Ecological Informatics 36 (2016) 126–134

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Field validation of an invasive species Maxent model

Amanda M. West [a,*], Sunil Kumar [a], Cynthia S Brown [b], Thomas J. Stohlgren [a], Jim Bromberg [c]

[a] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523, USA
[b] Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523, USA
[c] Rocky Mountain National Park, U.S. National Park Service, Estes Park, CO 80517, USA

ARTICLE INFO

Article history:
Received 16 April 2016
Received in revised form 27 October 2016
Accepted 4 November 2016
Available online 5 November 2016

Keywords:
Biological invasions
Field validation
GLM
Maxent
Habitat suitability
Regression analysis
Model comparison
Bromus tectorum

ABSTRACT

Accurate and reliable predictions of invasive species distributions are urgently needed by land managers for developing management plans and monitoring new potential areas of establishment. Presence-only species distribution models are commonly used in these evaluations, however they are rarely tested with independent data over time or compared with presence-absence models fit with the same presence data. Using Maxent, we developed a presence-only model of invasive cheatgrass (*Bromus tectorum* L.) distribution in Rocky Mountain National Park, Colorado, USA in 2007 fit with limited data, and then tested the model with independent presence and absence data collected between 2008 and 2013. This model was verified using threshold dependent and threshold independent evaluation metrics. Next, we developed a Maxent model with cheatgrass presence data from 2007 through 2013 (i.e. Maxent 2013), and compared this model to a presence-absence method (i.e., generalized linear model; GLM 2013) using the same data. Threshold dependent and threshold independent evaluation metrics suggested Maxent 2013 outperformed GLM 2013, and a two-tailed Wilcoxon signed rank test indicated relative probability outputs were not significantly different between the models in geographic space. Based on known presences and absences of cheatgrass collected in the field, the Maxent 2013 and GLM 2013 relative probability outputs were highly correlated at absence locations but less correlated at presence locations. A Kappa comparison of Maxent 2007 and Maxent 2013 binary output provides evidence that Maxent is robust when fit with limited data. Our results indicate Maxent is an appropriate model for use when land management objectives are supported by limited resources and thus require a conservative, but highly accurate estimate of habitat suitability for invasive species on the landscape.

1. Introduction

The importance of predicting species distributions is increasing rapidly with global changes and their influences on native ecosystems. Scientists or land managers may need to locate and protect populations of a rare species or identify habitat that may be threatened by an invasive species, to name two of many reasons for the need of accurate predictive tools. Distributions of species vary according to an array of biological and physical conditions underlying the fundamental niche (Hutchinson, 1957), and correlative species distribution models (SDMs) provide a tool that enhances our understanding of this niche in geographic space. The maximum entropy model (Maxent; Phillips et al., 2006) is one of the most widely used presence-only SDMs; as of 04/15/2016, searching for '"maxent" and "species distribution"' in Web of Science yields 1292 results. This approach has demonstrated comparable ability to predict a species' range to models that use both locations where the species is known to occur and known not to occur (i.e., presence-absence models; Elith et al., 2006). Presence-only models use background points rather than true absences, and do not assume that absence precludes the possibility of occurrence (Evangelista et al., 2008; Kumar et al., 2009). Much uncertainty exists with absences, since they may indicate either unsuitable habitat or suitable habitat into which the species has not yet dispersed (Jarnevich et al., 2015).

While many of these models have been determined to effectively predict where species are likely to occur, they may not be rigorously validated. Many species habitat models use a subset of the original data to validate the model (Elith et al., 2006; Fielding and Bell, 1997). In such cases, the data are partitioned into training data to generate model predictions and testing data that are used to assess the accuracy of the model predictions. If the testing data are sufficiently predicted correctly by the model, then the model is considered to accurately predict the species' range. Since the testing data are a random sub-sample of the original dataset, information cannot be obtained on the accuracy of the model when applied to a larger region than that from which the original data came. Improved model evaluation can be obtained by incorporating independent field based presence and absence data, but this method is rarely used, particularly for invasive plant species (see Costa et al., 2010 and Rebelo and Jones, 2010 for examples using reptiles and bats, respectively).

* Corresponding author.
E-mail address: amanda.west@colostate.edu (A.M. West).

http://dx.doi.org/10.1016/j.ecoinf.2016.11.001

Model comparisons can be used to evaluate multiple SDMs using both threshold dependent and threshold independent evaluation metrics. The area under the receiver operating characteristic (ROC) curve (AUC) is a commonly used threshold independent metric for evaluation of SDMs fit to true presence and absence data (Elith et al., 2006; Evangelista et al., 2008; Hosmer and Lemeshow, 2000; Swets, 1988). Test AUC (AUC_{TEST}) measures the ability of model predictions to discriminate between observed presence and absence for a test dataset (e.g., data held aside in a 10-fold cross validation split or independent test data), regardless of the absolute value of the predictions (Fielding and Bell, 1997; Elith and Graham, 2009). However, the use of AUC has its drawbacks. A low AUC value may indicate low discrimination between presences and absences even with a model that fits the data accurately (Lobo et al., 2008). AUC values also provide no information on the spatial distribution of incorrectly predicted presences and absences of a species (Lobo et al., 2008). Thus, AUC is useful in measuring how well presence locations can be discriminated from absences based on predictor variables, while providing little information about how well the model predictions fit the species distribution.

While AUC provides the ability of a model to discriminate between presences and absences, additional metrics can be used to evaluate SDMs developed using threshold selection methods based on study objectives. In the case of invasive plant species, management objectives may be tied to gaining the best possible understanding of where a given species exists on the landscape currently, which would encourage a model threshold based on maximizing sensitivity. Sensitivity measures the percentage of correctly classified presences, while specificity measures the percentage of correctly classified absences. Percent correctly classified (PCC) index considers both sensitivity and specificity. The true skill statistic (TSS = sensitivity + specificity − 1) places equal weight on model sensitivity and specificity, with values ranging between −1 and 1 (Allouche et al., 2006). Values above zero indicate better model performance than chance alone. Often, studies using SDMs rely on these threshold dependent metrics to evaluate and compare model performance and do not consider alternative indicators of model robustness, including comparisons in geographic space. Examples of model comparisons in geographic space include relative probability raster output comparisons using the two-tailed Wilcoxon signed-rank test and regression analysis, and binary raster output comparisons using the Kappa statistic.

The focus of this study was to examine a Maxent model developed with limited presence- only data for an invasive species, and evaluate its usefulness in a management context using threshold independent, threshold dependent, and geographic similarity comparison metrics. We modeled the distribution of cheatgrass because of the concern that land managers have about spread of this non-native species throughout high elevation plant communities (Bromberg et al., 2011; West et al., 2015). While modeling potential ranges of other species may be of interest as well, cheatgrass was of high priority to land managers in our study area, Rocky Mountain National Park. Tied to management objectives, the primary motivation of this study was to determine whether the predicted Maxent relative probabilities were strong indicators of where cheatgrass would be present. We used an independent presence and absence dataset collected during new field campaigns to validate initial Maxent model predictions, highlighting statistical robustness that cannot be obtained from partitioning the original data into training and testing subsets. Finally we combined the newly collected field data with the existing cheatgrass presence (and absence) data and compared Maxent to a commonly used presence-absence model, generalized linear model (GLM).

