



Iterative species distribution modeling results in the discovery of novel populations of a rare cold desert perennial

Israel T. Borokini^{1,2,*}, Kenneth Nussear³, Blaise Petitpierre⁴, Thomas E. Dilts⁵, Peter J. Weisberg⁵

¹Ecology, Evolution and Conservation Biology graduate program, Department of Biology, University of Nevada, Reno, Reno NV 89557, USA

²Department of Integrative Biology, University of California, Berkeley, Berkeley CA 94720, USA

³Department of Geography, University of Nevada, Reno, Reno NV 89557, USA

⁴Institute of Earth Surface Dynamics, University of Lausanne, Lausanne 1015, Switzerland

⁵Department of Natural Resources and Environmental Science, University of Nevada, Reno, Reno NV 89557, USA

ABSTRACT: Niche modeling for rare and range-restricted species can generate inaccurate predictions leading to an overestimation of a species geographic distribution. We used an iterative ensemble modeling approach and model-stratified field surveys to improve niche model formulation and better understand the ecological drivers of *Ivesia webberi* distribution. *I. webberi* is a US federally threatened herbaceous species, narrowly distributed in the western Great Basin Desert. Niche models for *I. webberi* were fitted using 10 replicates each of 6 modeling algorithms, while geographical projections of habitat suitability were generated using weighted ensembles of models with optimal performance. The resulting model projections were used to guide field surveys for 5 yr, generating additional spatial data, which were added to the existing dataset for subsequent modeling. Model performance across iterations was investigated and niche differences in the spatial dataset were explored. Model-guided field surveys resulted in the discovery of several new locations of *I. webberi* and an expansion of the species known range by 63 km. Model performance was higher in the earlier overfitted niche models. Overfitting was corrected in the final models, and predicted habitat suitability reduced from 5.98 % in the 2015 model to 3.34 % in the 2020 model. Findings show that *I. webberi* niche is associated with biotic, topographic and bioclimatic variables. Furthermore, a partial overlap was observed between environmental conditions of the initial and the new locations (Schoener's $D = 0.47$), which can be decomposed into 93 % of niche stability. This indicates that the majority of the newly discovered locations are within the environmental niche of the initial data.

KEY WORDS: Habitat suitability · Iterative ensemble modeling · Niche overlap · Field validation surveys · Niche stability · *Ivesia webberi* · Great Basin Desert

1. INTRODUCTION

Limited empirical information on the geographical distributions of taxa (Wallacean shortfall; Whittaker et al. 2005) can impact the assessment of species rarity resulting in misguided conservation prioritiza-

tions (Coddington et al. 2009). Field surveys, especially those conducted using random sampling strategies, can generate additional biodiversity data to mitigate this; however, such surveys are costly, time-consuming, and ineffective for rare species because human resources are limited (Hirzel & Guisan 2002,

*Corresponding author: iborokini@berkeley.edu

Guisan et al. 2006). Therefore, scientists and conservation managers have considered other cost-effective methods to stratify and prioritize field surveys using, for example, expert opinion and quantitative niche modeling. Species distribution models (SDMs) can relate the occurrences of taxa to their ecological conditions to quantify the realized niche, i.e. species known locations due to environmental tolerance observed in the field (Hutchinson 1957). These SDMs generate geographic predictions of species habitat suitability that can be used to stratify and optimize sampling efficiency (Chiffard et al. 2020). Moreover, integrating the new spatial data from model-guided sampling can reduce spatial bias in subsequent modeling iterations, improve the predictive accuracy of SDMs for rare species, and reliably identify biologically relevant environmental factors (Singh et al. 2009, Rinnhofer et al. 2012).

Understanding the distribution of rare species is critical for effective conservation planning, but with few, incomplete and biased spatial data, it can be challenging to model the niches of rare species with high predictive accuracy (Hernandez et al. 2006, Wisz et al. 2008), a condition referred to as the rare species modeling paradox (Lomba et al. 2010). This is because fewer occurrence points in a spatial dataset indicates low prevalence, which weakens the analytical power of the models and inflates bias in SDMs (Vaughan & Ormerod 2003). Furthermore, correlative species distribution models include the underlying assumption that species are in equilibrium with their environment (i.e. temporal and spatial stationarity) and that all important and biologically relevant variables have been included in the niche model (Elith & Leathwick 2009). This presents challenges to modeling rare species because the inclusion of many predictors when occurrences are few can lead to model overfitting (Wisz et al. 2008, Jarnevich et al. 2015). Moreover, limited natural history knowledge makes predictor variable selection challenging and potentially subjective for rare species (Aranda & Lobo 2011). Consequently, poorly fit models and misjudgments of model predictions can lead to over- or underestimation of the species niche and result in poorly informed management decisions (Ramesh et al. 2017, Burns et al. 2020). Despite the development of several statistical methods to reduce prediction errors in SDMs, the most practical way is to increase occurrence data for rare species, which is inevitably linked with data collection during field surveys. Therefore, geographical predictions of SDM for rare species should not be treated as truth, but can be used as hypotheses for further ecological or biogeo-

graphical investigations (Stockwell & Peterson 2002, Jarnevich et al. 2015, Sofaer et al. 2019).

The discovery of new locations of targeted species from SDM-guided field surveys is well documented in the literature (e.g. de Siqueira et al. 2009, Williams et al. 2009, Särkinen et al. 2013, Burns et al. 2020). These novel discoveries underscore the importance of SDMs as an important conservation tool. SDMs have been used to evaluate the degree of species rarity (Broennimann et al. 2006) and identify areas that may serve as future climatic refugia (Sousa-Silva et al. 2014). Furthermore, SDMs are also used to advance scientific knowledge of species–environment relationships (Jiménez-Valverde et al. 2011) and identify niche-constraining environmental factors (Gorban et al. 2011). SDM predictions are often integrated into models of population and landscape genetics (e.g. Ikeda et al. 2017, Banerjee et al. 2019), and spatial phylogenetics (e.g. Thornhill et al. 2017). Beyond conservation uses, newly discovered occurrences may have significant ecological contributions to the understanding of the overall species niche. For example, additional occurrence points may be found either within the existing realized niche space or in areas with different ecological conditions, thus expanding the species environmental niche. The COUE (centroid shift, overlap, unfilling and expansion) framework can be used to quantify realized niches of species from different ranges and categorize the niche position of newly discovered occurrences (Broennimann et al. 2012). This framework has been used to investigate niche dynamics between the native and invaded ranges of invasive species (Broennimann et al. 2012, Strubbe et al. 2013), as well as niche evolution vs. conservatism between sister taxa (Villegas et al. 2021).

The aim of this study was to assess the benefits of using an iterative sampling approach that alternates between niche modeling and model-guided field surveys versus a presence/absence modeling approach using only data available at the onset of the study to predict the distribution of a rare plant (*Ivesia webberi* A. Gray). Therefore, we asked the following questions: (1) Which environmental variables determine the distribution of *I. webberi* and how does the species–environment relationship change with each iteration of the SDMs given additional spatial data? (2) Do additional distribution data alter habitat suitability map projections across modeling iterations? (3) Is the environmental niche conserved throughout the modeling iterations? (4) Do modeling iterations improve the predictive accuracy of species distribution models for *I. webberi*?

2. MATERIALS AND METHODS

2.1. Study species and study area

Ivesia webberi is a US federally listed threatened perennial forb restricted to the eastern foothills of the Sierra Nevada and the adjacent western edge of the Great Basin Desert. *I. webberi* was estimated to have originated between 1.3 and 3.8 million yr ago (Töpel et al. 2012) and may be one of the many Great Basin Desert neoendemic and phylogenetically young taxa that have not had enough time to fully colonize their range (Kraft et al. 2010, Thornhill et al. 2017). At the outset of our study, it was known from 23 spatially aggregated locations, occurring in or near ephemeral washes and dry forest meadow gaps in mostly gently sloped areas (Witham 2000). These ‘presence’ locations were visited multiple times between 2015 and 2020, and therefore are not prone to positional error. The locations have experienced varying degrees of biological invasion pressures from *Bromus tectorum*, *Taeniatherum caput-medusae* and *Poa bulbosa*, as well as disturbances from wildfires, cattle grazing and off-highway vehicle use.

The study extent was defined by a 60 km buffer from marginal ranges of populations known as of 2015. The species produces achenes which are not adapted for long-range dispersal; therefore, the study area was restricted in order to mask out expansive adjacent unsuitable areas of playas in the central Great Basin Desert. This modeling decision was guided by natural history, which indicates that the species grows in sparsely vegetated low sagebrush (*Artemisia arbuscula*) communities in mid-elevation areas of the western Great Basin Desert and the adjacent northern Sierra Nevada eastern foothills (Federal Register 2014). Climatic conditions in these sites are characterized by relatively mild winters and hot summers (Svejcar et al. 2017). Temperatures range from an average of -5.8°C in the winter to an average of 28°C in the summer, and annual precipitation varies between 25

and 33 cm, most of which falls as snow or rain during the winter months.

