academicJournals

Vol. 5(3), pp. 122-134, March 2013
DOI: 10.5897/IJBC12.057
ISSN 2141-243X © 2013 Academic Journals
http://www.academicjournals.org/IJBC

International Journal of Biodiversity and Conservation

Full Length Research Paper

Estimating suitable environments for invasive plant species across large landscapes: A remote sensing strategy using Landsat 7 ETM+

Kendal E. Young[1,4]*, Laurie B. Abbott[1], Colleen A. Caldwell[2] and T. Scott Schrader[3,5]

[1]Department of Animal and Range Sciences, New Mexico Cooperative Fish and Wildlife Research Unit, New Mexico State University, Las Cruces, NM 88003, USA.
[2]U.S. Geological Survey, New Mexico Cooperative Fish and Wildlife Research Unit, New Mexico State University, Las Cruces, NM 88003, USA.
[3]Department of Fish, Wildlife and Conservation Ecology, New Mexico Cooperative Fish and Wildlife Research Unit, New Mexico State University, Las Cruces, NM 88003, USA.
[4]U.S. Forest Service, Hathaway Pines, CA 95233, USA.
[5]U.S.D.A. Jornada Experimental Range, Las Cruces, NM 88003, USA.

Accepted 11 March, 2013

The key to reducing ecological and economic damage caused by invasive plant species is to locate and eradicate new invasions before they threaten native biodiversity and ecological processes. We used Landsat Enhanced Thematic Mapper Plus imagery to estimate suitable environments for four invasive plants in Big Bend National Park, southwest Texas, using a presence-only modeling approach. Giant reed (*Arundo donax*), Lehmann lovegrass (*Eragrostis lehmanniana*), horehound (*Marrubium vulgare*) and buffelgrass (*Pennisteum ciliare*) were selected for remote sensing spatial analyses. Multiple dates/seasons of imagery were used to account for habitat conditions within the study area and to capture phenological differences among targeted species and the surrounding landscape. Individual species models had high (0.91 to 0.99) discriminative ability to differentiate invasive plant suitable environments from random background locations. Average test area under the receiver operating characteristic curve (AUC) ranged from 0.91 to 0.99, indicating that plant predictive models exhibited high discriminative ability to differentiate suitable environments for invasive plant species from random locations. Omission rates ranged from <1.0 to 18%. We demonstrated that useful models estimating suitable environments for invasive plants may be created with <50 occurrence locations and that reliable modeling using presence-only datasets can be powerful tools for land managers.

Key words: Invasive species, presence-only modeling, Landsat 7 ETM+, Maxent.

INTRODUCTION

Invasive species introductions have led resource managers to explore predictive theories of community attributes that increase their susceptibility to invasion (Rejmanek and Richardson, 1996). Unfortunately, universal predictive theories are not available because traits associated with invasion potential vary by geographic location, species and habitat (Alpert et al., 2000; Sakai et al., 2001). Nevertheless, habitats that are subject to altered disturbance regimes, have a history of prior invasion, provide adequate soil and water resources, or experience

*Corresponding author. E-mail: kyoung@nmsu.edu or kendalyoung@fs.fed.us. Tel: (209) 795-1381. Ext: 355. Fax: (209) 795-6849

high frequency of propagule introductions and are at greater risk to invasive plant incursion (Alpert et al., 2000; Rejmanek, 2000).

The U.S. National Park Service reported that exotic plants infested approximately 1.1 million ha in national parks (National Park Service, 2005). For example, over 65 exotic plant species that could potentially influence environmental quality, biotic health and ecosystem integrity were documented in Big Bend National Park, (Big Bend National Park, 1998; National Parks Conservation Association, 2003).

The key to reducing ecological and economic damage caused by invasive species is to locate and eradicate new invasions before they threaten native biodiversity and ecological processes (Stohlgren et al., 1999). Costs associated with ground-based reconnaissance preclude resource managers from conducting comprehensive inventories across large landscapes. Further, some invasive plants remain dormant between introduction and expansion and may go undetected during surveys. If surveys are not conducted at sufficient intervals, these undetected plants can rapidly spread once environmental conditions are favorable. Unfortunately, as the severity of the invasion rapidly increases, so does the cost of managing invasive plants.