Our objectives were to: (1) generate an initial potential habitat suitability model for cheatgrass using Maxent fit with presence-only data, and use field sampling to test the predictions; (2) compare Maxent and GLM model predictions fit with the split-sample approach using threshold-dependent and threshold-independent metrics and comparisons in geographic space, and (3) identify the best fit model for management purposes.

2. Materials and methods

2.1. Study area

The study was conducted in Rocky Mountain National Park (referred as the Park hereafter), near the Colorado Front Range in the southern region of the Rocky Mountains. The elevation of the Park ranges from approximately 2300 m (7500 ft) in Estes Park to over 4300 m (14,100 ft) on Longs Peak. The Park is situated at latitudes of approximately 40°10′N to 40°32′N and longitude of 105°31′W to 105°41′ W (Peet, 1981). One main road traverses the Park running generally east to west, while additional roads run along the eastern border of the Park. The backcountry is accessible through 578 km (359 miles) of trails as they meander throughout the Park. Grasslands, shrub lands, and forests as well as rocky, non-vegetated areas were included in the study region. All of the sampling sites occurred within the Park and ranged in elevation from 2490 m to 3540 m.

The Park experiences an arid climate east of the continental divide with average annual precipitation of approximately 400 mm in Estes Park at the east side of the Park (WRCC, 2009). Approximately 480 mm of precipitation fall annually in Grand Lake at the west side of the Park (WRCC, 2009). Most of the total precipitation comes in the form of summer rain although the west side of the Park receives much more winter snowfall (WRCC, 2009). The growing season is short with snow often occurring into early June and returning in September and the potential for snow any month of the year. Average high temperatures in July are 25.7 °C with lows around 7.8 °C (WRCC, 2009). Average temperatures for the month of January range from a high of 3.5 °C to a low around −8.7 °C (WRCC, 2009). Extremely rapid changes in weather are a common occurrence in the Park.

2.2. Field methods

Cheatgrass presence data ($n = 21$) were collected in the Park using a modified Whittaker plot design between 1993 and 2007 (Stohlgren et al., 1995). A presence-only model for cheatgrass was developed in Maxent using these data (see Maxent 2007 in Modeling procedure). Relative probability output from this model was used to stratify field samples taken in 2008 through 2013; these field samples would later be used to validate the model. To stratify the field samples, random site coordinates in Universal Transverse Mercator (UTM) projection were generated in ArcGIS 9.2 (ESRI Inc., Redlands, CA, USA) and stratified among five relative probability classes (>0.1, 0.1–0.3, 0.3–0.5,0.5–0.8, and 0.8–1.0) of cheatgrass habitat suitability from Maxent 2007 (Bromberg et al., 2011). The coordinates were also stratified among vegetation communities and elevation to capture the available environment for cheatgrass; these covariates were two of the most influential environmental predictors from Maxent 2007. Distance to the nearest road or trail was also one of the top three environmental predictors, but was not used for stratifying sample locations. An array of distances from roads and trails would automatically be captured in the randomness of the stratified sampling. Elevation was grouped into six classes (<2500 m, 2500–2700 m, 2700–2900 m, 2900–3100 m, 3100–3300 m, >3300 m) for the purpose of stratifying site locations. Elevation of randomly generated sites ranged from 2396 m to 4023 m. Sites actually visited ranged from 2490 m to 3540 m in elevation. Missing presences of cheatgrass at higher elevations was not a concern, since the highest recorded specimen in Colorado was collected in 2004 at approximately 3050 m (Rocky Mountain Herbarium). That is substantially lower in elevation than many of the highest sites visited in this study. Distance to the nearest road or trail of randomly generated sites ranged from 30 m to 12,046 m with the farthest site visited at 8574 m from a road or trail. The sites were stratified among six vegetation communities, which comprised non-vegetated, shrubland, grassland, deciduous forest, coniferous forest, and tundra.

Sites were visited during the summers of 2008–2013, during early July to early September ($n = 298$). We used a Garmin ETrex Vista GPS unit to navigate to the UTM coordinates using NAD83 datum to match the reference system in which the locations were originally generated. Once at a particular UTM coordinate, we searched for any cheatgrass within a 30 × 30 m area to match the resolution of the environmental variable layers used in Maxent 2007. We spent approximately 10 to 20 min at each plot to thoroughly scour for signs of cheatgrass within the plot. Sites with minimal vegetation required less time to search for the grass than those is dense grasslands and shrublands. Plots significantly infested with cheatgrass also required much less time to determine if the grass was present. Since the purpose of this study was to generate habitat suitability models, we recorded the presence or absence of cheatgrass, but not abundance at each site.

2.2.1. Environmental data

Geospatial raster layers of environmental data were created in ArcGIS (ESRI; Redlands, CA), which resulted in 36 unique covariates to be included in the SDMs (Appendix 1). These covariates included elevation from a digital elevation model (DEM) and five other topographic covariates derived from this DEM (these were selected based on field observations of cheatgrass growth habit in the study area), five spectral indices covariates derived from Landsat 7 ETM + remotely sensed imagery, three covariates derived from MODIS remotely sensed imagery, distance to roads, distance to streams, overland distance to water, solar radiation, and vegetation community types (a categorical variable; see Appendix 1 for further description). The DEM and Landsat layers had a 30 m spatial resolution; the MODIS layers were resampled from 250 m to 30 m to match the other covariates. Prior studies have highlighted the importance of spatial resolution considerations in SDMs (Gillingham et al., 2012; West et al., 2015). A covariate correlation analysis was conducted in the Software for Assisted Habitat Modeling (SAHM; Morisette et al., 2013; West et al., 2016) for assessing multicollinearity among environmental variables. When two variables had a Pearson, Spearman, or Kendall correlation coefficient, $|r| \geq 0.70$, only one of the pair was selected for model development (Dormann et al., 2013), based on percent deviances explained from a univariate generalized additive model (GAM) with the predictor, relative importance of each variable, and expert knowledge of cheatgrass growth habit in the study area.