2.2. Distribution data

We began species distribution modeling in 2015 with 23 occurrence points and 758 absence points obtained from the Nevada Natural Heritage Program (NNHP; Table 1). The absence points represent areas where *I. webberi* was not detected during historical surveys by NNHP botanists and citizen scientists. Additional spatial points were added following iterative modeling and field validation cycles in predicted suitable habitats. In all modeling iterations, the absence points were thinned using a 7.5 km distance in *spThin* R package version 0.2.0 (Aiello-Lammens et al. 2015) to reduce the effects of spatial aggregation and mitigate low prevalence in the spatial dataset. Additionally, absence points within 5 km of an occurrence point were removed to avoid false negatives. The remaining absence points were merged with the presence points for niche modeling (Table 1).

2.3. Predictor variables

A total of 72 predictor variables describing edaphic, topographic, land cover, vegetative cover and climatic factors were assembled for fitting space distribution models (SDM) for *I. webberi* (see Table S1 in the Supplement at www.int-res.com/articles/suppl/n050p047_supp.pdf). To avoid overfitting and maintain a 1:10 ratio of predictor variables to occurrence points (Harrell et al. 1996), the full set of predictor variables was reduced to 6 uncorrelated predictors (Table 2) using a combination of the Kendall r correlation coefficient ($r > 0.6$), feature selection runs in *Boruta* R package version 4.0.0 (Kursa & Rudnicki 2010) and recursive feature elimination algorithm in *caret* R package version 6.0-78 (Kuhn 2008).

Table 1. Iterative niche modeling with increasing number of presence and absence points for *Ivesia webberi*

Year	Presence points	Raw absence points	Thinned absence	Predictor variables used for final modeling
2015	23	758	53	Perennial herbaceous cover, Topographic Position Index (TPI) and annual evapotranspiration
2018	26	1652	90	Perennial herbaceous cover, TPI and cosine aspect
2019	27	1881	75	Perennial herbaceous cover, TPI and cosine aspect
2020	32	2289	102	Perennial herbaceous cover, TPI and summer mean precipitation

Table 2. Descriptions of 6 uncorrelated predictor variables used to fit preliminary niche models for *Ivesia webberi*. The 3 predictor variables used for the iterative niche models were selected from this pool. All predictors were resampled to 30 m resolution

Predictor variable	Relationship with species
Cumulative actual evapotranspiration (AET)	An estimate of the amount of water removed from an area by both evaporation and transpiration. Cumulative AET, a direct predictor, is a proxy estimate of plant productivity
Cosine aspect	Higher values indicate north-facing slopes which receive less sunlight
Perennial herbaceous vegetative cover	A spatial vegetative cover delineation representing native grasses, perennial forbs and cacti, which includes areas of <i>I. webberi</i> distribution. It is considered a representation of biotic interactions and accounts for community assemblage in sites harboring <i>I. webberi</i>
Minimum monthly temperature	A direct predictor that potentially influences plant distribution (Araújo & Rozenfeld, 2014). Vegetative and seed regeneration of <i>I. webberi</i> are dependent on cold stratification that characterizes late winter and early spring seasons
Summer seasonal precipitation	A direct predictor that potentially influences plant distribution. Summer precipitation causes surface runoffs which facilitate localized gravity-enhanced seed dispersal and colonization of empty niches. Precipitation and temperature in winter and spring seasons influence the phenology of <i>I. webberi</i>
Topographic position index (TPI)	A scale-dependent variable describing the elevation of a cell in relation to the mean elevation of the neighboring cells. At the scale of 333 m, TPI distinguishes between mountains and valleys in the study area. The study area is characterized by topographic heterogeneity which can limit dispersal and distribution, and also act as proxy for microclimatic conditions

The climatic variables (cumulative actual evapotranspiration [AET], minimum monthly temperature and summer seasonal precipitation) were down sampled from the Parameter-elevation Relationships on Independent Slopes Model (PRISM) climatic data normals (1971–2000) (Daly et al. 2008), from 800-m to 30-m spatial resolution using the Climatic Water Deficit Toolbox (Dilts et al. 2015) and ordinary kriging. The cosine-transformed aspect, ranging from -1 (south-facing slope) to $+1$ (north-facing slope), was derived from slope using the formula: $\theta \times \cos(\alpha)$, where θ is slope (in percentage) and α is aspect (in radians), while slope was calculated from the 1 arc-second digital elevation models (DEM; USGS 2017). Perennial herbaceous vegetative cover, a vegetation type raster layer, was obtained from the Multi-Resolution Land Characteristics (MRLC) development of the 2016 US National Land Cover Database (NLCD; Xian et al. 2013). Topographic Position Index (TPI) was calculated from the DEM using the formula described by Weiss (2001).

2.4. Iterative ensemble niche modeling and model-based sampling

The SDMs were fitted at 30 m resolution to capture the landscape and ecological heterogeneity in the study area, particularly in the *I. webberi* locations that

occur within forest gaps. An ensemble modeling approach was used in all niche modeling iterations. The use of multi-algorithm ensemble models renders predictions less susceptible to biases, assumptions or limitations of any individual algorithm while broadening the types of environmental response functions that can be identified (Araújo & New 2007). SDMs have been developed from a wide range of modeling techniques including regression, classification and machine learning algorithms (Lauzeral et al. 2012). Because these algorithms have different predictive performances under different contingencies (Li & Wang 2013), fitting of niche models using different algorithms and combining their model parameters to build a consensus or ensemble model is often recommended (Marmion et al. 2009). Ten replicates of 6 algorithms (Boosted Regression Trees, Random Forests, Maximum Entropy, Artificial Neural Networks, Generalized Additive Models and Generalized Linear Models) were fitted using the *biomod2* R package (Thuiller et al. 2009). All statistical packages were implemented in R statistical software version 4.0.2 (R Core Team 2020). See Table S2 for modeling details.

Model performance was evaluated using 4 metrics: (a) true skill statistic (TSS; Allouche et al. 2006), (b) area under the curve (AUC) of the receiver operating characteristics plot (Hanley & McNeil 1982), (c) TSS-based specificity and (d) Boyce Index (Boyce et al. 2002). In each modeling iteration, 3 predictors, se-

lected from the 6 uncorrelated variables, were used to fit the niche models. Models were fitted with 80% of the data with 20% used for k-fold cross-validation (Araújo et al. 2005, Thuiller et al. 2009). Model replicates with TSS ≥ 0.7 were averaged into ensemble models which were used to produce geographic projections of habitat suitability (Marmion et al. 2009, Thuiller et al. 2009). On the habitat suitability maps, cells with ≥ 0.5 occurrence probability were considered suitable to delineate areas with higher habitat suitability values for field validation surveys. Uncertainty in habitat suitability projections was visualized on maps of coefficients of variation from the iterative niche ensemble models (Hortal 2008).

Habitat suitability maps produced by the SDMs were used to guide field validation surveys to areas of high predicted probability of occurrence. The non-thinned absence points were overlaid on the predicted habitat map and predicted suitable and unsuitable areas that had not been previously surveyed were selected for field validation. To increase chances of detection, field validation surveys were done between May and June of each year when the plants were in bloom. Additional spatial data from the field surveys were used in the subsequent modeling iteration and site selection for post-modeling field validation. The iterative modeling and field surveys were repeated for 5 yr. For each newly discovered population, we calculated the distance to the nearest previously known occurrence with the *FNN* R package version 1.1.3 (Beygelzimer et al. 2019).

The relative importance of the predictor variables in all iterative SDMs was evaluated using the jackknife test (Phillips et al. 2006), while species–environment relationships were described with partial response curves using the evaluation strip method (Elith et al. 2005) as implemented in the *biomod2* R package. We assessed the trends and statistical significance of the model performance across the years of iterative niche modeling to investigate whether additional spatial data improved the overall predictive accuracy of the iterative ensemble SDMs. Mean scores of the 4 model performance metrics for each of the 6 algorithms (10 replicates each) were regressed against the years of iterative SDMs using multivariate multiple linear regression (MMLR). The statistical significance of the MMLR models was corrected using the Tukey post-hoc test.

We also assessed the reliability of these iterative SDM predictions by checking for model overfitting with a spatial cross-validation approach using block partitioning. Spatial block partitioning is a nonrandom allocation of spatial data to reduce the effect of

spatial bias and autocorrelation in ecological models (Valavi et al. 2019). The entire study area was divided into 6 equal latitudinal and longitudinal bins, which were then clustered into 3 spatial blocks. Two spatial blocks were used for model training, while the third block was used for testing. Spatial block partitioning was done in the *blockCVR* package version 2.1.4 (Valavi et al. 2019), whereas the niche models were conducted in *biomod2* R package version 3.5.1 using similar model tuning as used for the iterative SDMs. Partitioning our relatively small spatial dataset could only meet the requirements for modeling with Random Forest, Maximum Entropy and Artificial Neural Networks, which were then used for the spatial block-based ensemble niche modeling. Overfitting was assessed as the difference between the block (training) and test AUC values (Warren & Seifert 2011).

