 ±2.7	0.6 ±0.5	0.8 ±0.7	3.5 ±2.0	4.8 ±3.4
Precipitation of driest qtr.	2.0 ±2.8	0.5 ±0.5	0.4 ±0.2	4.3 ±3.7	2.9 ±2.3
Precipitation of wettest qtr.	1.9 ±2.2	1.0 ±1.2	0.8 ±0.8	2.7 ±2.8	2.9 ±2.5
Slope	0.8 ±0.9	0.1 ±0.1	0.1 ±0.1	1.2 ±0.5	1.8 ±0.9
Aspect	0.3 ±0.4	0.1 ±0.1	0.2 ±0.2	0.6 ±0.5	0.5 ±0.4

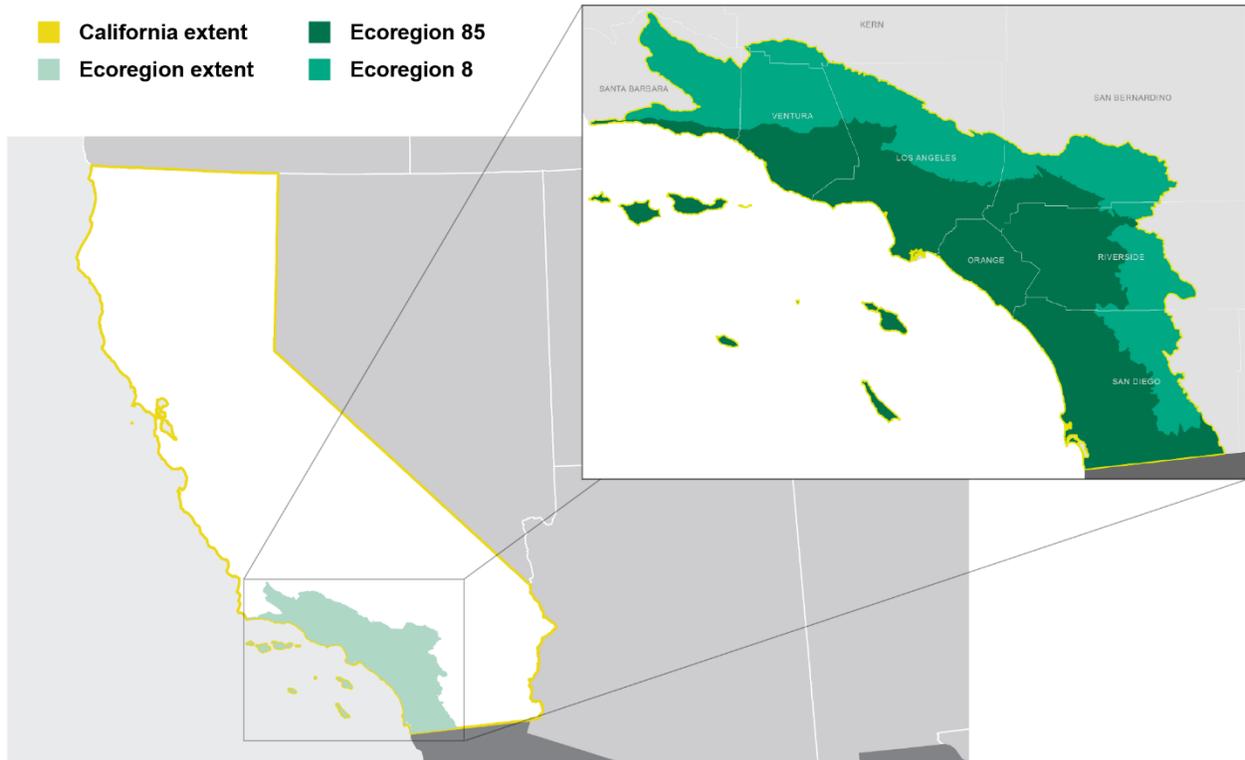
**Table 8**

Mean, minimum, and maximum values of the leading contributing variables across all ISHB models, calculated within the suitable area defined by the minimum training presence threshold.

Variable	Mean	Minimum	Maximum
Min temp. coldest month (°C)	3.97 ±1.13	0.36 ±2.56	7.27 ±2.18
Percent impervious surface	14.0 ±10.1	0 ±0	85.1 ±22.3
Isothermality (%)	52.9 ±2.5	43.8 ±3.3	62.3 ±4.0
Dry-season NDVI	0.19 ±0.04	-0.1 ±0.15	0.57 ±0.06