2.2.2. Modeling procedure

All statistical modeling algorithms were executed in SAHM. We used Maxent (version 3.3.3; Phillips et al., 2006; http://www.cs.princeton.edu/~schapire/maxent/) presence-only model for the 2007 cheatgrass data because of its better performance than other modeling methods; it also performs well even with small sample sizes (Elith et al., 2006; Kumar et al., 2009). Maxent determines patterns in data given constraints placed on the system, and then selects the most likely configuration of the system based on maximizing Shannon's entropy (Merow et al., 2013; Phillips et al., 2006). Maxent automatically includes variable interactions and can consider continuous and categorical predictor variables. To optimize the Maxent model, we used the ENMeval R package (Muscarella et al., 2014) to select a regularization multiplier (i.e. 2.0) and feature types (i.e. hinge, product, linear, quadratic) based on changes in Akaike's Information Criterion (AIC; Anderson and Burnham, 2002) and mean AUC. Additionally, we tested several thresholds to optimize the models (i.e., threshold = 0.5, Sensitivity = Specificity, Maximizes (sensitivity + specificity) / 2, and Minimizes distance between ROC plot and (0, 1); Freeman and Moisen, 2008) before selecting the final threshold (i.e., Minimizes distance between ROC plot and (0, 1)), which falls in line with study objectives (i.e., to maximize sensitivity). Maxent generates a logistic output that can be interpreted as an estimate of relative probability of species distribution in geographic space (Elith et al., 2006), with values that vary from 0 (lowest probability) to 1 (highest probability). In this study, a relative probability output was also produced from GLM 2013, and both will be referred to hereafter as relative habitat suitability.

After we had additional presence cheatgrass data from field campaigns in 2008 to 2013 we used Maxent to generate a final model using the full presence dataset with spatially autocorrelated points tested and removed in ArcGIS using the Global Moran's *I* (Legendre and Legendre, 1998) tool ($n = 157$; Maxent 2013). Additionally, we fit a GLM with these presence data and absence data ($n = 162$) collected from the 2008 to 2013 field campaigns (GLM 2013). GLM is a generalized ordinary linear regression approach that specifies a relationship between the mean of a random variable and a function of the linear combination of predictors (McCullagh and Nelder, 1989). Within SAHM, we fit the GLM model including squared and interaction terms to make it comparable to the Maxent model, and used the stepwise AICc simplification method for optimization. For both the Maxent 2013 and GLM 2013 models, we selected a threshold that minimizes the distance between the ROC plot and (0, 1), which was consistent with Maxent 2007.

2.2.3. Model validation

We tested Maxent 2007 with presence and absence data collected in the 2008 to 2013 field campaigns. For Maxent 2013 and GLM 2013 we used 10-fold cross-validation procedure to test the models. We also compared the relative habitat suitability generated by Maxent 2007 to the actual presences found in the field within each relative habitat suitability class in 2008 to 2013. This allowed us to examine the numbers and percentages of cheatgrass presence in 2008 to 2013 that fell within each relative habitat suitability class from the 2007 model.

2.2.4. Evaluation of model performance

We used threshold-independent (i.e., AUC) and threshold-dependent (i.e., sensitivity, specificity, percent correctly classified, and TSS) measures of model accuracy to evaluate model performance (Franklin, 2009). An AUC value of 0.5 shows that model predictions are not better than random; <0.5 are worse than random; 0.5–0.7 indicates poor performance; 0.7–0.9 reasonable/moderate performance; and >0.9, high performance (Peterson et al., 2011). For testing the differences in predictions of habitat suitability between Maxent 2013 and GLM 2013, we used a two-tailed Wilcoxon signed-rank test in R statistical package (Randin et al., 2006; R Core Team 2012). We generated 1000 random points throughout the Park and extracted the relative habitat suitability values from each of the three models to run the test (H_a = true location shift is not equal to 0). To further compare the relative habitat suitability outputs from the Maxent 2013 and GLM 2013 models, we extracted the predicted value for each respective model at every cheatgrass presence or absence point sampled, and compared these values in a regression analysis. Finally, to evaluate the similarity of quantity and similarity of location between first the Maxent 2013 and GLM 2013 binary raster outputs (i.e. value of 0 or 1 at each raster cell based on threshold), and then the Maxent 2007 and Maxent 2013 binary outputs, we used the Kappa statistic tool in the Model Comparison Kit (Visser and de Nijs, 2006).

3. Results

When tested with an independent presence and absence dataset, Maxent 2007 had robust evaluation metrics: AUC_{TEST} 0.80, PCC 0.74, sensitivity 0.81, specificity 0.68, and TSS 0.49 (Table 1). The likelihood of detection of cheatgrass in 2008 to 2013 field campaigns increased with the higher values of predicted relative habitat suitability in Maxent 2007 (Fig. 1). Even though the lowest relative habitat suitability class (<0.1) was more thoroughly sampled than others, cheatgrass was only found within 5% of the random stratified sampling points for this relative habitat suitability class. In the highest relative habitat suitability class (0.7 to 1.0), cheatgrass was detected in 81% of the locations sampled, the greatest proportion of any of the relative habitat suitability

Table 1
Comparison of cheatgrass models with field collected presence/absence test data and data partitioning.[a]

	Threshold independent (±SD)	Threshold dependent (±SD)			
Model	AUC	PCC	Sensitivity	Specificity	TSS
Maxent 2007 (2013 test data)	0.80	74.00	0.81	0.68	0.49
Maxent 2013	0.96	89.46	0.92	0.89	0.81
(10-fold cross-validation)	(±0.008)	(±0.94)	(±0.05)	(±0.009)	(±0.05)
GLM 2013	0.83	76.19	0.80	0.73	0.53
(10-fold cross-validation)	(±0.09)	(±8.21)	(±0.09)	(±0.12)	(±0.16)

[a] Maxent 2007 is the Maxent model trained using cheatgrass presence data up to 2007 ($n = 21$); test results using field data from 2008–2013 (presence = 180, absence = 232) are presented. Maxent 2013 and GLM 2013 are Maxent and generalized linear models, respectively, trained using 67% of the data and tested with remaining 33%; results presented are averages of 10 replicate runs (±SD). AUC is Area Under the ROC (receiver operating characteristic) Curve; PCC is percent correctly classified; and TSS is True Skill Statistic.

classes. As relative habitat suitability increased by class, the proportion of plots containing cheatgrass also increased (Fig. 1).

Maxent 2013 outperformed GLM 2013 based on 10-fold cross validation evaluation metrics (Table 1) however both models were robust with deviance explained values of 0.44 and 0.69, respectively. The two-tailed Wilcoxon signed rank test for comparison of the Maxent 2013 and GLM 2013 models had a $p < 0.001$, a pseudo-median < 0.001 (i.e., the median of the difference between a sample from the first model and a sample from the second model), and a 95% confidence interval (C.I.) of 0.0000021, 0.000057.

The Maxent 2013 and GLM 2013 relative habitat suitability outputs had higher agreement in predicting absences of cheatgrass than presences; Pearson's correlation between model predicted relative habitat suitability, and cheatgrass presences and absences were 0.20, and 0.84, with C.I. [0.07, 0.32] and C.I. [0.80, 0.87], respectively. Maxent 2013 predicted a higher number of cells in the <0.10 relative habitat suitability class compared to GLM 2013; conversely, GLM 2013 predicted a higher number of cells in the 0.70 to 1.0 relative habitat suitability class compared to Maxent 2013 (Fig. 2).