Detecting suitable invasive plant environments using remotely sensed data

Remotely sensed data can provide a cost-effective tool to estimate environments suitable for invasive plants across large landscapes. Once potential areas of suitable habitat for invasive species are predicted, selective ground reconnaissance can be effectively used for verification and control. Potential suitable areas can also be frequently monitored to determine if undetected plants exist in the area. There has been a substantial increase in the use of remotely sensed and GIS data to model invasive species distributions or potential habitats as well as identify locations that may be at risk of plant incursion (Joshi et al., 2004; Franklin, 2009). This increase coincides with improved development and implementation of classification techniques, remote sensing apparatus and computer technology (Lass et al., 2005).

The success of remote sensing data to detect invasive plants or habitats is dependent upon the sensors' spatial and spectral (bandwidth) resolution and the sensors' repeat cycle (temporal resolution). Sensors that yield high spatial resolution data (<5 m) with hyperspectral (>100 spectral bands) capabilities have the highest likelihood of detecting microhabitats or rare plants (Marcus et al., 2003; Lass et al., 2005; Lawrence et al., 2006). However, these data also tend to be expensive and their relatively small swath size (ground area represented within the image) requires extensive computer processing time and storage for analyses of large areas. As such, landscape scale modeling endeavors often compromise bandwidth and spatial resolution considerations.

Landsat Thematic Mapper (Landsat TM) and Landsat Enhanced Thematic Mapper Plus (Landsat 7 ETM+) are multispectral, medium spatial resolution (30 m) sensors and are well suited for modeling endeavors across large landscapes. Landsat TM and Landsat 7 ETM+ have been extensively used to model vegetation types across a variety of landscapes. In some cases, these sensors can identify individual plant species with unique spectral or temporal characteristics (Parker-Williams and Hunt, 2002). Dewey et al. (1991) compared Dyer's woad (*Isatis tinctoria*) locations with spectral classes created from Landsat 5 TM data in northern Utah and observed strong associations between plant locations and 10 spectral classes. The authors demonstrated a remotely sensed predictive model that provided resource managers with a tool for estimating the plant's potential distribution. In the Great Basin, cheatgrass (*Bromus tectorum*) was modeled using Landsat TM and Landsat 7 ETM+ data (Bradley and Mustard, 2005). Shafii et al. (2003) predicted yellow starthistle (*Centaurea solstitalis*) using a land use classification based on Landsat data and other GIS datasets.

Landsat TM and Landsat 7 ETM+ data are best used for detecting plants that have patch sizes of 0.5 ha or larger (Everitt and DeLoach, 1990; Everitt et al., 1992; Anderson et al., 1993). Similar spectral signatures between targeted plants and the surrounding environment, changes in soil color or moisture, and low plant densities hinder discrimination efforts for invasive plants. However, seasonal differences in plant phenology may enhance the detection of invasive plants due to flowering or green-up at different periods than the native surrounding vegetation. Multiple dates of imagery allow the detecting of these phenological differences between targeted plants and the surrounding landscape. Price et al. (2002) noted that increasing the number of Landsat TM bands by using multiple dates of imagery improved discrimination accuracy of grassland types. As such, imagery dates should correspond to critical phenological phases of the targeted plant (Zhang et al., 2003; Joshi et al., 2004).

Presence-only predictive models

Predictive habitat models are grounded in ecological niche theory and are quantitatively related to the likelihood of a species occurrence given a set of predictor variables (Franklin, 2009). One can discriminate between response values (e.g., species presence and absence) using a set of environmental predictors. Many analytical approaches have been applied to determine statistical relationships between species and predictor variables (Guisan and Zimmermann, 2000; Elith, 2002; Franklin, 2009). Several of the analytical approaches (for example, general linear and logistic regression models) require presence/absence data (Pearce and Ferrier, 2000; Manel et al., 2001; Elith, 2002). This requirement is a serious limitation for species distribution models, since many species have detection probabilities <1. A sampling problem is the failure to detect a species' presence in suitable habitats with the implication that non-detection occurrences represent species' absence (Mackenzie et al., 2002; Weidong and Swihart, 2004). As such, there is a growing interest in using presence-only data in modeling efforts (Elith, 2002; Graham et al., 2004; Argaez et al., 2005).