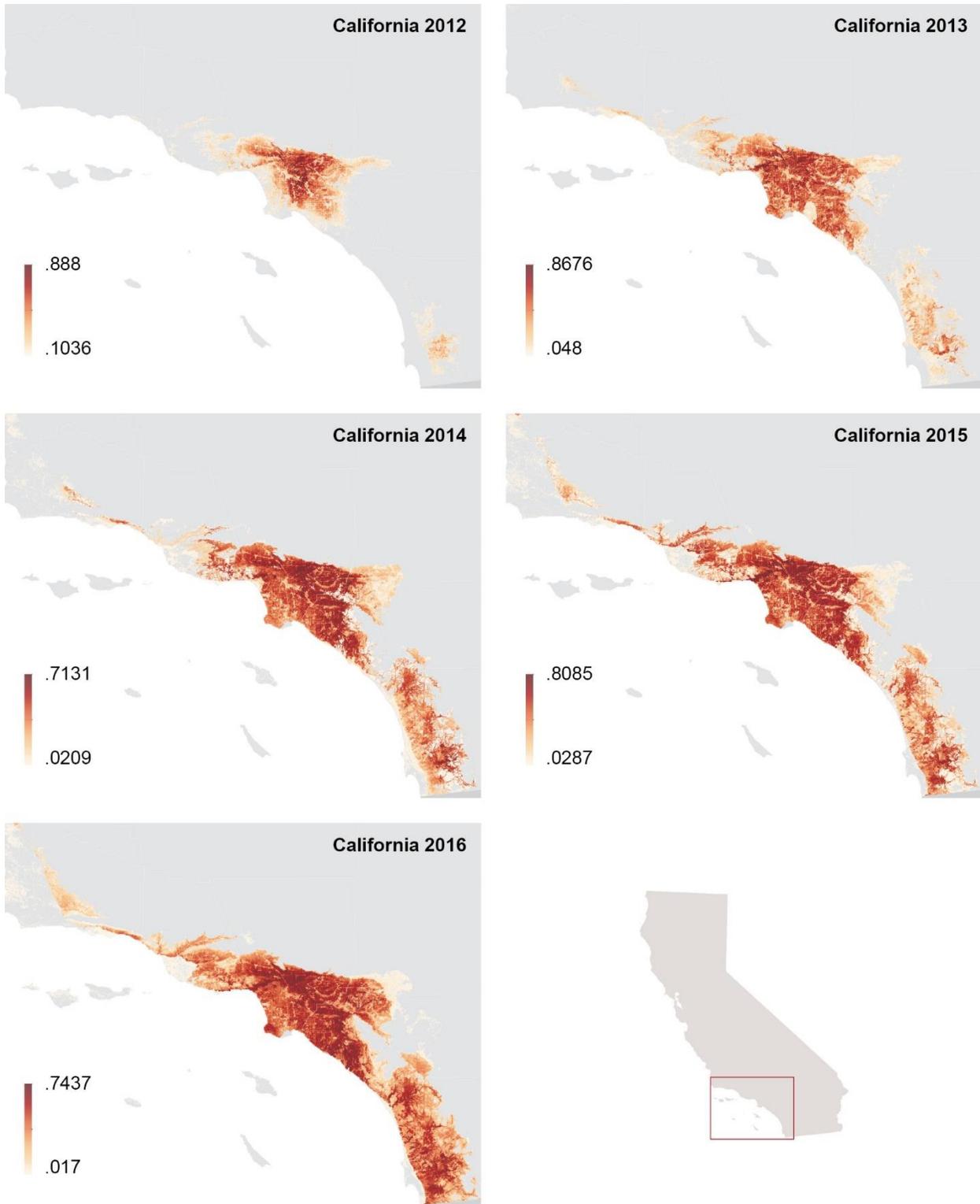
**Figure 1**

Model outputs were produced for two spatial extents: California and Ecoregions (8 and 85 combined).



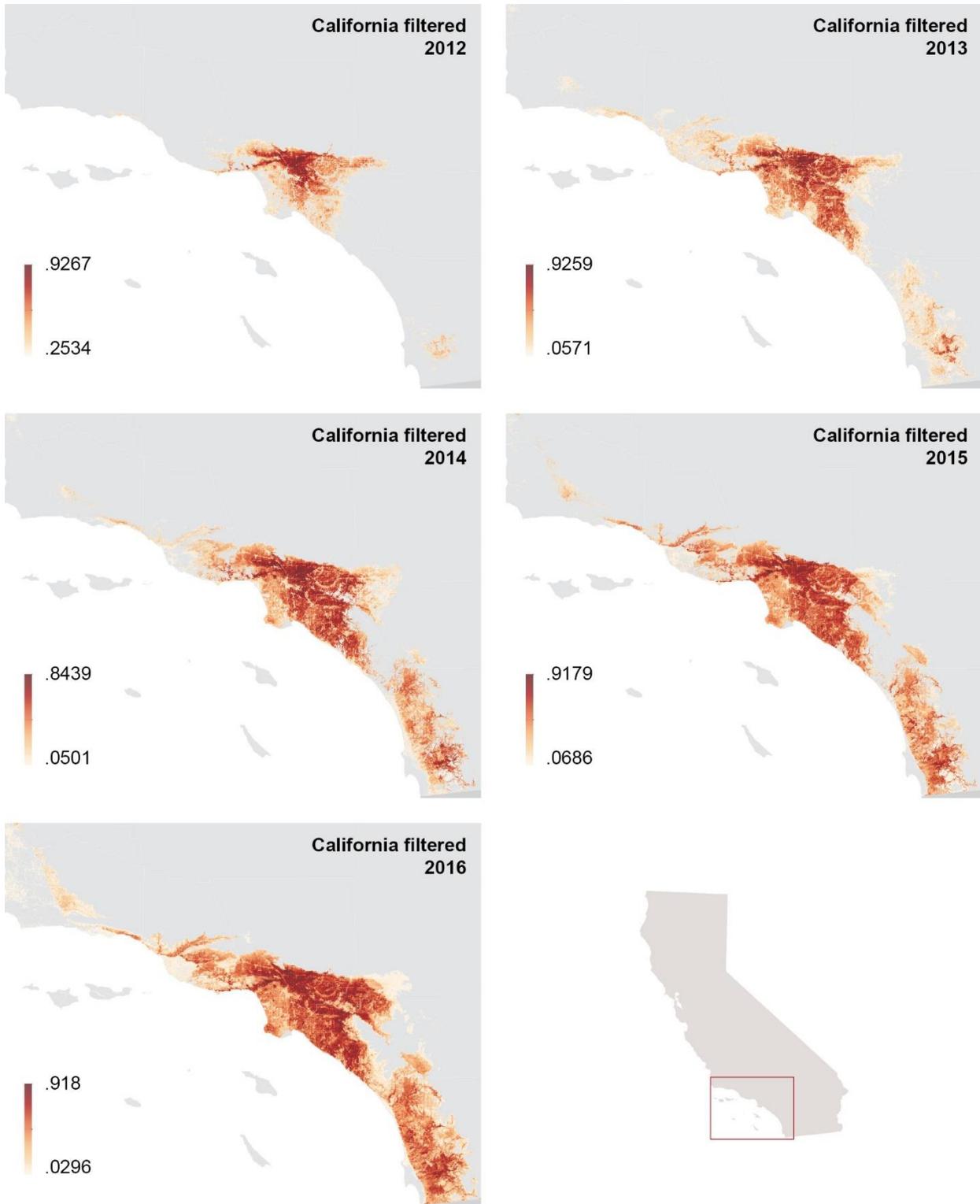
**Figures 2a-2e**

California unfiltered model outputs. Gray areas are unsuitable, based on the minimum training presence threshold.