Predicted relative habitat suitability outputs from Maxent 2007, Maxent 2013, and GLM 2013 had similar spatial patterns (Fig. 3), however the area encompassed by each relative habitat suitability class was highly dissimilar, which is in agreement with Fig. 2 results. When the binary output rasters were compared, the overall Kappa statistic was 0.77 between the Maxent 2013 and GLM 2013 models (Appendix 2a), and 0.73 between Maxent 2013 and Maxent 2007 (Appendix 2b), respectively.

Elevation was the covariate with the highest relative contribution in Maxent 2013 and GLM 2013, however distance to roads and trails was the most important covariate in Maxent 2007. The latter variable was not retained by the GLM 2013 model. The formula for the GLM model was as follows:

Response ~ elevation^2 + flow direction + mean evi^2 + vegetation type 12 + elevation + slope + slope^2 + flow accumulation^2 +-elevation: slope + flow direction: elevation.where: represents interaction between two covariates.

Vegetation type 12 (i.e., subalpine) was the most important covariate in the GLM 2013 model and was also important in Maxent 2007 and Maxent 2013; as cheatgrass habitat suitability decreased, subalpine vegetation cover increased (Appendices 3 and 4). Slope and flow direction were important in the GLM 2013 model but not in Maxent 2007 or Maxent 2013 (Table 2).

4. Discussion and conclusions

Using a suite of model validation tests, this study suggests that a Maxent model fit with limited presence-only data can provide robust estimates of habitat suitability for invasive species on the landscape. Maxent 2007 fit with a small number of presence points resulted in predictions that were robust to an independent test dataset collected in 2008 to 2013. Maxent 2013 relative probability output indicated the habitat suitability of cheatgrass is more restricted in geographic space across the Park when compared to the same output from GLM 2013; this result was expected because Maxent 2013 placed more constraints (i.e. covariates) on the data than GLM 2013 (see Royle et al., 2012). However, the two-tailed Wilcoxon signed rank test based on 1000 randomly selected points from Maxent 2013 and GLM 2013 relative habitat suitability outputs provides strong evidence that there was minimal difference in overall spatial predictions between the two models. Given this result, we were surprised that Maxent 2013 and GLM 2013 relative habitat suitability outputs were highly correlated at known absence locations but not as strongly correlated at known presence locations. This is likely due to the tendency for Maxent to assign higher habitat probability values at presence locations than GLM; however this result

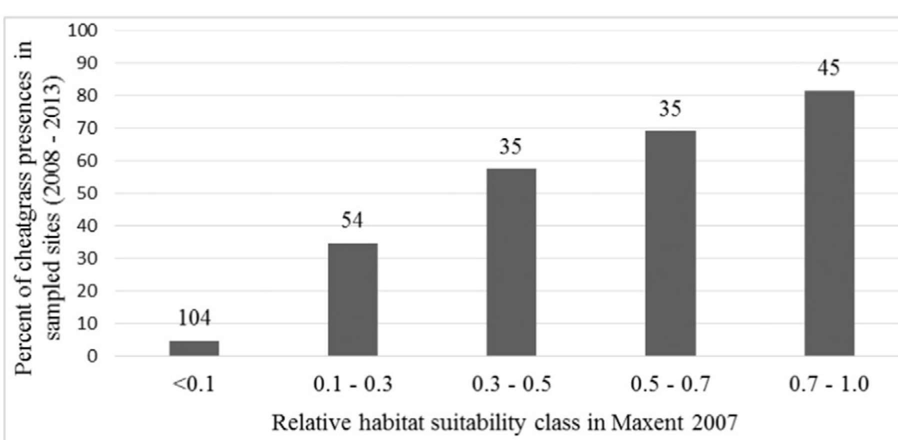


Fig. 1. Bars represent the percent of cheatgrass occurrences in sampled sites for each relative habitat suitability class; classes were generated from the logistic output of Maxent 2007. The numbers above each bar represent the total number of sites visited within each relative habitat suitability class.

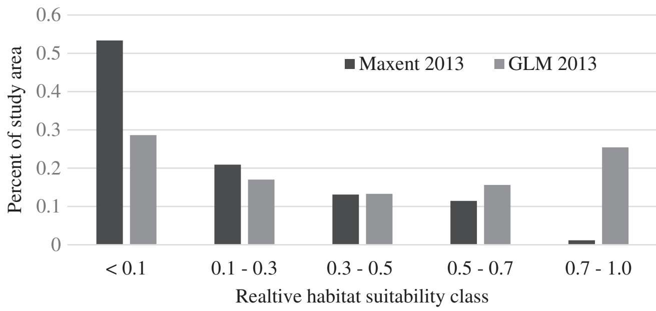


Fig. 2. Percentage of study area in different relative habitat suitability classes from Maxent 2013 and GLM 2013 output.

warrants further investigation. In agreeance with prior studies, topography, distance to roads and trails, and vegetation community type influenced where cheatgrass is found in the Park (Bromberg et al., 2011).

Rebelo and Jones (2010) supported the use of presence-only modeling for a rare bat species with limited data; their test also used an independent test to validate the Maxent model. Other SDM comparisons have highlighted the importance of independent model validation (Gastón and García-Viñas, 2011; Gies et al., 2015). Our results suggest that a Maxent presence-only model can accurately forecast the habitat suitability of the generalist, invasive species cheatgrass and builds on previous model comparison studies (Long et al., 2009). Nonetheless, model comparisons are important (see Brotons et al., 2004 where presence-absence models are more accurate for a generalist bird species, and Gastón and García-Viñas (2011) where penalized logistic regression model results did not differ significantly from Maxent results when tested