2.5. Assessment of the change in *I. webberi* niche across modeling iterations

We used the COUE framework to investigate the position of the new locations relative to the initial niche of *I. webberi*. The COUE framework, based on the principal component analysis (PCA), allows for direct comparison of species–environment relationships (Broennimann et al. 2012). To calculate the niche metrics, a kernel density function is applied to smoothen the varying sampling sizes of the 2 sets of occurrence points within a PCA gridded environmental space (Broennimann et al. 2012, Petitpierre et al. 2012). We calculated niche overlap (Schoener's *D*), stability, expansion and unfilling between the initial (2015) and a combination of all new (2018–2020) *I. webberi* locations, based on the environmental space of the 6 uncorrelated predictor variables. Schoener's *D* is calculated from the environmental occupancy of the 2 niches and it ranges from 0 (no overlap) to 1 (total overlap). No overlap and total overlap represent niche divergence and similarity, respectively (Brown & Carnaval 2019). In this study, niche stability represents the proportion of the environmental space in the newly discovered locations available in the initial occurrences' environmental space, whereas niche expansion represents the proportion of the environmental space in the new locations that are not available in the initial locations. A niche unfilling estimate was used to investigate whether the new occurrences only colonized a limited portion of the environmental space of the initial occurrences (Petitpierre et al. 2012,

Guisan et al. 2014). Given that habitat suitability map projections of the initial occurrences were used for field validation surveys, the COUE framework was implemented using a niche similarity test, which assumes that the environmental niches in the new occurrences are similar to the initial occurrences (Liu et al. 2020, Pili et al. 2020). The niche similarity test generated random estimates of Schoener's D , niche stability, expansion and unfilling, using 1000 randomizations of the niche positions of the initial and newly discovered occurrences. These randomizations were used to check if the observed niche overlap and stability were higher and if observed niche expansion and unfilling were lower than expected by chance. Furthermore, we extracted the initial niche density values at the new locations to quantify the degree of niche stability or expansion in the new locations. Niche density values range from 0 to 1, with 0 representing new locations outside the initial niche (i.e. niche expansion) and 1 representing new locations in the core of the initial niche. The niche similarity test was run with the development version 3.2.1 of the *ecospat* R package available on *github* (Di Cola et al. 2017).

Additionally, we quantified the number of predicted suitable raster cells (≥ 0.5 probability of occurrence) in the habitat suitability maps. We also performed a niche overlap analysis on the geographic projections of habitat suitability between the 2015 (initial) and 2020 (final) model iterations for each iterative niche model using the I similarity metric, which is based on Hellinger distance (Warren et al. 2008). The I similarity metric ranges from 0 to 1, representing the degree of pairwise similarity in niche model projections. This map-based niche overlap test is a cell-by-cell comparison with

a randomization test of geographical predictions of the 4 iterative SDMs (Warren et al. 2008) and it was performed in the *dismo* R package (Hijmans et al. 2017).

3. RESULTS

3.1. Environmental variables associated with the ecological niche of *Ivesia webberi* and species–environment relationship change across the space distribution models (SDMs)

Throughout the iterative SDMs from 2015 to 2020, the perennial herbaceous vegetative cover consistently contributed the most to the fitted distribution of *I. webberi* (Fig. 1). In the 2020 model iteration, *I. webberi* showed an asymmetric and threshold response curve for perennial vegetative cover, with suitable sites occurring in areas with moderate ($>20\%$) to high native perennial forb cover (Fig. 2a). TPI was the second most important predictor across all model iterations. The response curve for TPI is bimodal and asymmetric, illustrating that *I. webberi* occurs on sites that are either gentle lateral valleys or ridges (Fig. 2c). Cumulative AET was the third most important predictor in the 2015 niche model iteration (Fig. 1a). The cosine-transformed slope aspect, a proxy for exposure to sunlight, came third in the 2018 and 2019 iterations (Fig. 1c,d), while summer seasonal precipitation was the third most important predictor for the 2020 iteration (Fig. 1d). The response curve for summer seasonal precipitation shows a threshold response, where the probability of *I. webberi* occurrence was maximized at >25 mm summer precipitation, beyond which the curve flattened (Fig. 2b).

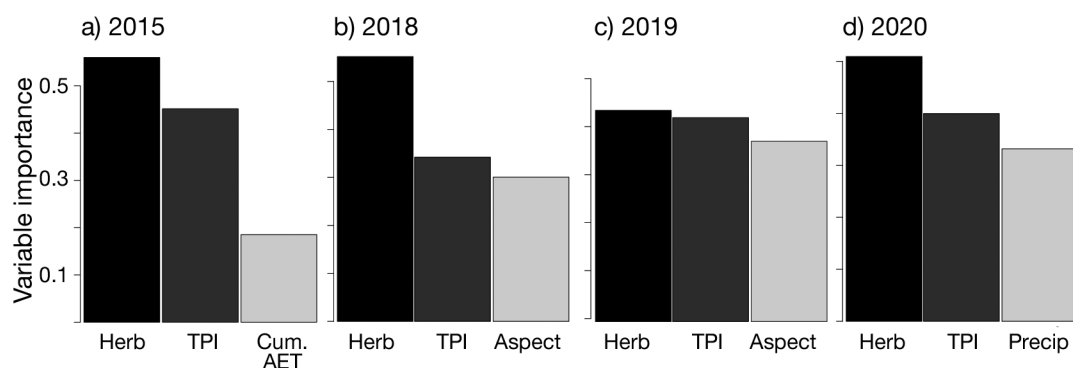


Fig. 1. Variable contributions to the iterative niche modeling for *Ivesia webberi* from (a) 2015 to (d) 2020. The 3 predictors used for each year of iterative modeling were selected from the preliminary modeling. Herb: perennial herbaceous vegetative cover; TPI: Topographic Position Index at 333 m; Cum. AET: cumulative actual evapotranspiration; Aspect: cosine-transformed aspect; Precip: summer mean precipitation

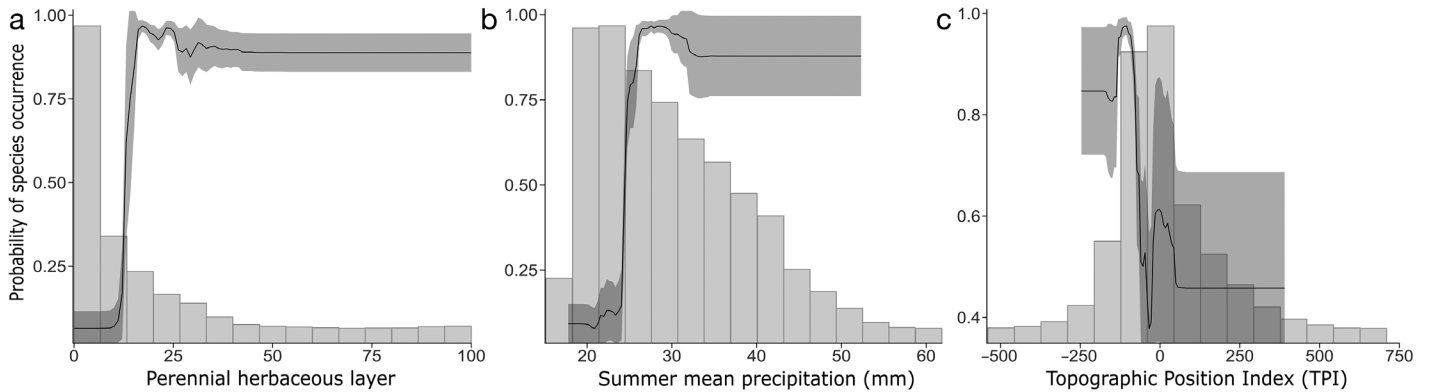


Fig. 2. Partial response plots showing the predicted probability of *Ivesia webberi* occurrence in (a) Perennial herbaceous vegetative cover, (b) summer mean precipitation and (c) Topographic Position Index. The partial response plots were generated using the Boosted Regression Trees, while the histograms represent the predicted values from 10 000 randomly sampled background points from the 3 variables used for the niche modeling. The black line represents the average response for the 10 modeling replicates, and the grey shading shows the SD. The partial response plots for each of the 10 model replicates of the 6 SDM algorithms are included in Fig. S1 in the Supplement

3.2. Impact of modeling iterations and field surveys on the distribution and ecological niche of *I. webberi*

The iterative ensemble SDMs and model-guided field surveys resulted in the discovery of 7 new locations of *I. webberi* (30.4 % of the initial dataset), while 2 additional new locations (8.7 % of the initial dataset) were discovered opportunistically by local botanists. The distance from the new locations to the closest known locations ranged from 30 m to 63 km (Table 3). As a result, the northern distribution range of the species was expanded by 63 km (Table 3). However, the percentage of the suitable raster cells in the ensemble habitat projections decreased from 5.98 % in 2015 to 3.34 % in 2020 (Fig. 3). Despite the decrease in the percentage of suitable grid cells, niche overlap be-

tween the geographical projections of the 2015 and 2020 model iterations was high (Hellinger's $I = 0.89$). The model projections also predicted higher probability of *I. webberi* occurrence in locations near the center of the study area (Fig. 3). Prediction uncertainties (coefficients of variation) were relatively low across all 4 projections (Fig. 4).