Considerable research has focused on the creation of species distribution models using presence and pseudo-absence (e.g., background data) (Zaniewski et al., 2002; Elith et al., 2006; Oliver and Wotherspoon, 2006). Background locations are randomly selected points generated throughout the entire study area or region which may or may not include locations where the species was present. As such, analyses of background data are sensitive to the size of the analytical area (VanDerWal et al., 2009; Phillips et al., 2009). Unfortunately, background locations do not incorporate ecological knowledge of the species-habitat relationships. Nevertheless, this type of modeling is effective for modeling species' potential distribution, including rare species (Elith et al., 2006).

The goal of this study was to investigate methods to assist in the early detection of invasive plants over a large area. Using data and analytical methods that are typically accessible to most land management agencies, we evaluated the efficacy of using Landsat 7 ETM+ imagery to estimate suitable environments for invasive plants across Big Bend National Park. Since reliable absence data for the target plants were not available, a presence-only modeling and validation approach was employed.

METHODS

Big Bend National Park is located in the northern Chihuahuan Desert in southwest Texas. The Park encompasses roughly 324,154 ha with elevation ranging from 518 m at the Rio Grande to 2,388 m at Emory Peak in the Chisos Mountains (Big Bend National Park, 2004). Temperatures range from over 37°C in the summer to -17°C in the winter; although average summer and winter temperatures are 27°C and 3°C, respectively (Cochran and Rives, 1985). Mean annual precipitation in the Park is approximately 33 cm, of which 75% falls from April to September as heavy thunderstorms (Cochran and Rives, 1985).

Big Bend National Park has three major environmental zones: the Chihuahuan Desert, the Rio Grande and its riparian corridor, and the Chisos Mountains (National Park Service, 1983). The Rio Grande defines 190 km of the Park's southern boundary.

Creosotebush (*Larrea tridentata*) and lechuguilla (*Agave lechuguilla*) are the dominant plant species in the Chihuahuan Desert community, occupying approximately 72% of the Park (Plumb, 1991). Riparian vegetation along the river is dominated by saltcedar (*Tamarix* spp.), mesquite (*Prosopis* spp.), cottonwood (*Populus deltoides*), willow (*Salix* spp.), tree tobacco (*Nicotiana glauca*), Bermuda grass (*Cynodon dactylon*), and giant reed (*Arundo donax*). The Chisos Mountains are located near the center of the Park and considered the southernmost mountains in the continental United States (Wauer, 1996). Dominant woodland vegetation includes alligator juniper (*Juniperus deppeana*), red-berry juniper (*J. pinchotii*), weeping juniper (*J. flaccida*), gray oak (*Quercus grisea*), Emory oak (*Q. emoryi*), Grave's oak (*Q. gravesii*) and Chisos oak (*Q. graciliformis*). Mexican pinyon (*Pinus cembroides*), Ponderosa pine (*P. ponderosa*), quaking aspen (*Populus tremuloides*) and Arizona cypress (*Cupressus arizonica*) are also found in isolated patches.

Predictive habitat models

Maxent software version 3.3.1 (http://www.cs.princeton.edu/~schapire/maxent/) was used to create predictive habitat models (Phillips et al, 2004, 2006). Maxent uses the principle of maximum entropy to estimate the target probability distribution that has the broadest distribution compatible with the information available (Phillips et al., 2004; Dudík et al., 2007; Phillips and Dudík, 2008). The Maxent procedure employs a maximum-likelihood to generate a probability distribution grid of the analysis area. The program begins with a uniform distribution, and performs multiple iterations that increase the probability of the species sample locations (Yost et al., 2008). Maxent uses pixels with known species occurrence records and randomly selected background points to constitute sample points. Landsat 7 ETM+ and GIS datasets provided the environmental variables measured at each sample point. A more detailed description of how the program functions can be found in the software tutorial, help section and from Phillips et al. (2004, 2006, 2009), Phillips and Dudík (2008) and Elith et al. (2010).