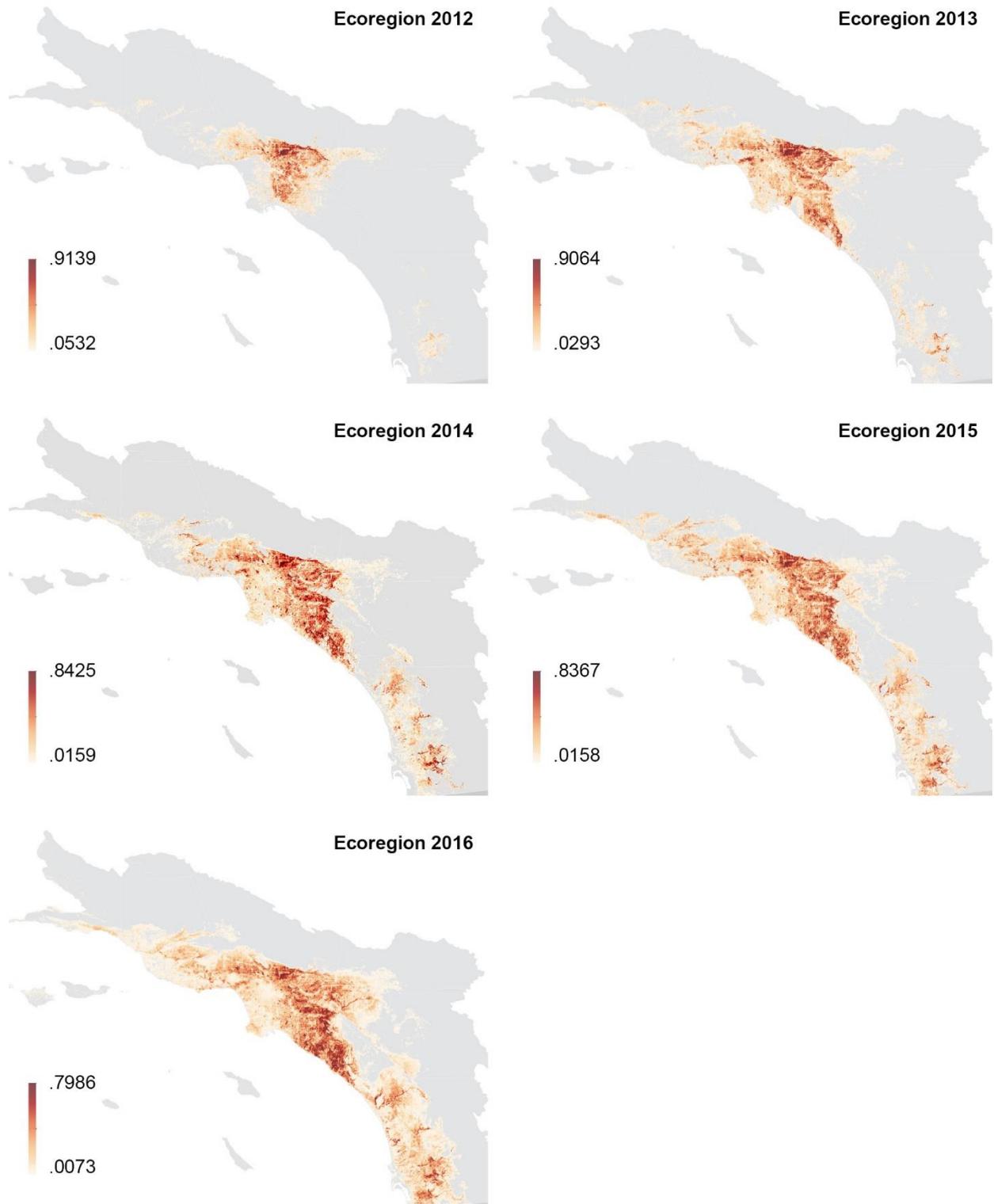
**Figures 3a-3e**

California filtered model outputs. Gray areas are unsuitable, based on the minimum training presence threshold.