3.3. *I. webberi* niche dynamics across modeling iterations

The first 2 principal component analysis (PCA) axes, both representing topo-climatic gradients, explained 49 % of the variation in the data (Fig. 5, Table S3), while the third axis, representing the perennial vegetation cover, explained an additional 17.4 % of the

Table 3. Niche density, predicted habitat suitability and distance of new locations to the nearest neighbor in the initial points

Location name	Finding	Year	Niche density	Predicted habitat suitability	Distance from known location (km)
Wildcat Hill	Opportunistic: discovered by Bureau of Land Management (BLM) staff during land surveys	2018	0.27	0.73	8.07
Unit 6 extension	Model: predicted suitable sites near known location	2018	0.87	0.59	0.38
Smoke Creek Road	Opportunistic: discovered during California Native Plant Society vegetative surveys	2019	0.86	0.37	62.98
Unit 4 extension	Model: high predicted suitability	2020	0.38	0.62	0.03
South end of HJWA	Model: high predicted suitability	2020	0.58	0.64	2.99
HJWA south end #2	Model: suitable sites near known location	2020	0.71	0.30	2.26
Private land discovery #1	Model: suitable sites near known location	2020	0.21	0.21	1.39
Private land discovery #2	Model: suitable sites near known location	2020	0.36	0.28	2.23
New Smoke Creek Road	Model: suitable sites near known location	2020	0.83	0.15	1.22

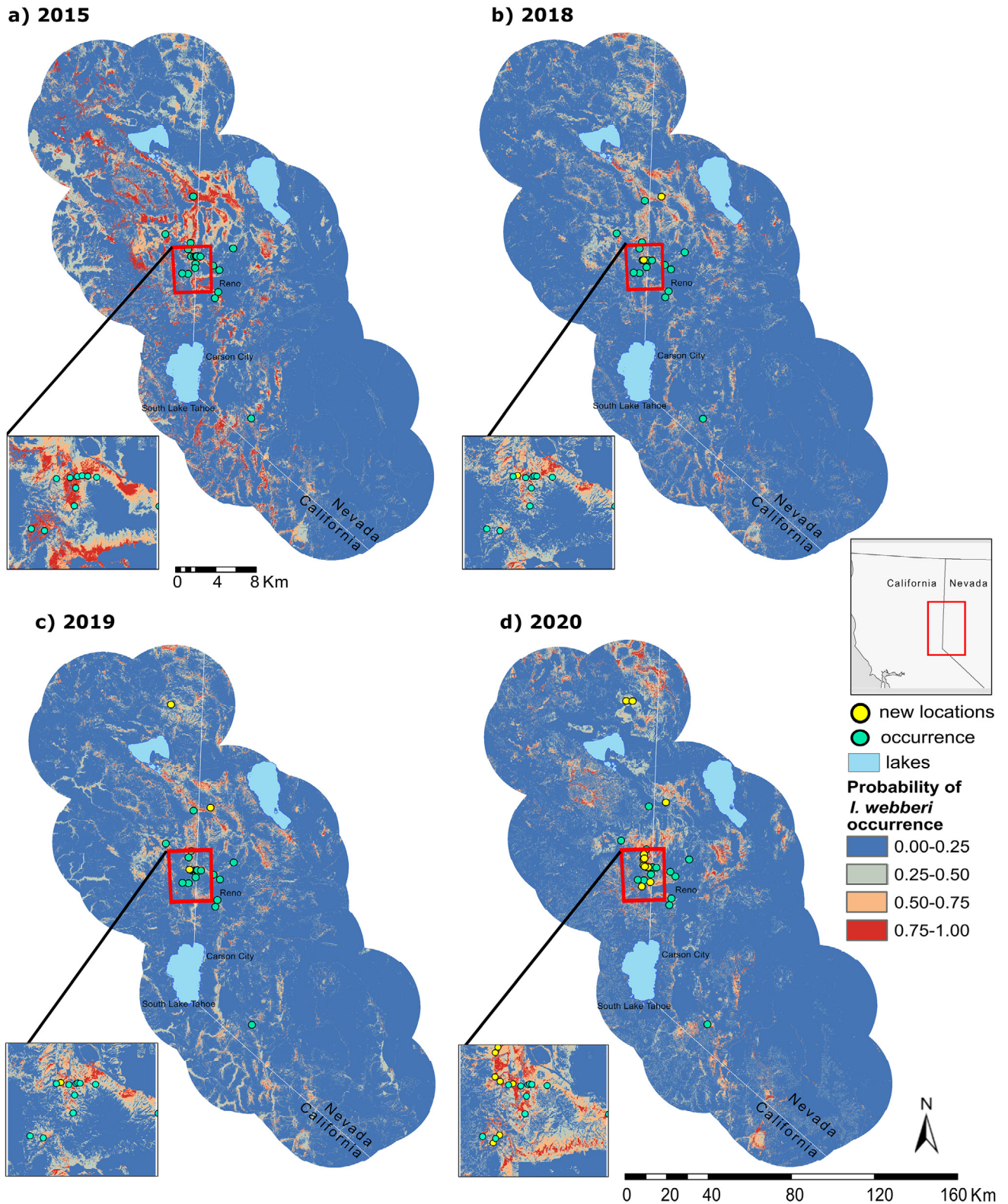


Fig. 3. Predicted geographical distribution of *Ivesia webberi* in the western Great Basin Desert, with both the original and new occurrence points overlay. Red pixels: predicted high probability areas of *I. webberi* occurrence; orange and grey pixels: intermediate to low probability of species occurrence; blue pixels: zero to low probability of species occurrence (non-suitable areas). Green occurrence points: original *I. webberi* occurrence points; yellow occurrence points: novel *I. webberi* populations