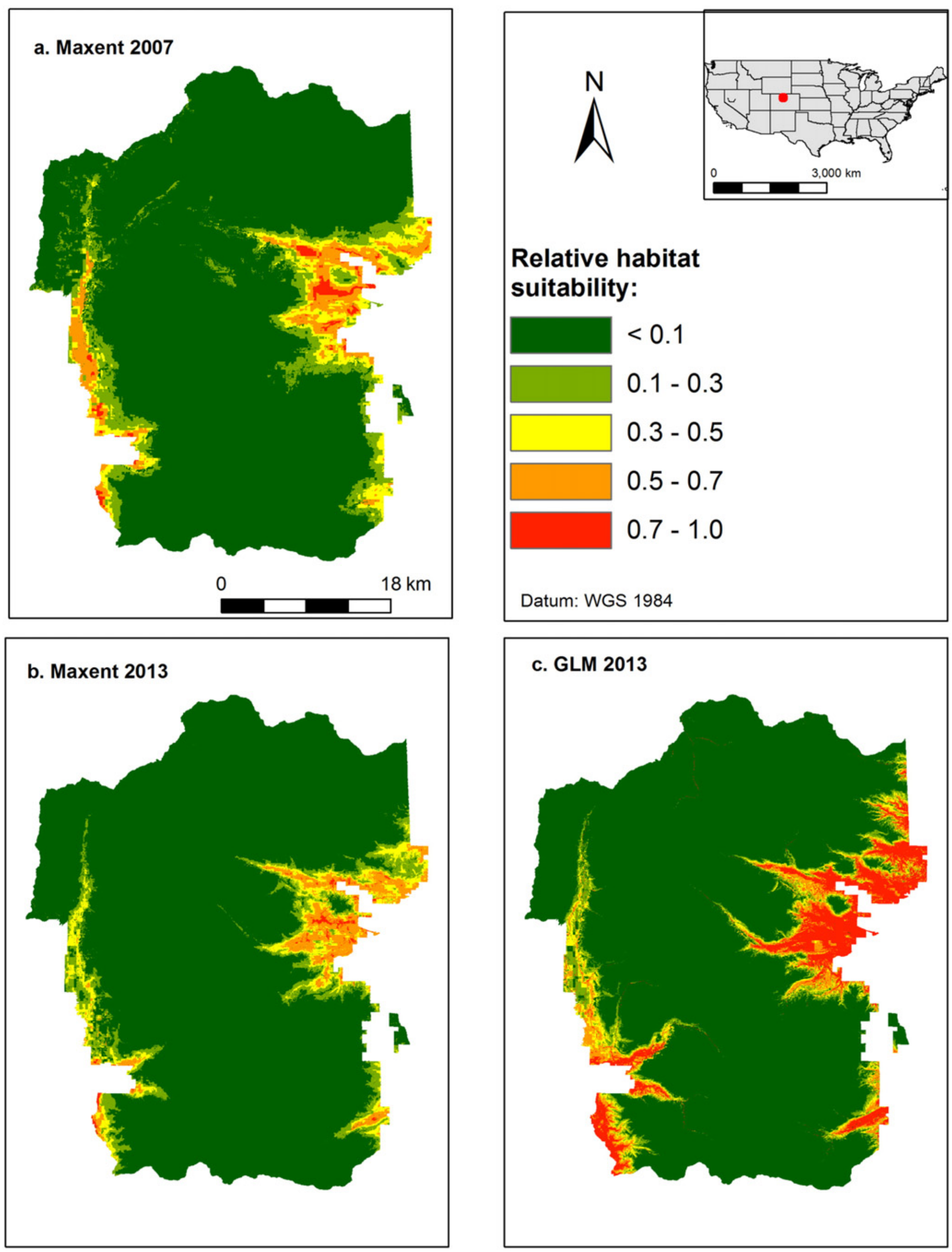


Fig. 3. a–c: Relative habitat suitability for cheatgrass in Rocky Mountain National Park; (a) Maxent model including cheatgrass presence data through the year 2007 (Maxent 2007), (b) Maxent model including cheatgrass presence data from 2007 to 2013 (Maxent 2013), and (c) GLM model including cheatgrass presence and absence data from 2007 to 2013 (GLM 2013).

Table 2
Relative importance (percent contribution) of different environmental predictors in Maxent 2007, Maxent 2013, and GLM 2013. EVI is enhanced vegetation index.

Variable	Maxent 2007	Maxent 2013	GLM 2013
Distance to roads/trails	65.87	13.27	–
Subalpine vegetation type	15.15	9.23	11.38
Elevation	14.05	67.86	73.81
Range EVI	4.92	1.13	–
Overland distance to water	–	2.77	–
Mean EVI	–	1.93	1.69
Flow accumulation	–	1.26	1.21
Slope (degrees)	–	1.04	5.95
Peak EVI	–	1.04	–
Flow direction	–	0.47	5.96

with presence-absence data). Conservative estimates like those provided by the Maxent models can be more useful to land management agencies who seek accurate and reliable predictions that can be used to prioritize areas for invasive species management (Underwood et al., 2004).

More presences of cheatgrass were found at higher probabilities using the random stratified design, which was an initial indication that higher predicted relative habitat suitability does in fact correlate with greater likelihood of presence. A higher proportion of sites visited with cheatgrass present fell into the higher relative habitat suitability classes, also providing evidence that increased predicted relative probabilities indicate a realized increased chance of species presence on the ground. The calibration of the data demonstrated that the proportions of sites visited containing cheatgrass in fact matched the expected ranges within each relative habitat suitability class. It is likely that cheatgrass has not yet fully expanded into its suitable habitat range in the Park, which would result in lower occurrences than expected. Additionally, the Fern Lake wildfire of 2012 encompassed an area of 14 km^2 on the east side of the Park, which may have had positive (e.g., increased N availability and decreased competition) or negative (e.g., depleted seed bank) effects on established cheatgrass populations in that area.

Stratified random sampling was useful for the field validation because it allowed for sampling a much larger area than from where the original data were collected, capturing a wider range of environmental variability in the Park. In Maxent 2007, new regions of the Park where data had not previously been collected were predicted to have high probability of cheatgrass habitat suitability. Cheatgrass was in fact found in many of these high probability regions sampled in 2008 to 2013. Two new regions predicted to have high probabilities of habitat suitability on the east side of the park were validated by our field sampling. A third area that was predicted to have a high probability of habitat suitability on the west side of the Park did not have cheatgrass present at any of the random stratified points from the surveys in 2008 to 2013. However, Park staff found cheatgrass nearby (pers. comm., Dyan Hardin, Rocky Mountain National Park, 13T 0428662 4457565 NAD83). Stratified sampling did not detect cheatgrass in some of the areas with a high predicted cheatgrass habitat suitability. This may have either been due to scaling issues with the model not being able to predict areas far from the initial survey area, or simply lower propagule pressure and dispersal in these areas. The western road corridor in the Park was predicted to have high habitat suitability for cheatgrass. Since most of the cheatgrass in the Colorado Rockies is creeping from the foothills and plains on the east side of the Park, it may not have reached the western side of the Park yet, where only one cheatgrass presence point was found during field sampling.

While Maxent has been used by researchers to make predictions about species distributions, it can be a valuable tool for land managers as well. Maxent predicted the likelihood of cheatgrass presence based on a small initial set of data points. The field validation of the model demonstrated that the predictions were quite good from this initial small dataset. With the limited time and resources that land managers often have for data collection, Maxent can help them determine potential species ranges based on a quick initial assessment of a species. For cheatgrass, land managers can make inferences about potential presence based on model relative probabilities and environmental factors such as elevation. Such information would be useful to managers in helping prioritize the allocation of time and resources. There are always uncertainties in any model predictions (Jarnevich et al., 2015), which is evident from the probability classes and proportions of sites found with cheatgrass in those classes. With such uncertainties, land managers should not solely base their decisions on models, but rather use them to help guide their management efforts.