The program default settings as described by Phillips and Dudík (2008) were used for analyses, including the default regularization multiplier to reduce over-fitting (Pearson et al., 2007). Default settings included maximum iterations = 500, convergence threshold = 10^{-5}, number of background points = 10,000 and auto features that would select the appropriate feature function (linear, quadratic, product, threshold or hinge) based on the number of presence records (Phillips et al., 2006; Pearson et al., 2007; Peterson et al., 2008). Program outputs include a logistic probability surface, and tabular and graphical representations of model performance and variable contribution.

Input occurrence data

Four invasive plant species were selected for remote sensing spatial analyses: giant reed, Lehmann lovegrass (*Eragrostis lehmanniana*), horehound (*Marrubium vulgare*) and buffelgrass (*Pennisteum ciliare*). These species were National Park Service priority species, had been or were currently undergoing evaluation within Big Bend National Park, with known spatial locations of occurrence in the Park. Reliable spatial information on plant absences was not available.

Giant reed, native to eastern Asia (Bell, 1997), is thought to be first introduced into the U.S. near Los Angeles, California, for erosion control along drainage canals (Hoshovsky, 1987). In Big Bend National Park, giant reed displaces native species and forms dense stands along waterways (Photo 1). Stands create flood-control problems, increase fire hazard, and reduce biodiversity and habitat for wildlife. Lehmann lovegrass originates from southern Africa (Anable et al., 1992), and was introduced in the southwest in 1932 (Cox et al., 1988). Starting in the early 1950s, commercial seed growers produced large quantities of Lehmann lovegrass which was planted from Texas to Arizona to prevent soil erosion and provide livestock forage (Cox et al., 1988; McClaran and Anable, 1992). Lehmann lovegrass spreads aggressively into desert and grassland communities where it excludes native plants, especially after disturbance. This plant alters fire frequency and intensity, and reduces biodiversity (Photo 2). Horehound is native to northern Africa, Asia and Europe, and is now common across most parts of

Photo 1. Giant reed along the Rio Grande, Big Bend National Park.

Photo 2. Lehmann lovegrass clump in southern New Mexico.

Photo 3. Horehound inflorescence, Big Bend National Park.

Photo 4. Buffelgrass clump along a culvert in Big Bend National Park.

the U.S. (Simon et al., 1984). Horehound colonizes disturbed sites and reduces native biodiversity (Photo 3). Buffelgrass is native to Africa, Asia, and the Middle East and was introduced into the southwestern U.S. for livestock forage (Holt, 1985; Cox et al., 1988; Ibarra et al., 1995). Buffelgrass colonizes disturbed sites and reduces biodiversity by altering fire frequency and intensity, crowds out native grasses and competes for limited resources (Photo 4).

Personnel of Big Bend National Park provided spatial data (collected by non-probability sampling procedures) on the occurrences of the target invasive plants. This dataset included present and historical plant locations (areas where eradication and remediation efforts were applied) from opportunistic sightings and planned roadside surveys from 2001 to 2005. Additional ground surveys were conducted in June 2006 throughout the Park to ascertain additional locations of target species. A systematic sampling approach was used to optimize the chance of detecting additional invasive plant populations. Roads, trails, campgrounds, facilities and infrastructure sites, disturbed sites, arroyos, and springs were systematically surveyed throughout the Park. Additional surveys (randomly selected) occurred away from features described above in undisturbed areas. Once a targeted species was located, Global Positioning System (GPS) locations (UTM Zone 13, GRS 1980 Spheroid and NAD 83 Datum) and digital photographs were collected. Spatial data provided by Park personnel and from the additional ground survey efforts were consolidated for predictive modeling.

Presence-only datasets and non-probability sampling procedures offer a variety of challenges for creating predictive models, since most species distribution models rely on the collection of unbiased samples (Franklin, 2009; Elith et al., 2010).

We reduced the bias associated with presence-only datasets by eliminating duplicate records (spatial autocorrelation) and by restricting the selection of our background samples to areas that were

Table 1. Dates and seasons of imagery used to model potential invasive plant habitat in Big Bend National Park.