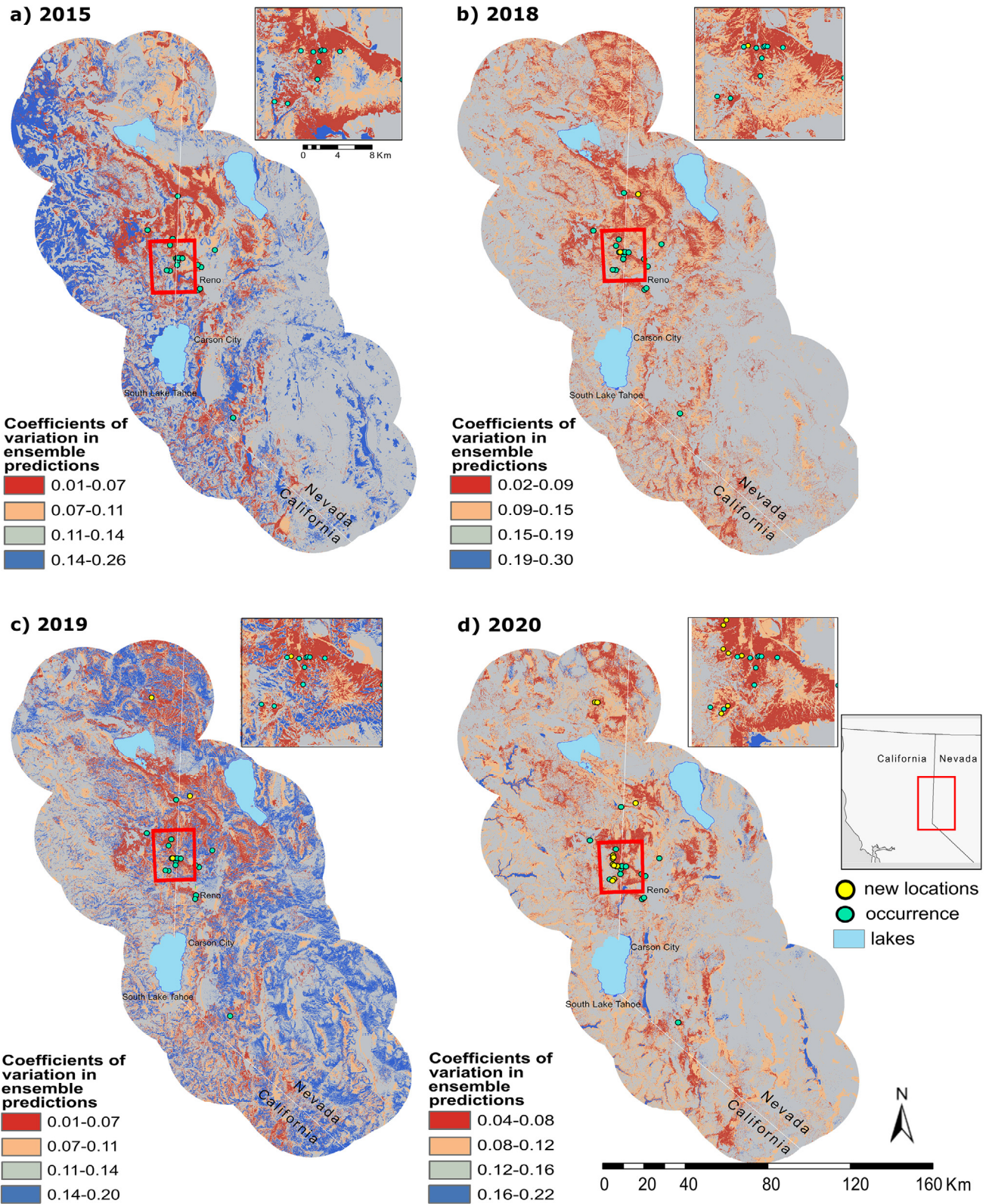


Fig. 4. Coefficients of variation within the ensemble predictions of *Ivesia webberi* in the western Great Basin Desert, with both the original and new occurrence points overlay. Red pixels: low prediction uncertainty areas; orange and grey pixels: intermediate to low prediction uncertainty areas; blue pixels: high prediction uncertainty areas in model predictions

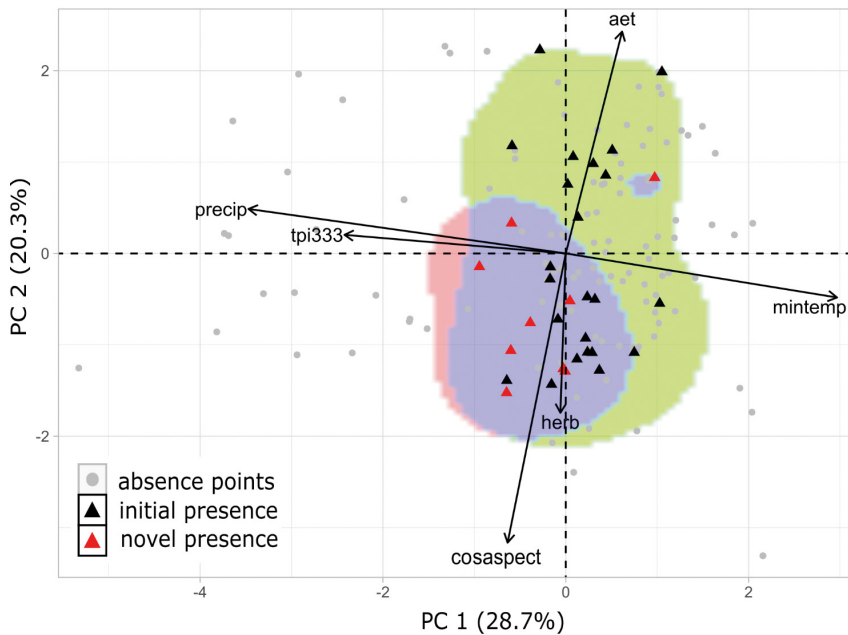


Fig. 5. Principal component analysis biplot of the environmental predictors (see Fig. 1 and Table 2 for definitions) that influence *Ivesia webberi* niche in the western Great Basin Desert. These absence and presence data are combinations of initial (black) and new locations (red). Green area: niche occupied only by the initial occurrences (unfilling); blue area: niche occupied by both initial and novel occurrences (stability); pink area: niche occupied only by the new locations (expansion)

variance (Table S3). The PCA niche similarity test shows that the environmental niche of the new occurrences overlaps that of the initial occurrences with marginal significance (Schoener's $D = 0.47$; $p = 0.05$). This finding was corroborated by the niche stability result, showing that the environmental niche of the new occurrences is similar to the initial occurrences (niche stability = 0.93; Fig. 5), although this high value was marginally significant ($p = 0.09$; Fig. 6). Furthermore, the values of the new locations in the niche density of the initial occurrences ranged from 0.21 to 0.87 (Table 3). This shows that these new points are found within the initial niche, indicating niche stability. However, niche changes between the initial and new occurrences were due to unfilling (estimate = 0.47; $p = 0.11$) rather than expansion (estimate = 0.07; $p = 0.09$; Fig. 6). The majority of the randomized niche overlap and stability estimates were lower than the observed values (Fig. 6a,b), while

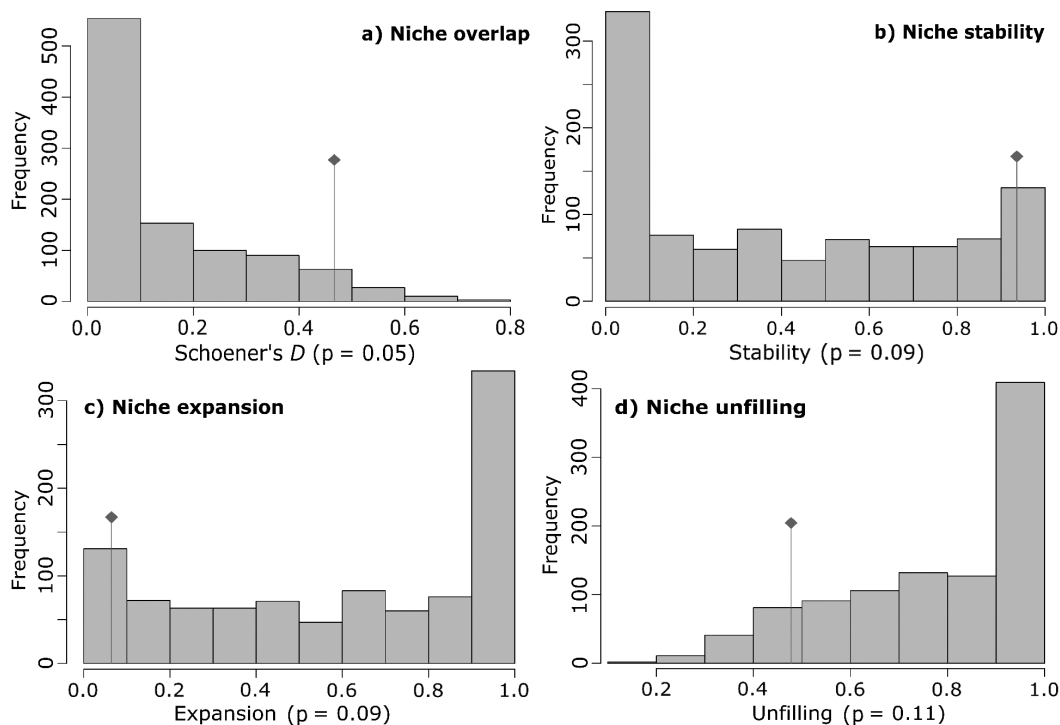


Fig. 6. Randomized values for (a) niche overlap, measured as the Schoener's D , (b) niche stability, (c) niche expansion and (d) niche unfilling between the initial and novel occurrence locations for *Ivesia webberi*. The \blacklozenge on each plot represents the actual niche metric. For each niche estimate, 1000 randomizations were done using a niche similarity test that randomly shifts the centroids of the initial and novel realized niches

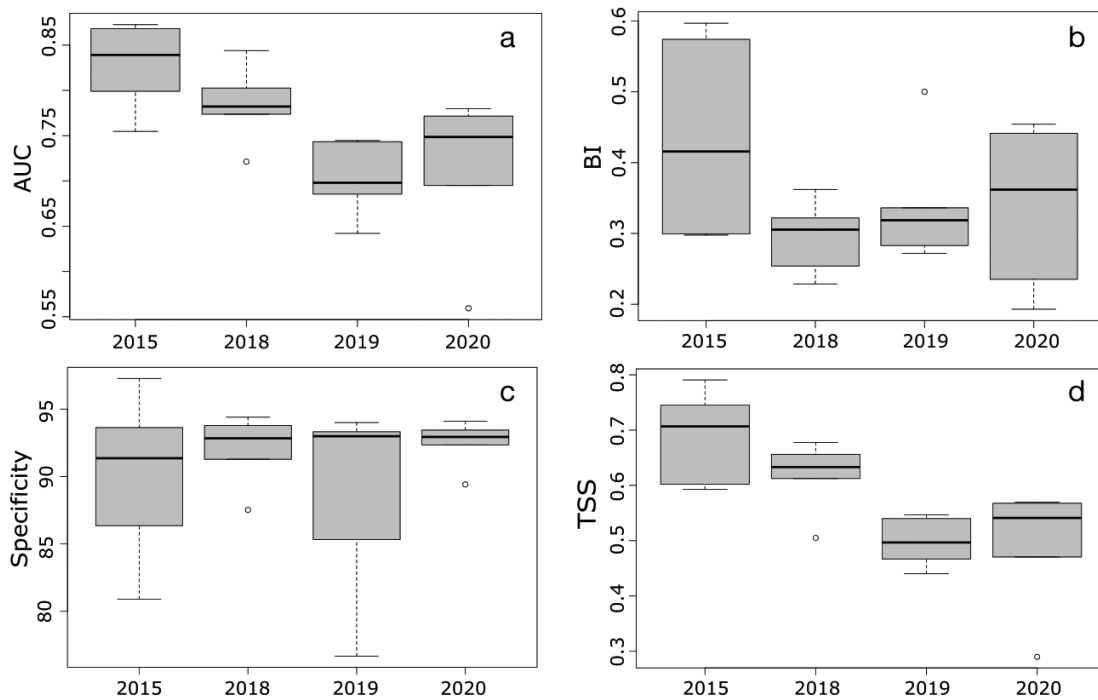


Fig. 7. Boxplots showing the performance of species distribution models ($n = 10$ replicates each for 6 algorithms) in (a) area under curve (AUC) of the receiver operating characteristic plot, (b) Boyce index (BI), (c) specificity, and (d) true skill statistic (TSS) across the years of iterative niche modeling (shown on x-axes). The bold lines represent median values, while the boxes illustrate the lower and upper quartiles. The lower and upper whiskers represent the lowest and highest values, respectively