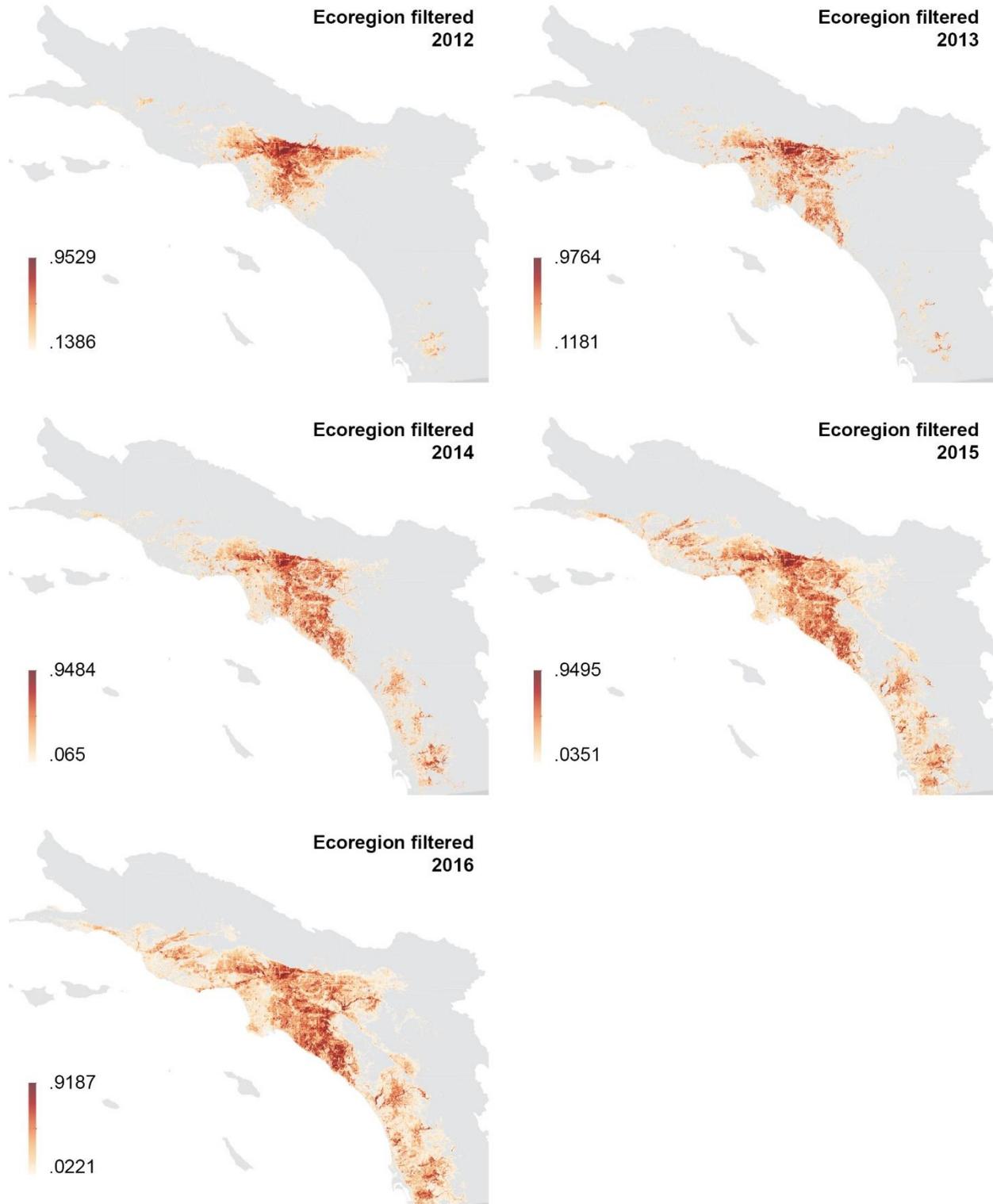
**Figures 4a-4e**

Ecoregion unfiltered model outputs. Gray areas are unsuitable, based on the minimum training presence threshold.



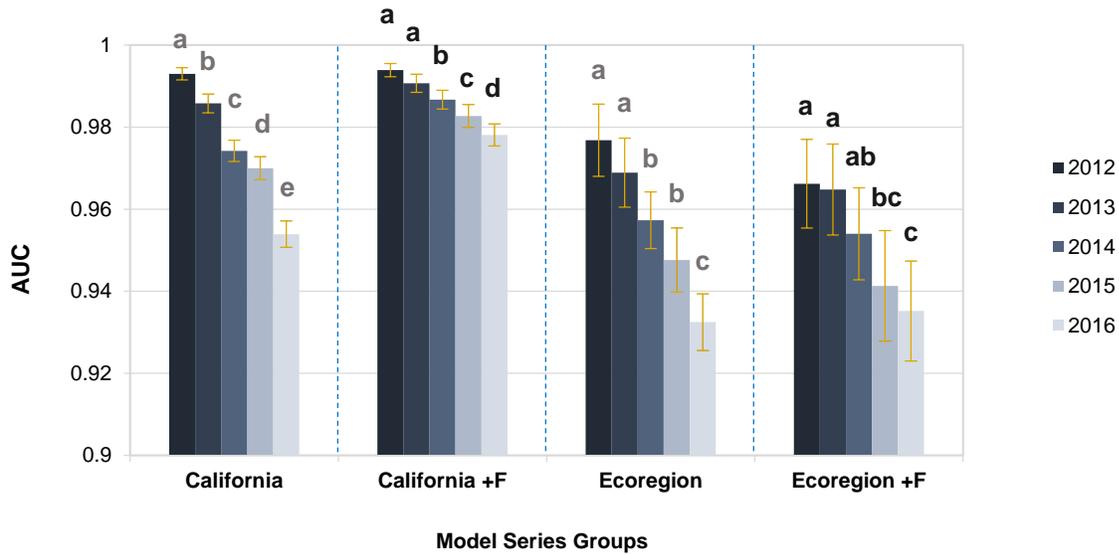
**Figures 5a-5e**

Ecoregion filtered model outputs. Gray areas are unsuitable, based on the minimum training presence threshold.