Maxent predictions have been made for other species, but similar field based validations have rarely been performed (e.g., Costa et al., 2010; Rebelo and Jones, 2010). Examples of the field validation of other SDMs do exist (Dennis and Eales, 1999; Fielding and Haworth, 1995; Randin et al., 2006) but are not common. It is possible that other species with widespread distributions but apparent environmental constraints may also be predicted well by the model, but this information is not known. Prior studies have compared species with limited distributions to those that can thrive in a greater range of environmental conditions (Evangelista et al., 2008; Hernandez et al., 2008). Even a widespread species such as cheatgrass will have constrained distributions in less desirable environments. Cheatgrass is widespread throughout the Great Basin (Knapp, 1996; Mack, 1981), but appears to be more constrained in a high elevation range such as Rocky Mountain National Park. The model should be validated in various physical and climatic conditions to see if it can consistently make correct predictions in numerous types of environments. In addition to testing the model in different environments, other similar generalist as well as specialist species should be included in model validations to determine what types of species best fit the model predictions.

Appendix 1
Environmental variable GIS layers included in analysis.

Environmental variable	Spatial resolution	Data source
Elevation (DEM)[a]	30 m	NED seamless data
Slope	30 m	Derived from the DEM
Eastness	30 m	Derived from the DEM
Northness	30 m	Derived from the DEM
Flow accumulation	30 m	Derived from the DEM
Flow direction	30 m	Derived from the DEM
NDVI (2001)[b]	30 m	Landsat 7 ETM+
Brightness index (2001)[b]	30 m	Landsat 7 ETM+
Greenness index (2001)[b]	30 m	Landsat 7 ETM+
Moistness index (2001)[b]	30 m	Landsat 7 ETM+
Wetness index (2001)[b]	30 m	Landsat 7 ETM+
Mean EVI[c]	250 m	MODIS (resampled to 30 m)
Peak EVI[c]	250 m	MODIS (resampled to 30 m)
Range in EVI[c]	250 m	MODIS (resampled to 30 m)
Distance from roads and trails[d]	30 m	Created in ArcGIS
Distance from streams[d]	30 m	Created in ArcGIS
Overland distance to water[d]	30 m	Flows tools (Theobald et al., 2006)
Solar radiation[d]	30 m	Created in ArcGIS
Vegetation community type[e]	30 m	Landfire data

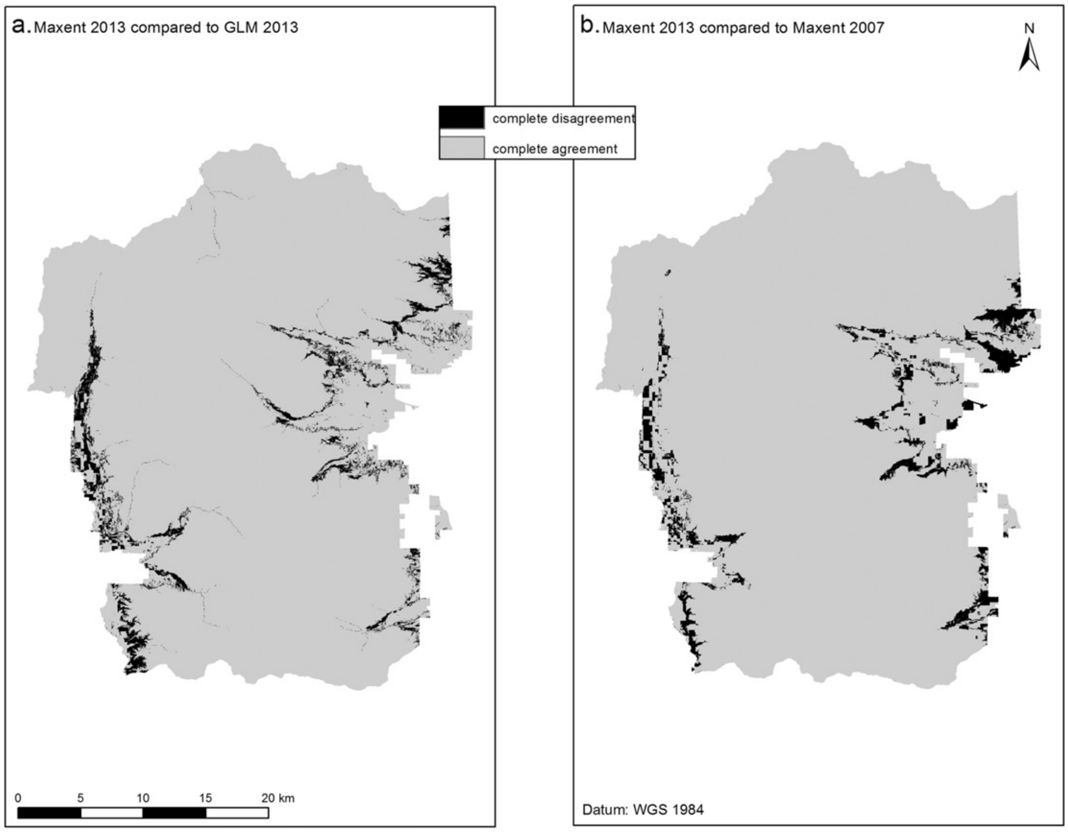
[a] Digital Elevation Model (NED or National Elevation Dataset is the primary elevation dataset used by the USGS, http://ned.usgs.gov/).

[b] Spectral indices derived from Landsat 7 ETM+ satellite imagery; NDVI is Normalized DifferenceVegetation Index, http://earthexplorer.usgs.gov/.

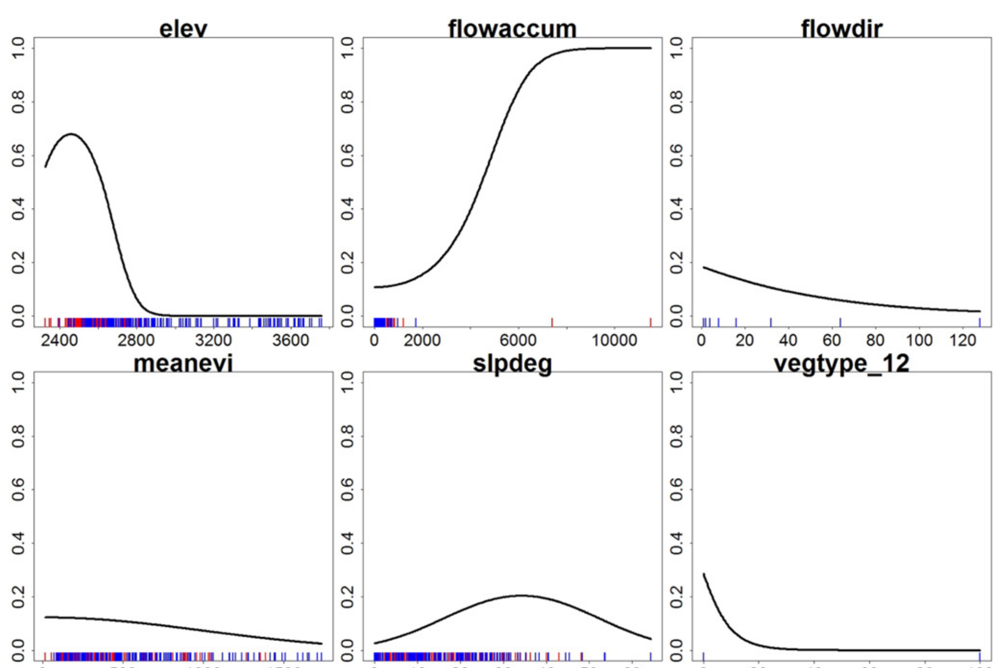
[c] Moderate Resolution Imaging Spectroradiometer, http://modis.gsfc.nasa.gov/.