the majority of the randomized expansion and unfilling estimates were higher than the observed values (Fig. 6c,d).

3.4. Effect of the modeling iterations on the predictive accuracy and reliability of SDMs for *I. webberi*

Fig. 7 shows the mean performance metrics for the iterative ensemble SDMs between 2015 and 2020. The true skill statistics (TSS)-based model performance scores significantly decreased from 0.70 in 2015 to 0.60 in the 2020 model iterations (Tukey post-hoc: $p = 0.01$). Similarly, the area under the curve (AUC) scores significantly decreased from 0.83 to 0.72 between the 2015 and 2020 model iterations (Tukey post-hoc: $p = 0.02$). However, both Boyce Index and specificity showed non significant ($p > 0.05$) changes between the 2015 and 2020 model iterations (0.43 to 0.34 and 90.14 to 92.53, respectively). The predictive performance of the spatial block niche modeling for the 2015 spatial data indicates model overfitting ($AUC_{\text{BLOCK}} = 0.81$, $AUC_{\text{TEST}} = 0.47$), in contrast to the 2020 spatial data which did not exhibit overfitting ($AUC_{\text{BLOCK}} = 0.47$, $AUC_{\text{TEST}} = 0.52$).

4. DISCUSSION

Within a 5-yr period, our iterative modeling approach resulted in the discovery of 9 novel locations (representing 39% of the initial known distribution) and a 63 km expansion of the predicted geographical range of a federally threatened perennial forb. The discovery of new locations from model-guided field surveys is frequently reported for rare species in the literature and highlight the importance of SDMs and model-guided field surveys in conservation. As a result of enlarged occurrence datasets and known ranges, many threatened species have subsequently been delisted from the US Endangered Species Act (Keinath et al. 2014, Sofaer et al. 2019). Additionally, with sufficient spatial data, models can reliably identify biologically relevant ecological factors that support species persistence and predict their potential distributions. In this study, the number of *Ivesia webberi* occurrences increased by 39% (from $n = 23$ to $n = 32$) and *Ivesia webberi* patch occupancy in many of the new locations compares well with those of the original locations. Therefore, findings from this study can guide decisions on future *I. webberi* management. Moreover, previous studies have also reported major revisions to conservation management and

reserve designs due to the additional biodiversity data from model-guided field surveys (Platts et al. 2010), including decisions regarding translocation of species of conservation concern (Draper et al. 2019). Findings of multiple analyses show that the majority of the new locations are found within the environmental niche of the initial occurrences. We observed high niche stability (93%) and low niche expansion (7%) between the environmental conditions in the initial and new occurrences. Moreover, both the initial niche density values of the new locations and the niche dynamics plot (Fig. 5) illustrate the position of the new locations within the realized niche space of the initial occurrences. This is not surprising, considering that the field validation surveys that resulted in the discovery of these novel locations were based on initial models. The observed niche overlap and stability estimates are higher than the majority of the randomly generated niches, whereas niche unfilling and expansion are lower than most of the random niches generated in the similarity test (Fig. 6). In spite of the nonsignificant randomization results, these findings provide partial support for niche similarity between the initial and novel occurrences. The marginally significant randomizations ($0.05 < p < 0.15$) could be attributed to a limited statistical power due to the low number of occurrences and high degree of geographical similarity in both the initial and new datasets (Brown & Carnaval 2019). The unfilled portion of the niche (Fig. 5) suggests that there may be more *I. webberi* locations yet to be discovered or suitable habitat yet to be colonized due to the species limited dispersal capacity.

Additional spatial data can significantly impact the predictive performance of iterative niche models, due to their effect on model parameters (Guisan et al. 2006). In this study, we observed changes in model performance and geographical projections, despite the minimal changes in the 3 predictors used across all model iterations. Specificity is based on omission error rates, which represent the percentage of false negatives in the spatial data. Therefore, slight increases in specificity across the model iterations suggest that the additional spatial data slightly reduced presence–absence ratio in the overall spatial data and also reduced the model omission errors (Lauzeral et al. 2012, Chiffard et al. 2020). However, the reduction of AUC and TSS, and Boyce Index values in all but the final model iteration may be attributed to overfitting due to insufficient occurrences in the dataset. Additional spatial datasets from multi-year sampling may have corrected model overfitting, but they also resulted in reduced SDM performance. This is consistent

with previous studies that also reported reduced niche performance when correcting overfitting in niche models (Guisan et al. 2006, Peterson et al. 2007). Therefore, a fair performance assessment for iterative niche modeling should focus on model generalizability as the primary measure of performance as opposed to model fit for any given year. A rigorous approach to assessing model generalizability (or lack of overfitting) is to use spatially independent data for model validation, as in the spatial block niche modeling approach employed in this study. Secondly, some of the additional absence points were sampled from areas that were predicted to be suitable. This can introduce noise into spatial data used for iterative niche modeling because the absence of *I. webberi* in these predicted suitable sites may be due to dispersal limitation (Lobo et al. 2010, Lauzeral et al. 2012). Field observations support the suitability of some of these surveyed sites because they have similar edaphic and topographic features, and the occurrence of common associates like *Balsamorhiza hookeri*, *Artemisia arbuscula*, *Antennaria dimorpha* and *Phlox longifolia*. McCune (2016) reported similar circumstances where common floristic associates of several studied plants were found in sites predicted to be suitable. Therefore, the inclusion of such absence points in iterative niche models can result in the underprediction of the potential niche and a reduction in model performance (Araújo & Peterson 2012).

The biology of a species may also affect the predictive performance of niche models (Marmion et al. 2009, Regos et al. 2019), particularly the performance of iterative SDMs following the addition of new spatial data (Guisan et al. 2006, Lauzeral et al. 2012). Despite its relatively restricted geographical range, *I. webberi* is locally abundant in occurrence locations and it exhibits mixed mating system (Borokini et al. 2021). These traits suggest high colonization potential and wider niche breadth (Grant & Kalisz 2020), which fits the description of satellite-type species (Hanski 1982, Collins et al. 1993). For satellite-type species, low dispersal capacity limits the full colonization of suitable habitat and may reduce predictive performance of SDMs (Edwards et al. 2005). Araújo & Peterson (2012) cautioned that areas of commission errors should be interpreted carefully for species with fewer occurrences because they may represent suitable habitats that are yet to be colonized (i.e. potential niche). This may be true for the neo-endemic *I. webberi*, which may not yet be in equilibrium with its suitable environment (Araújo & Pearson 2005) because it has not yet fully colonized its range (Kraft et al. 2010, Thornhill et al. 2017). To

reduce spatial bias in iterative SDMs, additional spatial data must be collected using stratified sampling from both sites with predicted higher and low probabilities of species occurrence (Edwards et al. 2005, Guisan et al. 2006). Additionally, absence points too close to presence points in ordination space (thus sharing similar environmental conditions) should be excluded from subsequent modeling.

A combination of biotic and topo-climatic variables contributes to the niche of *I. webberi*. Throughout modeling iterations, perennial herbaceous cover and TPI consistently contributed the most to *I. webberi* distribution, while cumulative AET, cosine aspect and summer seasonal precipitation also contributed to the species niche in model iterations. The perennial herbaceous cover may have constrained *I. webberi* niche to areas of suitable vegetative community, thus representing a biotic component of the species niche. Vegetative land cover is reported in the literature as an important predictor of habitat suitability for rare plants (Gogol-Prokurat 2011, McCune 2016). TPI illustrates topographic heterogeneity, which impacts microclimatic conditions and influences plant distribution and diversity in high-altitude and heterogeneous landscapes (Chardon et al. 2014, Thornhill et al. 2017). The greater probability of *I. webberi* occurrence in areas with higher cosine aspect in the 2018 and 2019 model iterations shows that the species prefers cooler north-facing slopes which receive less sunlight. Though topographic variables are not proximal (Austin 2002), they have been used successfully as spatial delineators and to represent missing climatic variables especially in high-altitude areas, map species habitat suitability, reduce niche model overprediction and increase model performance (Lassueur et al. 2006, Fois et al. 2018).