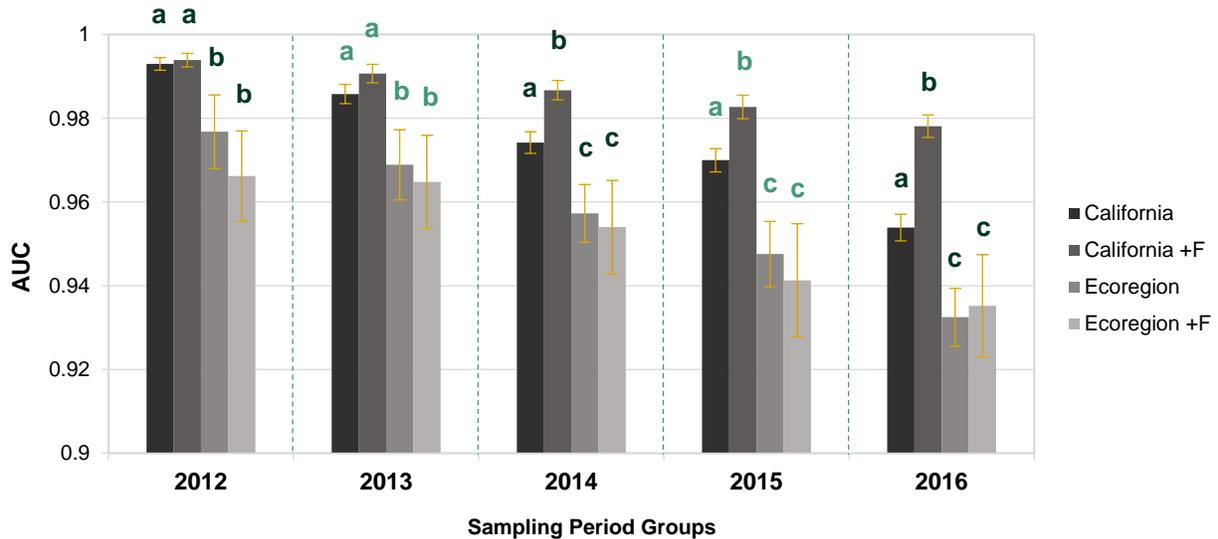
**Figure 6**

Mean area under the receiver operating curve (AUC). Models are grouped by series (spatial extent and spatial filtering). Shared letters indicate that there is no statistically significant difference between the sampling periods (within that model series only) as determined by one-way ANOVA and post hoc Tukey tests. Filtered models are indicated by +F.



**Figure 7**

Mean area under the receiver operating curve (AUC). Models are grouped by length of sampling period. Shared letters indicate that there is no statistically significant difference between those model series (within that sampling period only) as determined by two-way ANOVA and post hoc Tukey tests. Filtered models are indicated by +F.



## APPENDIX

Longitude, latitude, and year observed for 150 model-independent ISHB test records. Thirty records were randomly selected from each year of sampling. Records are listed by original identification number (OID).

OID	Longitude	Latitude	Year	OID	Longitude	Latitude	Year
0	-117.063398997	32.5506940184	2015	37	-117.797007186	33.6505815464	2016
1	-116.881833594	32.7742398614	2014	38	-117.820096428	33.6526805684	2016
2	-116.872387995	32.7784379054	2014	39	-117.802254741	33.6526805684	2016
3	-116.884982127	32.7794874164	2014	40	-117.801205230	33.6526805684	2015
4	-116.882883105	32.7794874164	2014	41	-117.800155719	33.6526805684	2015
5	-116.880784083	32.7794874164	2014	42	-117.831641049	33.6537300794	2013
6	-116.879734572	32.7826359494	2014	43	-117.830591538	33.6537300794	2013
7	-117.031913667	33.1216280024	2015	44	-117.829542027	33.6537300794	2013
8	-117.029814645	33.1226775134	2015	45	-117.828492516	33.6547795904	2013
9	-117.019319535	33.1405192004	2014	46	-117.802254741	33.6547795904	2015
10	-117.025616601	33.1436677334	2014	47	-117.801205230	33.6547795904	2015
11	-117.019319535	33.1436677334	2014	48	-117.835839093	33.6558291014	2013
12	-117.163102542	33.2591139434	2015	49	-117.833740071	33.6558291014	2013
13	-117.249162444	33.2633119874	2016	50	-117.828492516	33.6558291014	2013
14	-117.703600707	33.5435314244	2013	51	-117.811700340	33.6558291014	2016
15	-117.707798751	33.5498284904	2014	52	-117.833740071	33.6568786124	2014
16	-117.706749240	33.5508780014	2014	53	-117.821145939	33.6568786124	2016
17	-117.713046306	33.5519275124	2016	54	-117.813799362	33.6568786124	2016
18	-117.708848262	33.5519275124	2013	55	-117.811700340	33.6568786124	2016
19	-117.707798751	33.5519275124	2013	56	-117.590253519	33.6579281234	2015
20	-117.714095817	33.5529770234	2013	57	-117.643778580	33.6621261674	2016
21	-117.713046306	33.5529770234	2013	58	-117.943938726	33.6631756784	2015
22	-117.711996795	33.5529770234	2014	59	-117.943938726	33.6652747004	2015
23	-117.710947284	33.5529770234	2013	60	-117.624887382	33.6715717664	2016
24	-117.708848262	33.5529770234	2014	61	-117.650075646	33.6726212774	2016
25	-117.741383103	33.5571750674	2016	62	-117.664768800	33.6810173654	2015
26	-117.741383103	33.5582245784	2016	63	-117.916651440	33.6820668764	2013
27	-117.739284081	33.5582245784	2016	64	-117.940790193	33.7187997614	2014
28	-117.720392883	33.5907594194	2013	65	-117.938691171	33.7198492724	2013
29	-117.868373934	33.6295913264	2014	66	-117.943938726	33.7208987834	2016
30	-117.794908164	33.6358883924	2015	67	-117.939740682	33.7208987834	2013
31	-117.650075646	33.6390369254	2014	68	-117.938691171	33.7250968274	2014
32	-117.797007186	33.6400864364	2016	69	-117.950235792	33.7261463384	2016
33	-117.842136159	33.6442844804	2014	70	-117.950235792	33.7271958494	2015
34	-117.843185670	33.6453339914	2015	71	-117.937641660	33.7292948714	2014
35	-117.668966844	33.6453339914	2015	72	-117.951285303	33.7303443824	2016
36	-117.948136770	33.6474330134	2015	73	-117.899859264	33.7534336244	2013