Year	Season	Scene 3140	Scene 3040
1999	Fall	September 30	September 23
2000	Summer	May 27	July 23
2000	Fall	October 02	November 28
2001	Spring	April 28	March 20
2001	Fall	October 21	October 14
2002	Spring	March 30	April 8
2003	Spring	January 28	March 26

surveyed for invasive plants (Phillips et al., 2009; Elith et al., 2010). All duplicate records that were within 30 m of each other were removed to ensure that each Landsat 7 ETM+ pixel would only host one occurrence record. Model predictions were then projected to the areas that were not searched by using the projection facilities in Maxent.

Environmental datasets

We used Landsat 7 ETM+ spectral data to represent environmental variables for this study. The non-classified spectral data were free of classification errors (typical in vegetation community or landscape maps) and provided a dataset that was free from biases associated with interpretations of suitable or unsuitable areas. Likewise, Landsat 7 ETM+ provided a computationally efficient scale for spatial analyses across the entire park.

The use of Landsat 7 ETM+ limited the likelihood of detecting small populations or individual plants, as the plant's reflectance value would have been masked by more dominant reflectance values that occurred in the same pixel. Importantly, Landsat 7 ETM+ did not limit the likelihood of detecting landscape features that are associated with invasive plant occurrences. As such, the predictive habitat models are based on discriminating spectral values associated with landscape features at invasive plant locations and spectral values at background locations.

Further, some of the target invasive plants had unique vegetation phenology or distinct habitat associations. As such, multiple dates and seasons of imagery (Table 1) were used to account for dynamic habitat conditions that result from discrete rainfall events. Multiple dates and seasons of imagery also provided a means to capture phenological differences among the targeted species and the surrounding landscape.

We acquired Fall 1999, Summer and Fall 2000, Spring and Fall 2001, Spring 2002, and Spring 2003 Landsat 7 ETM+ data from America View (http://www.americaview.org) and the Texas View Remote Sensing Consortium (http://www.texasview.org). Data were obtained pre-processed to Level 1-G, which includes radiometric and geometric correction. Two Landsat scenes were required for complete coverage of the Park: WRS3140 and WRS3040. Imagery was acquired in UTM Zone 13, WGS 84 projection and was reprojected to UTM Zone 13, GRS 1980 Spheroid and NAD 83 Datum.

Image mosaic was performed in Erdas Imagine 9.0 using feather blending of overlapping regions and clipped to the Park boundary. Landsat 7 ETM+ bands 1, 2, 3, 4, 5 and 7 were extracted into Generic ASCII raster format for each year/season mosaic for data analyses needed by Maxent software. This resulted in 42 potential spectrally based environmental variables (seven season/years with six bands each) used for analyses. Finally, a Digital Elevation Model (DEM) clipped to the Park boundary and extracted into Generic ASCII raster format was used to estimate altitude associations resulting in 43 potential environmental variables used for analyses.

Variable selection

Maxent has a method for variable selection (regularization parameter). The intent of the regularization parameter is to determine the most parsimonious model possible, that is, a model that provides a balance between the extremes of having too few parameters (under - fitting) and models that have too many parameters (over-fitting) (Burnham and Anderson, 1992). Maxent fits a penalized maximum likelihood model analogous to Akaike's Information Criterion (Elith et al., 2010). Maxent's regularization parameter is fairly stable with regards to correlated variables, reducing the need to remove correlated variables (Elith et al., 2010).

Maxent also uses a jackknife approach to evaluate which variables are most important in the model. In addition, Maxent creates response curves that evaluate the contribution of a variable in relation to the mean of all other variables for the occurrence locations, and provides a tabular output of variable percent contribution.

Analyses provided in Maxent were used to guide model variable selection (default regularization parameter and model variable contribution). Initially, a model was created that included all 43 environmental variables (global or full model). We noted between 16 and 37 variables were retained, with several variables contributing <1% to the full model. We then created another spatial model that excluded variables contributing <1% to the full model. The percent contribution of variables in this revised model was inspected. Variables with <3% contribution to the revised model were then excluded, and another spatial model was created. The intermediate step of excluding variables with <1% contribution was necessary because some variables that were slightly under 3% contribution in the full model would return ≥3% after the first variable elimination step. These iterative steps greatly reduced the number of unnecessary variables retained in the models. We then tested the resulting model for variable correlation, since a model with fewer variables (thus higher percent contributions) that were correlated may mislead interpretations. Pearson's correlation coefficient was calculated using Proc Corr in Statistical Analysis Software (SAS 9.1). Correlated variables ($r \geq 0.8$) were removed from the final model by retaining the variable with the highest model contribution, as determined by Maxent analyses. The random seed function in Maxent was not used for this variable selection process, thus ensuring that the same dataset was used for each elimination step.