Summer seasonal precipitation and cumulative AET, the bioclimatic variables, represent the availability of water and energy which governs the timing of spring regeneration and seed germination in *I. webberi*. Summer seasonal precipitation may play an important role in *I. webberi* seed dispersal, as has been observed for spring-germinating plants in other cold deserts of the world (Chen et al. 2019). Field observations show that *I. webberi* seeds are dispersed by gravity-assisted surface run-off due to summer precipitation, resulting in the colonization of interspace microsites and decommissioned roads and trails. This localized seed movement due to summer precipitation was also reported for *I. tweedyi* and *I. lycopodioides* var. *scandularis* (Moseley 1993, Pollak 1997). Taken together, the SDM predictions are congruent with field observations that *I. webberi* suit-

able habitats are found on gentle slopes and ridges dominated by native perennial forbs, herbs, annual grasses and fewer stands of native shrubs, interspersed with bare ground or gravel-covered microsites. Unfortunately, these sites are vulnerable to anthropogenic disturbances and colonization by invasive species which have altered wildfire regimes in the Great Basin Desert (Chambers et al. 2014, Morris & Rowe 2014).

Species with small population size and restricted geographical distributions are more vulnerable to future environmental changes and are frequently targets of conservation priority (Lomba et al. 2010, Sousa-Silva et al. 2014). In this study, we explored the efficacy of 2 complementary approaches for addressing the challenges associated with SDMs for rare species: iterative ensemble modeling and model-guided field sampling. These 2 complementary approaches can reduce spatial bias, allow for model fine tuning that can improve model performance and increase the chances of detecting novel locations that can either fill the realized niche space or expand the species niche breadth, and hence, the known geographical distribution. Improved model performance will enhance reliable assessment of species-environment relationships. Iterative SDMs are particularly important for guiding future efforts to improve species distribution datasets and allow for a tighter integration of models with data, leading ultimately to more accurate and ecologically meaningful SDMs.

Acknowledgements. This research was funded through a grant from the U.S. Fish and Wildlife Service CFDA program (award number F15AC00784), with additional funding support from the Northern California Botanists (NCB) Botany research scholarship program and the Society for Conservation Biology (SCB) graduate student research fellowship. The authors are grateful to the Nevada Natural Heritage Program for supplying the species occurrence data and absence points for modeling. Additional data from Sarah Kulpa of the U.S. Fish and Wildlife Service-Reno Field Office, Valda Lockie of the Bureau of Land Management-Susanville CA Office and the California Native Plant Society were also appreciated. This work also benefited from the expert advice of Dr. Town Peterson, Dr. Olivier Broennimann, August (Tianxiao) Hao and Damien Georges.

LITERATURE CITED

- ✦ Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38:541–545
- ✦ Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43:1223–1232

- Aranda SC, Lobo JM (2011) How well does presence-only-based species distribution modelling predict assemblage diversity? A case study of the Tenerife flora. *Ecography* 34:31–38
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends Ecol Evol* 22: 42–47
- Araújo MB, Pearson RG (2005) Equilibrium of species' distributions with climate. *Ecography* 28: 693–695
- Araújo MB, Peterson AT (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology* 93:1527–1539
- Araújo MB, Rozenfeld A (2014) The geographic scaling of biotic interactions. *Ecography* 37:406–415
- Araújo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species–climate impact models under climate change. *Glob Change Biol* 11:1504–1513
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Model* 157:101–118
- Banerjee AK, Mukherjee A, Guo W, Ng WL, Huang Y (2019) Combining ecological niche modeling with genetic lineage information to predict potential distribution of *Mikania micrantha* Kunth in South and Southeast Asia under predicted climate change. *Glob Ecol Conserv* 20:e00800
- Beygelzimer A, Kakadet S, Langford J, Arya S, Mount D, Li S (2019) FNN: fast nearest neighbor search algorithms and applications. R package version 1.1.3. <https://CRAN.R-project.org/package=FNN>
- Borokini IT, Klingler KB, Peacock MM (2021) Life in the desert: the impact of geographic and environmental gradients on genetic diversity and population structure of *Ivesia webberi*. *Ecol Evol* 11:17537–17556
- Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA (2002) Evaluating resource selection functions. *Ecol Modell* 157:281–300
- Broennimann O, Thuiller W, Hughes G, Midgley GF, Alkemade JMR, Guisan A (2006) Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Glob Change Biol* 12:1079–1093
- Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B and others (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Glob Ecol Biogeogr* 21:481–497
- Brown JL, Carnaval AC (2019) A tale of two niches: methods, concepts, and evolution. *Front Biogeogr* 11:e44158
- Burns PA, Clemann N, White M (2020) Testing the utility of species distribution modelling using Random Forests for a species in decline. *Austral Ecol* 45:706–716
- Chambers JC, Bradley BA, Brown CS, D'Antonio C and others (2014) Resilience to stress and disturbance, and resistance to *Bromus tectorum* L. invasion in cold desert shrublands of western North America. *Ecosystems* 17: 360–375
- Chardon NI, Cornwell WK, Flint LE, Flint AL, Ackerly DD (2014) Topographic, latitudinal and climatic distribution of *Pinus coulteri*: geographic range limits are not at the edge of the climate envelope. *Ecography* 37:1–12
- Chen Y, Shi X, Zhang L, Baskin JM, Baskin CC, Liu H, Zhang D (2019) Effects of increased precipitation on the life history of spring- and autumn-germinated plants of the cold desert annual *Erodium oxyrhynchum* (Geraniaceae). *AoB Plants* 11:plz004
- Chiffard J, Marciau C, Yoccoz NG, Mouillot F and others (2020) Adaptive niche-based sampling to improve ability to find rare and elusive species: simulations and field tests. *Meth Ecol Evol* 11:899–909
- Coddington JA, Agnarsson I, Miller JA, Kuntner M, Hormiga G (2009) Undersampling bias: the null hypothesis for singleton species in tropical arthropod surveys. *J Anim Ecol* 78:573–584
- Collins SL, Glenn SM, Roberts DW (1993) The hierarchical continuum concept. *J Veg Sci* 4:149–156
- Daly C, Halbleib M, Smith JI, Gibson WP and others (2008) Physiographically sensitive mapping of temperature and precipitation across the conterminous United States. *Int J Climatol* 28:2031–2064
- de Siqueira MF, Durigan G, de Marco P Jr, Peterson AT (2009) Something from nothing: using landscape similarity and ecological niche modeling to find rare plant species. *J Nat Conserv* 17:25–32
- Di Cola V, Broennimann O, Petitpierre B, Breiner FT and others (2017) ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40:774–787
- Dilts TE, Weisberg PJ, Dencker CM, Chambers JC (2015) Functionally relevant climate variables for arid lands: a climatic water deficit approach for modelling desert shrub distributions. *J Biogeogr* 42:1986–1997
- Draper D, Marques I, Iriondo JM (2019) Species distribution models with field validation, a key approach for successful selection of receptor sites in conservation translocations. *Glob Ecol Conserv* 19:e00653
- Edwards TC Jr, Cutler DR, Zimmermann NE, Geiser L, Alegria J (2005) Model-based stratifications for enhancing the detection of rare ecological events. *Ecology* 86: 1081–1090
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40:677–697
- Elith J, Ferrier S, Huettmann F, Leathwick JR (2005) The evaluation strip: a new and robust method for plotting predicted responses from species distribution models. *Ecol Modell* 186:280–289
- Federal Register (2014) Endangered and threatened wildlife and plants; threatened species status for *Ivesia webberi*; Final Rule. 50 CFR Part 17. 79 (106): 31878–31883. https://www.fws.gov/sites/default/files/federal_register_document/2014-12627.pdf (first accessed Oct 2014)
- Fois M, Cuenca-Lombrana A, Fenu G, Bacchetta G (2018) Using species distribution models at local scale to guide the search of poorly known species: review, methodological issues and future directions. *Ecol Modell* 385:124–132
- Gogol-Prokurat M (2011) Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. *Ecol Appl* 21:33–47
- Gorban AN, Pokidysheva L, Smirnova E, Tyukina T (2011) Law of the minimum paradoxes. *Bull Math Biol* 73: 2013–2044
- Grant AG, Kalisz S (2020) Do selfing species have greater niche breadth? Support from ecological niche modeling. *Evolution* 74:73–88
- Guisan A, Broennimann O, Engler R, Vust M, Yoccoz NG, Lehmann A, Zimmermann NE (2006) Using niche-based models to improve the sampling of rare species. *Conserv Biol* 20:501–511
- Guisan A, Petitpierre B, Broennimann O, Daehler C, Kueffer C (2014) Unifying niche shift studies: Insights from biological invasions. *Trends Ecol Evol* 29:260–269
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29–36