OID	Longitude	Latitude	Year	OID	Longitude	Latitude	Year
74	-118.091919777	33.7765228664	2012	112	-117.972275523	33.9633358244	2012
75	-117.758175279	33.7975130864	2016	113	-117.430727847	33.9664843574	2016
76	-118.103464398	33.8248003724	2012	114	-117.984869655	33.9801280004	2012
77	-118.102414887	33.8248003724	2012	115	-117.979622100	33.9801280004	2012
78	-118.103464398	33.8258498834	2012	116	-117.408688116	33.9906231104	2016
79	-118.102414887	33.8258498834	2012	117	-118.050988848	34.0168608854	2012
80	-117.719343372	33.8667808124	2014	118	-117.641679558	34.0651383914	2012
81	-117.717244350	33.8667808124	2015	119	-118.024751073	34.0724849684	2012
82	-117.767620878	33.8678303234	2015	120	-117.760274301	34.0745839904	2012
83	-117.725640438	33.8678303234	2016	121	-118.113959508	34.1249605184	2012
84	-117.719343372	33.8678303234	2014	122	-118.111860486	34.1260100294	2012
85	-117.765521856	33.8688798344	2016	123	-118.116058530	34.1270595404	2012
86	-117.719343372	33.8688798344	2014	124	-118.113959508	34.1281090514	2012
87	-117.717244350	33.8688798344	2014	125	-118.155939948	34.1291585624	2014
88	-117.709897773	33.8688798344	2015	126	-118.117108041	34.1302080734	2012
89	-117.764472345	33.8699293454	2016	127	-118.112909997	34.1302080734	2012
90	-117.760274301	33.8699293454	2016	128	-118.111860486	34.1312575844	2012
91	-117.713046306	33.8699293454	2015	129	-118.110810975	34.1312575844	2012
92	-117.708848262	33.8699293454	2015	130	-118.123405107	34.1375546504	2015
93	-117.755026746	33.8730778784	2015	131	-118.059384936	34.1396536724	2013
94	-117.716194839	33.8730778784	2016	132	-118.050988848	34.1438517164	2012
95	-117.739284081	33.8793749444	2016	133	-118.164336036	34.1511982934	2013
96	-118.179029190	33.8856720104	2013	134	-117.906156330	34.1553963374	2012
97	-118.196870877	33.8930185874	2012	135	-118.040493738	34.1616934034	2013
98	-117.867324423	33.9014146754	2014	136	-118.299722955	34.1658914474	2014
99	-117.866274912	33.9024641864	2013	137	-118.078276134	34.1763865574	2012
100	-117.886215621	33.9077117414	2016	138	-118.097167332	34.1774360684	2012
101	-117.883067088	33.9077117414	2016	139	-118.079325645	34.1795350904	2013
102	-117.889364154	33.9087612524	2014	140	-118.078276134	34.1795350904	2012
103	-118.368990681	33.9161078294	2015	141	-118.078276134	34.1826836234	2012
104	-117.836888604	33.9203058734	2013	142	-118.340653884	34.1963272664	2014
105	-117.831641049	33.9213553844	2013	143	-118.143345816	34.2015748214	2013
106	-117.829542027	33.9213553844	2013	144	-118.323861708	34.2613969484	2014
107	-117.832690560	33.9224048954	2013	145	-119.102598870	34.3296151634	2015
108	-117.831641049	33.9224048954	2013	146	-119.102598870	34.3380112514	2015
109	-118.194771855	33.9496921814	2012	147	-119.087905716	34.3390607624	2015
110	-117.940790193	33.9528407144	2012	148	-119.098400826	34.3527044054	2015
111	-117.941839704	33.9538902254	2012	149	-119.299906938	34.4272196864	2015

## REFERENCES

- Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, 222(15), 2796-2811.
- Bean, W. T., Stafford, R., & Brashares, J. S. (2012). The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography*, 35, 250-258.
- Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275, 73-77.
- Broennimann, O., Treier, U. A., Muller-Scharer, H., Thuiller, W., Peterson, A. T., & Guisan, A. (2007). Evidence of climatic niche shift during biological invasion. *Ecology Letters*, 10(8), 701-709.
- Brummer, T. J., Maxwell, B. D., Higgs, M. D., Rew, L. J., & Franklin, J. (2013). Implementing and interpreting local-scale invasive species distribution models. *Diversity and Distributions*, 19(8), 919-932.
- California Department of Food and Agriculture. (2016). *California Agricultural Statistics Review, 2015-2016*. Retrieved from: [www.cdffa.ca.gov/statistics](http://www.cdffa.ca.gov/statistics)
- Carrillo, D., Narvaez, T., Cossé, A. A., Stouthamer, R., & Cooperband, M. (2015). Attraction of *Euwallacea* nr. *fornicatus* (Coleoptera: Curculionidae: Scolytinae) to lures containing quercivorol. *Florida Entomologist*, 98(2), 780-782.
- Congalton, R. G. (1997). Exploring and evaluating the consequences of vector-to-raster and raster-to-vector conversion. *Photogrammetric Engineering and Remote Sensing*, 63(4), 425-434.
- Cowling, R. M., Rundel, P. W., Lamont, B. B., Arroyo, M. K., & Arianoutsou, M. (1996). Plant diversity in Mediterranean-climate regions. *Trends in Ecology & Evolution*, 11(9), 362-366.
- DeRose, R. J., Bentz, B. J., Long, J. N., & Shaw, J. D. (2013). Effect of increasing temperatures on the distribution of spruce beetle in Engelmann spruce forests of the Interior West, USA. *Forest Ecology and Management*, 308, 198-206.
- Eatough Jones, M., Kabashima, J., Eskalen, A., Dimson, M., Mayorquin, J. S., Carrillo, J. D., Hanlon, C. C., & Paine, T. D. (2017). Evaluations of Insecticides and Fungicides for Reducing Attack Rates of a new invasive ambrosia beetle (*Euwallacea* sp., Coleoptera: Curculionidae: Scolytinae) in Infested Landscape Trees in California. *Journal of Economic Entomology*, 110(4), 1611-1618.

- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Soberón, J., Williams, S., Wisz, M. S. & Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129-151.
- Elith J. and J. R. Leathwick. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677–97.
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330-342.
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E. and Yates, C. J. (2011), A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17, 43–57.
- Epanchin-Niell, R. S. & Liebhold, A. M. (2015). Benefits of invasion prevention: Effect of time lags, spread rates, and damage persistence. *Ecological Economics*, 116, 146-153.
- Eskalen, A., Stouthamer, R., Lynch, S. C., Twizeyimana, M., Gonzalez, A., & Thibault, T. (2013). Host Range of *Fusarium Dieback* and Its *Ambrosia Beetle* (Coleoptera: Scolytinae) Vector in Southern California. *Plant Disease*, 97(7), 938-951.
- Fourcade, Y., Engler, J. O., Rodder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS One*, 9(5), e97122.
- Franklin, J. (1998). Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, 9(5), 733-748.
- Gillespie, T. W., de Goede, J., Aguilar, L., Jenerette, D. E., Fricker, G. A., Avolio, M. L., Pincetl, S., Johnston, T., Clarke, L.W., & Pataki, D. E. (2016). Predicting tree species richness in urban forests. *Urban Ecosystems*, 1-11.
- Griffith, G. E., Omernik, J. M., Smith, D. W., Cook, T. D., Tallyn, E., Moseley, K., & Johnson, C. B. (2016). *Ecoregions of California: U.S. Geological Survey Open-File Report (2016–1021)*.
- Hernandez, P. A., Graham, C. H., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29(5), 773-785.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965-1978.

- Holcombe, T. R., Stohlgren, T. J., & Jarnevich, C. S. (2010). From Points to Forecasts: Predicting Invasive Species Habitat Suitability in the Near Term. *Diversity*, 2, 738-76.
- Homer, C. G., Dewitz, J. A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N. D., Wickham, J. D., & Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, 81(5), 345-354.
- Jiménez-Valverde, A., Lobo, J.M., & Hortal, J. (2008). Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, 14, 885–890.
- Jiménez-Valverde, A., Peterson, A. T., Soberón, J., Overton, J.M., Aragón, P., & Lobo, J.M. (2011). Use of niche models in invasive species risk assessments. *Biological Invasions*, 13(12), 2785-2797.
- Kramer-Schadt, S., Niedballa, J., Pilgrim, J.D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A.K., Augeri, D.M., Cheyne, S.M., Hearn, A.J., Ross, J., MacDonoald, D.W., Mathai, J., Eaton, J., Marshall, A.J., Semiadi, G., Rustam, R., Bernard, H., Alfred, R., Samejima, H., Duckworth, J.W., Breitenmoser-Wuersten, C., Belant, J.L., Hofer, H., & Wilting, A. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, 19(11), 1366-1379.
- Lippitt, C. D., Rogan, J., Toledano, J., Sangermano, F., Eastman, J. R., Mastro, V., & Sawyer, A. (2008). Incorporating anthropogenic variables into a species distribution model to map gypsy moth risk. *Ecological Modelling*, 210(3), 339-350.
- Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17(2), 145-151.
- Lombardero, M. J., Ayres, M. P., Ayres, B. D., & Reeve, J. D. (2000). Cold tolerance of four species of bark beetle (Coleoptera: Scolytidae) in North America. *Environmental Entomology*, 29(3), 421-432.
- Ludington, S., Moring, B.C., Miller, R.J., Stone, P.A., Bookstrom, A.A., Bedford, D.R., Evans, J.G., Haxel, G.A., Nutt, C.J., Flynn, K.S., & Hopkins, M.J. (2005). *Preliminary integrated geologic map databases for the United States, USGS Open-File Report (2005-1305)*. Retrieved from: [pubs.usgs.gov/of/2005/1305](https://pubs.usgs.gov/of/2005/1305)
- Lynch, S.C., Eskalen, A., & Gilbert, G.S. (2016). New Pest-Disease Complex Threatens California Forests. *Cal-IPC News*, 24(2).
- Mendel, Z., Protasov, A., Sharon, M., Zveibil, A., Yehuda, S.B., O'Donnell, K., Rabaglia, R., Wysoki, M., & Freeman, S. (2012). An Asian ambrosia beetle *Euwallacea fornicatus* and its