[d] Variables created in ArcGIS v.9.3 based on data provided by the National Park Service, Rocky Mountain National Park. Theobald, D.M., Norman, J.B., Peterson, E., Ferraz, S., Wade, A. & Sherburne, M.R. (2006) Functional linkage of water basins and streams (FLoWS) v1 user's guide: ArcGIS tools for network-based analysis of freshwater ecosystems. pp. 43. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado.

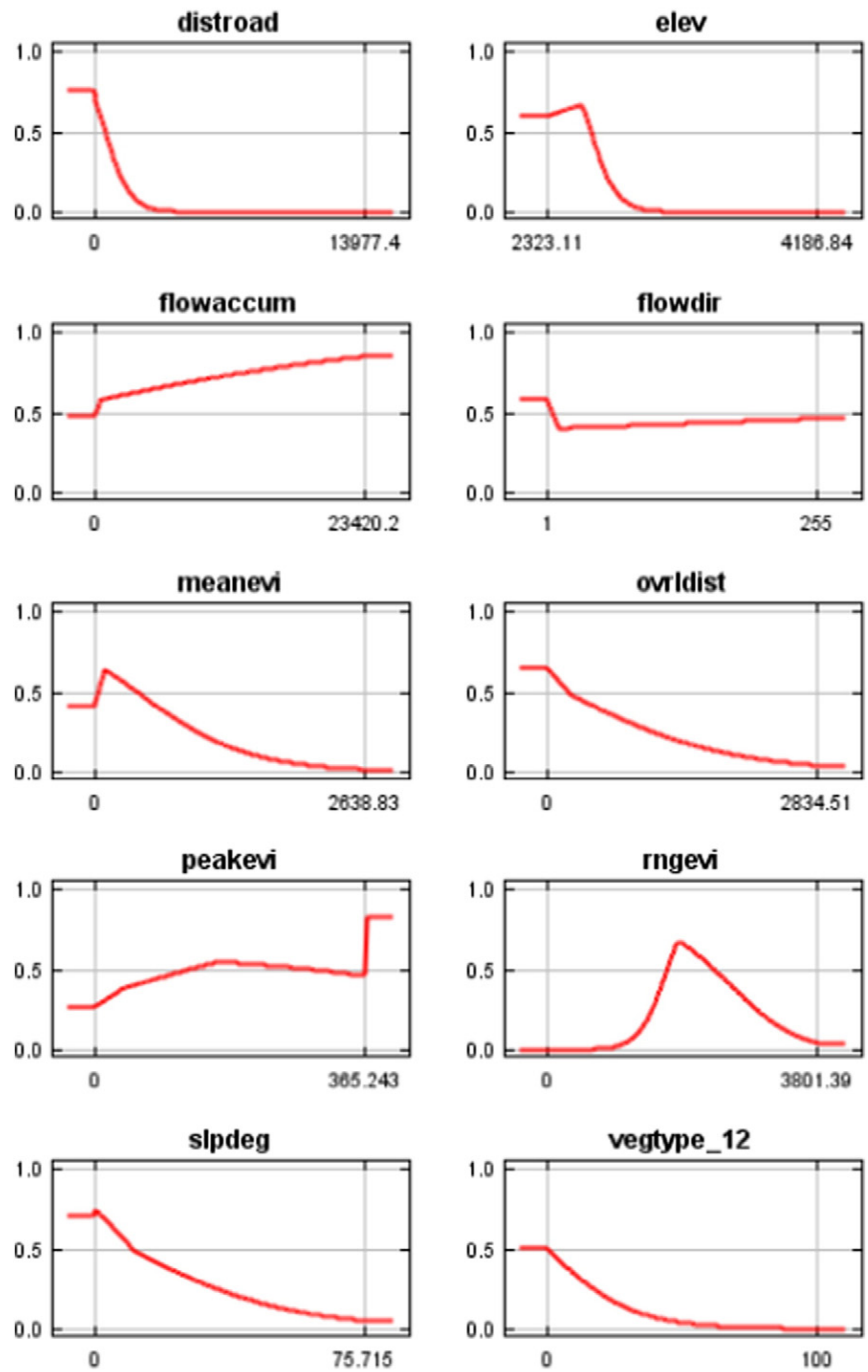
[e] Vegetation community type was a categorical variable derived from Landfire http://www.landfire.gov/NationalProductDescriptions21.php. Using ArcGIS, we classified Landfire categories and then created a continuous raster surface for each of the new 18 vegetation community type categories: (1) water, (2) snow/ice, (3) developed, (4) barren, (5) agriculture, (6) alpine/montane sparsely vegetated, (7) aspen forest, (8) pine/juniper forest, (9) lodgepole pine forest, (10) montane mixed conifer, (11) ponderosa pine forest, (12) subalpine, (13) lower montane/foothill shrubland, (14) alpine dwarf shrubland/alpine rangeland, (15) pine/juniper savannah, (16) sagebrush-steppe, (17) perennial graminoid/grassland, and (18) riparian.



Appendix 2. a-b. Kappa comparison of binary outputs for (a) Maxent 2013 and GLM 2013, and (b) Maxent 2013 and Maxent 2007.



Appendix 3. Response curves for covariates used to fit GLM 2013. Predicted value of habitat suitability is on the Y axis; value of given covariate is on the X axis.



Appendix 4. Response curves for covariates used to fit the Maxent 2013 model. Predicted value of habitat suitability is on the Y axis; value of given covariate is on the X axis.

Acknowledgements

Data for past sampling were provided by Jim Bromberg, Holley Zadeh Gardel, Amanda West, and the United States Geological Survey. Some of the cheatgrass presence data were published in Bromberg et al. (2011) and West et al. (2015). Assistance with database management was provided by Rick Shory, Natural Resource Ecology Laboratory, Colorado State University. Greg Newman, Natural Resource Ecology Laboratory, helped with the modified Whittaker plot design and sampling techniques. AW was supported by the US Geological Survey (Award No. G14AC00357 RMCESU G14AC00138) and the National Institute of Food and Agriculture, U.S. Department of Agriculture (Award No.- 2010-03280), SK was partially supported by funding from U. S. Geological Survey (Award No. G14AC00357 RMCESU G14AC00138) and NASA (Award No. NNH11AS09I). TS was partially supported by USDA CSREES/NRI 2008-35615-04666.

References

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.

Anderson, D., Burnham, K., 2002. Avoiding pitfalls when using information-theoretic methods. J. Wildl. Manag. 66, 912–918.