- Hanski I (1982) Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos* 38:210–221
- Harrell FE, Lee KL, Mark DB (1996) Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 15:361–387
- Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773–785
- Hijmans RJ, Phillips S, Leathwick J, Elith J (2017) Dismo: species distribution modeling. R package version 1.0-12. <http://CRAN.R-project.org/package=dismo>
- Hirzel A, Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecol Modell* 157:331–341
- Hortal J (2008) Uncertainty and the measurement of terrestrial biodiversity gradients. *J Biogeogr* 35:1335–1336
- Hutchinson GE (1957) Concluding remarks. *Cold Spring Harb Symp Quant Biol* 22:415–427
- Ikeda DH, Max TL, Allan GJ, Lau MK, Shuster SM, Whitham TG (2017) Genetically informed ecological niche models improve climate change predictions. *Glob Change Biol* 23:164–176
- Jarnevich CS, Stohlgren TJ, Kumar S, Morisette JT, Holcombe TR (2015) Caveats for correlative species distribution modeling. *Ecol Inform* 29:6–15
- Jiménez-Valverde A, Peterson AT, Soberón J, Overton JM, Aragón P, Lobo JM (2011) Use of niche models in invasive species risk assessments. *Biol Invasions* 13:2785–2797
- Keinath DA, Griscom HR, Andersen MD (2014) Habitat and distribution of the Wyoming pocket gopher (*Thomomys clusius*). *J Mammal* 95:803–813
- Kraft NJ, Baldwin BG, Ackerly DD (2010) Range size, taxon age and hotspots of neoendemism in the California flora. *Divers Distrib* 16:403–413
- Kuhn M (2008) Building predictive models in R using the *caret* package. *J Stat Softw* 28:1–26
- Kursá MB, Rudnicki WR (2010) Feature selection with the *Boruta* package. *J Stat Softw* 36:1–13
- Lassueur T, Joost S, Randin CF (2006) Very high-resolution digital elevation models: Do they improve models of plant species distribution? *Ecol Modell* 198:139–153
- Lauzeral C, Grenouillet G, Brosse S (2012) Dealing with noisy absences to optimize species distribution models: An iterative ensemble modelling approach. *PLOS ONE* 7:e49508
- Li XH, Wang Y (2013) Applying various algorithms for species distribution modelling. *Integr Zool* 8:124–135
- Liu C, Wolter C, Xian W, Jeschke JM (2020) Most invasive species largely conserve their climatic niche. *Proc Natl Acad Sci USA* 117:23643–23651
- Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33:103–114
- Lomba A, Pellissier L, Randin C, Vicente J, Moreira F, Honrado J, Guisan A (2010) Overcoming the rare species modelling paradox: a novel hierarchical framework applied to an Iberian endemic plant. *Biol Conserv* 143:2647–2657
- Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. *Divers Distrib* 15:59–69
- McCune JL (2016) Species distribution models predict rare species occurrences despite significant effects of landscape context. *J Appl Ecol* 53:1871–1879
- Morris LR, Rowe RJ (2014) Historical land use and altered habitats in the Great Basin. *J Mammal* 95:1144–1156
- Moseley RK (1993) Floristic inventory of subalpine parks in the Coeur d'Alene river drainage, northern Idaho. Report to Cooperative Challenge Cost-share Project. Idaho Panhandle National Forests and Idaho Department of Fish and Game. https://idfg.idaho.gov/ifwis/idnhp/cdc_pdf/moser93b.pdf (accessed Feb 2021)
- Peterson AT, Papeş M, Eaton M (2007) Transferability and model evaluation in ecological niche modeling: A comparison of GARP and Maxent. *Ecography* 30:550–560
- Petitpierre B, Kueffer C, Broennimann O, Randin C, Daehler C, Guisan A (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science* 335:1344–1348
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259
- Pili AN, Tingley R, Sy EY, Diesmos MLL, Diesmos AC (2020) Niche shifts and environmental non-equilibrium undermine the usefulness of ecological niche models for invasion risk assessments. *Sci Rep* 10:7972
- Platts PJ, Ahrends A, Gereau RE, McClean CJ and others (2010) Can distribution models help refine inventory-based estimates of conservation priority? A case study in the Eastern Arc forests of Tanzania and Kenya. *Divers Distrib* 16:628–642
- Pollak O (1997) Morphology and dynamics of alpine populations of *Ivesia lycopodioides* ssp. *scandularis* from the White Mountains of California. *Univ Cal White Mount Res Station Symp* 1:97–116
- R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Ramesh V, Gopalakrishna T, Barve S, Melnick DJ (2017) IUCN greatly underestimates threat levels of endemic birds in the Western Ghats. *Biol Conserv* 210:205–221
- Regos A, Gagne L, Alcaraz-Segura D, Honrado JP, Domínguez J (2019) Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Sci Rep* 9:4221
- Rinnhofer LJ, Roura-Pascual N, Arthofer W, Dejaco T and others (2012) Iterative species distribution modelling and ground validation in endemism research: an Alpine jumping bristletail example. *Biodivers Conserv* 21:2845–2863
- Särkinen T, Gonzáles P, Knapp S (2013) Distribution models and species discovery: the story of a new *Solanum* species from the Peruvian Andes. *PhytoKeys* 31:1–20
- Singh NJ, Yoccoz NG, Bhatnagar YV, Fox JL (2009) Using habitat suitability models to sample rare species in high-altitude ecosystems: a case study with Tibetan argali. *Biodivers Conserv* 18:2893–2908
- Sofaer HR, Jarnevich CS, Pearse IS, Smyth RL and others (2019) Development and delivery of species distribution models to inform decision-making. *Bioscience* 69:544–557
- Sousa-Silva R, Alves P, Honrado J, Lomba A (2014) Improving the assessment and reporting on rare and endangered species through species distribution models. *Glob Ecol Conserv* 2:226–237
- Stockwell DRB, Peterson AT (2002) Effects of sample size on accuracy of species distribution models. *Ecol Modell* 148:1–13
- Strubbe D, Broennimann O, Chiron F, Matthysen E (2013) Niche conservatism in non-native birds in Europe: niche

- unfilling rather than niche expansion. *Glob Ecol Biogeogr* 22:962–970
- ✦ Svejcar T, Boyd C, Davies K, Hamerlynck E, Svejcar L (2017) Challenges and limitations to native species restoration in the Great Basin, USA. *Plant Ecol* 218:81–94
- ✦ Thornhill AH, Baldwin BG, Freyman WA, Nosratinia S and others (2017) Spatial phylogenetics of the native California flora. *BMC Biol* 15:96
- ✦ Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* 32:369–373
- ✦ Töpel M, Antonelli A, Yesson C, Eriksen B (2012) Past climate change and plant evolution in Western North America: a case study in Rosaceae. *PLoS One* 7:e50358
- ✦ USGS (United States Geological Survey) (2017) 1 Arc-second Digital Elevation Models (DEMs) – USGS National Map 3DEP Downloadable Data Collection. U.S. Geological Survey. <https://www.usgs.gov/3d-elevation-program> (accessed Mar 2015)
- ✦ Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Aroita G (2019) blockCV: an R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods Ecol Evol* 10: 225–232
- ✦ Vaughan IP, Ormerod SJ (2003) Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conserv Biol* 17:1601–1611
- ✦ Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* 21: 335–342
- ✦ Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62:2868–2883
- ✦ Weiss A (2001) Topographic position and landform analysis. Poster presentation, ESRI User Conference, San Diego, California, USA. http://www.jennessent.com/downloads/TPI-poster-TNC_18x22.pdf
- ✦ Whittaker RJ, Araújo MB, Paul J, Ladle RJ, Watson JEM, Willis KJ (2005) Conservation biogeography: assessment and prospect. *Divers Distrib* 11:3–23
- ✦ Williams JN, Seo C, Thorne J, Nelson JK, Erwin S, O'Brien J, Schwartz MW (2009) Using species distribution models to predict new occurrences for rare plants. *Divers Distrib* 15: 565–576
- ✦ Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. *Divers Distrib* 14:763–773
- Witham CW (2000) Current knowledge and conservation status of *Ivesia webberi* A. Gray (Rosaceae), the Webber ivesia, in Nevada. Status Report prepared for the Nevada Natural Heritage Program. Department of Conservation and Natural Resources, Carson City, NV
- ✦ Xian G, Homer C, Meyer D, Granneman B (2013) An approach for characterizing the distribution of shrubland ecosystem components as continuous fields as part of NLCD. *ISPRS J Photogramm Remote Sens* 86:136–149

*Editorial responsibility: Hans Juergen Boehmer,
Suva, Fiji Islands*
Reviewed by: 2 anonymous referees

Submitted: May 1, 2022
Accepted: November 25, 2022
Proofs received from author(s): January 16, 2023