- novel symbiotic fungus *Fusarium* sp. pose a serious threat to the Israeli avocado industry. *Phytoparasitica*, 40(3), 235-238.
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36(10), 1058-1069.
- Newbold, T. (2010). Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography*, 34(1), 3-22.
- O'Donnell, K., Libeskind-Hadas, R., Hulcr, J., Bateman, C., Kasson, M. T., Ploetz, R. C., Konkol, J. L., Ploetz, J. N., Carillo, D., Campbell, A., Duncan, R. E., Lyanage, P. N. H., Eskalen, A., Lynch, S. C., Geiser, D. M., Freeman, S., Mendel, Z., Sharon, M., Aoki, T., Cossé, A. A., & Rooney, A. P. (2016). Invasive Asian *Fusarium* – *Euwallacea* ambrosia beetle mutualists pose a serious threat to forests, urban landscapes and the avocado industry. *Phytoparasitica*, 44(4), 435-442.
- Paine, T. D., Raffa, K. F., & Harrington, T. C. (1997). Interactions among scolytid bark beetles, their associated fungi, and live host conifers. *Annual Review of Entomology*, 42(1), 179-206.
- Pearson, R.G. 2010. Species' distribution modeling for conservation educators and practitioners. *Lessons in Conservation*, 3, 54-89.
- Peterson, A. T. (2003). Predicting the Geography of Species' Invasions via Ecological Niche Modeling. *The Quarterly Review of Biology*, 78(4), 419-433.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231-259.
- Phillips, S. J., Dudík, M., & Schapire, R. E. (2004, July). A maximum entropy approach to species distribution modeling. In *Proceedings of the Twenty-First International Conference on Machine Learning*. ACM.
- Phillips, S. J. and M. Dudík. 2008. Modeling of Species Distributions with Maxent: New Extensions and a Comprehensive Evaluation. *Ecography*, 31(2), 161-175.
- Phillips, S. J., Dudík, 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(1), 181-197.
- Pimentel, D. (Ed.). (2011). *Biological invasions: economic and environmental costs of alien plant, animal, and microbe species*. CRC Press.

- Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52(3), 273-288.
- Régnière, J., & Bentz, B. (2007). Modeling cold tolerance in the mountain pine beetle, *Dendroctonus ponderosae*. *Journal of Insect Physiology*, 53(6), 559-572.
- Rovzar, C., Gillespie, T. W., & Kawelo, K. (2016). Landscape to site variations in species distribution models for endangered plants. *Forest Ecology and Management*, 369, 20-28.
- RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL [www.rstudio.com](http://www.rstudio.com).
- Rushton, S. P., S. J. Ormerod, and G. Kerby. 2004. New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41(2), 193-200.
- Sidder, A. M., Kumar, S., Laituri, M., & Sibold, J. S. (2016). Using spatiotemporal correlative niche models for evaluating the effects of climate change on mountain pine beetle. *Ecosphere*, 7(7).
- Stouthamer, R., Rugman-Jones, P., Thu, P.Q., Eskalen, A., Thibault, T., Hulcr, J., Wang, L., Jordal, B.H., Chen, C., Cooperband, M., Lin, C., Kamata, N., Lu, S., Masuya, H., Mendel, Z., Rabaglia, R., Sanguansub, S., Shih, H., Sittichaya, W., & Zong, S. (2017). Tracing the origin of a cryptic invader: phylogeography of the *Euwallacea fornicatus* (Coleoptera: Curculionidae: Scolytinae) species complex. *Agricultural and Forest Entomology*.
- 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), e55158.
- Uden, D. R., Allen, C. R., Angeler, D. G., Corral, L., & Fricke, K. A. (2015). Adaptive invasive species distribution models: a framework for modeling incipient invasions. *Biological Invasions*, 17(10), 2831-2850.
- Umeda, C., Eskalen, A., & Paine, T. D. (2016). Polyphagous Shot Hole Borer and *Fusarium* Dieback in California. In *Insects and Diseases of Mediterranean Forest Systems* (pp. 757-767). Springer International Publishing.
- Ungerer, M. J., Ayres, M. P., & Lombardero, M. J. (1999). Climate and the northern distribution limits of *Dendroctonus frontalis* Zimmermann (Coleoptera: Scolytidae). *Journal of Biogeography*, 26(6), 1133-1145.
- U.S. Geological Survey (2016). The National Map, 3D Elevation Program. Retrieved from: [nationalmap.gov/3DEP](http://nationalmap.gov/3DEP).

- Václavík, T., & Meentemeyer, R. K. (2009). Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling*, 220(23), 3248-3258.
- Václavík, T., & Meentemeyer, R. K. (2012). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73-83.
- VanDerWal, J., Shoo, L. P., Graham, C., & Williams, S. E. (2009). Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, 220(4), 589-594.
- Vaughan, I. P., & Ormerod, S. J. (2003). Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conservation Biology*, 17(6), 1601-1611.
- Venette, R.C., Kriticos, D. J., Magarey, R. D., Koch, F. H., Baker, R. H. A., Worner, S.P., Raboteaux, N.N.G., McKenney, D.W., Dobesberger, E.J., Yemshanov, D., De Barro, .P.J., Hutchison, W.D., Fowler, G., Kalaris, T.M., & Pedlar, J. (2010). Pest Risk Maps for Invasive Alien Species: A Roadmap for Improvement. *BioScience*, 60(5), 349-362.
- Ward, D., & Morgan, F. (2014). Modelling the impacts of an invasive species across landscapes: a step-wise approach. *PeerJ*, 2, e435.
- Warren, D.L. & Seifert, S.N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21(2), 335–342.
- Welk, E. (2004). Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns: Purple loosestrife (*Lythrum salicaria*) in North America. *Ecological Modelling*, 179(4), 551-567.
- West, A.M., Kumar, S., Brown, C.S., Stohlgren, T.J., & Bromberg, J. (2016). Field validation of an invasive species Maxent model. *Ecological Informatics*, 36, 126-134.
- Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., & Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14(5), 763-773.
- Xian, G., Homer, C., Dewitz, J., Fry, J., Hossain, N., & Wickham, J. (2011). The change of impervious surface area between 2001 and 2006 in the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 77(8), 758-762.
- Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H., & Veran, S. (2013). Presence-only modelling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution*, 4(3), 236-243.