Bromberg, J.E., Kumar, S., Brown, C.S., Stohlgren, T.J., 2011. Distributional changes and range predictions of cheatgrass in Rocky Mountain National Park. Invasive Plant Sci. Manag. 4 (2), 173–182.

Brotons, L., Thuiller, W., Araujo, M.B., Hirzel, A.H., 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27, 437–448.

Costa, G.C., Nogueira, C., Machado, R.B., Colli, G.R., 2010. Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. Biodivers. Conserv. 19, 883–899.
Dennis, R.L.H., Eales, H.T., 1999. Probability of site occupancy in the large heath butterfly *Coenonympha tullia* determined from geographical and ecological data. Biol. Conserv. 87, 295–301.
Dormann, C.F., Elith, J., Bacher, S., et al., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46.
Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.
Elith, J., Graham, C.H., 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32 (1), 66–77.
Evangelista, P.H., Kumar, S., Stohlgren, T.J., Jarnevich, C.S., Crall, A.W., Norman III, J.B., Barnett, D.T., 2008. Modelling invasion for a habitat generalist and a specialist plant species. Divers. Distrib. 14, 808–817.
Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38–49.
Fielding, A.H., Haworth, P.F., 1995. Testing the generality of bird-habitat models. Conserv. Biol. 9 (6), 1466–1481.
Franklin, J., 2009. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, Cambridge, UK .
Freeman, E.A., Moisen, G., 2008. PresenceAbsence: an R package for presence absence analysis. J. Stat. Softw. 23:1–31. http://dx.doi.org/10.1038/nmat2803.
Gastón, A., García-Viñas, J.I., 2011. Modelling species distributions with penalised logistic regressions: a comparison with maximum entropy models. Ecol. Model. 222, 2037–2041.
Gies, M., Sondermann, M., Hering, D., Feld, C.K., 2015. A comparison of modelled and actual distributions of eleven benthic macroinvertebrate species in a Central European mountain catchment. Hydrobiologia 758, 123–140.
Gillingham, P.K., Huntley, B., Kunin, W.E., Thomas, C.D., 2012. The effect of spatial resolution on projected responses to climate warming. Divers. Distrib. 18, 990–1000.
Hernandez, P.A., Franke, I., Herzog, S.K., Pacheco, V., Paniagua, L., Quintana, H.L., Soto, A., Swenson, J.J., Tovar, C., Valqui, T.H., Vargas, J., Young, B.E., 2008. Predicting species distributions in poorly-studied landscapes. Biol. Conserv. 17, 1353–1366.
Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression. John Wiley and Sons, New York, New York, USA .
Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harb. Symp. Quant. Biol. 415–427 Pages 22.
Jarnevich, C.S., Stohlgren, T.J., Kumar, S., Morisette, J.T., Holcombe, T.R., 2015. Caveats for correlative species distribution modeling. Eco. Inform. 29, 6–15.
Knapp, P.A., 1996. Cheatgrass (*Bromus tectorum* L) dominance in the Great Basin desert: history, persistence, and influences to human activities. Glob. Environ. Chang. 6, 37–52.
Kumar, S., Spaulding, S.A., Stohlgren, T.J., Hermann, K.A., Schmidt, T.S., Bahls, L.L., 2009. Predicting habitat distribution for freshwater diatom *Didymosphenia geminata* in the continental United States. Front. Ecol. Environ. 7, 415–420.
Legendre, P., Legendre, L., 1998. Numerical Ecology. Second English ed. Elsevier Science, Amsterdam, The Netherlands .
Lobo, J.M., Jimenez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Glob. Ecol. Biogeogr. 17, 145–151.
Long, R.a., Muir, J.D., Rachlow, J.L., Kie, J.G., 2009. A Comparison of Two Modeling Approaches for Evaluating Wildlife–Habitat Relationships. J. Wildl. Manag. 73 (2), 294–302.
Mack, R.N., 1981. Invasion of *Bromus tectorum* L. Into Western North America: an ecological chronicle. Agro-Ecosystems 7, 145–165.
McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman and Hall .
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, 1058–1069.
Morisette, J.T., Jarnevich, C.S., Holcombe, T.R., Talbert, C.B., Ignizio, D., Talbert, M.K., Silva, C., Koop, D., Swanson, A., Young, N.E., 2013. VisTrails SAHM: visualization and workflow management for species habitat modeling. Ecography 36, 129–135.
Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M., Anderson, R.P., 2014. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol. Evol. 5, 1198–1205.
Peet, R.K., 1981. Forest vegetation of the Colorado Front Range. Vegetatio 45, 3–75.
Peterson, A. T., Soberon, R.G., Pearson, R.P., Anderson, E., Martinez-Meyer, M., Nakamura, and Araujo, M.B. 2011. Ecological Niches and Geographic Distributions. Princeton University Press, Princeton, New Jersey, USA.
Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259.
R Core Team, 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL. http://www.R-project.org/ .
Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? J. Biogeogr. 33: 1689–1703. http://dx.doi.org/10.1111/j.1365-2699.2006.01466.x.
Rebelo, H., Jones, G., 2010. Ground validation of presence-only modelling with rare species; a case study on barbastelles *Barbastella barbastellus* (Chiroptera: Vespertilinidae). J. Appl. Ecol. 47, 410–420.
Royle, J.A., Chandler, R.B., Yackulic, C., Nichols, J.D., 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods Ecol. Evol. 3, 545–554.
Stohlgren, T.J., Falkner, M., Schell, L.D., 1995. A modified-Whittiker nested vegetation sampling method. Vegetatio 117, 121–133.
Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285–1293.
Theobald, D.M., Norman, J.B., Peterson, E., Ferraz, S., Wade, A., Sherburne, M.R., 2006. Functional Linkage of Water basins and Streams (FLoWS) v1 User's Guide: ArcGIS tools for Network-based analysis of freshwater ecosystems. Natural Resource Ecology Lab. Colorado State University, Fort Collins, CO. Pgs, p. 43.
Underwood, E.C., Klinger, R., Moore, P.E., 2004. Predicting patterns of non-native plant invasions in Yosemite National Park, California, USA. Divers. Distrib. 10, 447–459.
Visser, H., de Nijs, T., 2006. The map comparison kit. Environ. Model. Softw. 21, 346–358.
West, A.M., Kumar, S., Wakie, T., Brown, C., Stohlgren, T.J., Laituri, M., Bromberg, J., 2015. Using high-resolution future climate scenarios to forecast *Bromus tectorum* invasion in Rocky Mountain National Park. PLoS One 10.
West, A.M., Evangelista, P.H., Jarnevich, C.S., Young, N.E., Stohlgren, T.J., Talbert, C., Talbert, M., Morisette, J., Anderson, R., 2016. Integrating remote sensing with species distribution models; mapping tamarisk invasions using the Software for Assisted Habitat Modeling (SAHM). J. Vis. Exp. (116), e54578 http://dx.doi.org/10.3791/54578.
Western Regional Climate Center (WRCC), 2009. http://www.wrcc.dri.edu .