Model performance

Maxent uses threshold-dependent and threshold-independent metrics to evaluate model performance. Threshold-dependent metrics require a known threshold to classify a response as presence/absence, or suitable/not suitable. Maxent output yields a variety of threshold-dependent values. Users can choose a threshold value based on their objectives (Fieldings and Bell, 1997). For example, research or management objectives may require greater emphasis placed on the ability to accurately predict species presence (sensitivity) rather than species absence (specificity). Under those conditions, a threshold weighted towards sensitivity would be selected.

Maxent threshold-dependent metrics uses a one-tailed binomial test to evaluate if a model performed significantly better than random (Phillips et al., 2006). Maxent uses the omission rate (fraction of the test localities that fall into pixels not predicted as suitable) and the proportion of all the pixels that are predicted as suitable habitat to estimate training and testing omission rates.

Threshold-independent metrics are not based on the selection of a specified threshold for classifying the predicted observation into binomial outcomes. Instead, model performance is evaluated across the continuum of thresholds from 0 to 1.0, e.g., receiver operating

Table 2. Number of geo-referenced target invasive plant occurrences (individual plants or populations) and number of training samples used to construct predictive habitat models in Big Bend National Park, TX.

Plant name	Number of known occurrence	Number of training sample	Number (percent) of test sample	
Giant Reed	31	26	5	(17%)
Lehmann lovegrass	253	208	45	(18%)
Horehound	26	21	5	(20%)
Buffelgrass	627	552	75	(12%)

characteristic (ROC) curves. ROC analyses are also independent of species prevalence (Pearce and Ferrier, 2000). The area under the ROC curve (AUC) provides a summary measure of the model's discrimination ability, that is, its ability to differentiate suitable from unsuitable habitats (Phillips et al., 2006). An AUC value of 0.5 indicates the model preformed no better than a random prediction; values between 0.5 and 0.7 indicate low discrimination ability; values between 0.7 and 0.9 indicate moderate discrimination ability; and values >0.9 indicate high discrimination ability (Pearce and Ferrier, 2000; Manel et al., 2001).

Data used to test model performance were obtained by randomly partitioning plant occurrence locations into training data and independent test data (Fieldings and Bell, 1997). Replicate models (n = 30) per plant species were conducted to assess the average behavior of the predictive habitat models (Phillips et al., 2006; Yost et al., 2008). Each partition was created through a bootstrap procedure where 12 to 20% of the occurrence locations for each target species were randomly selected with replacement. The random seed function in Maxent ensured a different random subset of background locations for each replication. Model performance was evaluated using the average testing omission rate, average test AUC metrics, and the average regularized training gain (Phillips et al., 2006). Gain can be interpreted as representing how much better the distribution fits the sample points than the uniform distribution. In addition, 95% confidence intervals were estimated for the average test AUC and omission rate.

The final model used to identify suitable environments for the target plant species was created by averaging 30 replicate grids in Maxent. The amount of suitable habitat for each invasive plant species was estimated by applying the average logistic threshold that provided an equal tradeoff between test data sensitivity and specificity (from Maxent output).

RESULTS

Invasive species occurrences

Approximately 400 h were spent systematically searching roads, trails, campgrounds, rivers, drainages, developed areas, and random locations for invasive plants. Combining this survey effort with known occurrences documented by Big Bend National Park, and removing duplicate records, a total of 937 invasive plant locations were documented for the four species of interest (Table 2).

The number of occurrence points used to generate individual species models ranged from 26 to 522 (Table 2). Small sample sizes (<50 occurrences) occurred with giant reed and horehound. As a result, only five samples were withheld to evaluate model performance for these two species. In contrast, there were >250 occurrences of Lehmann lovegrass and buffelgrass, which allowed for ≥45 samples to be used to evaluate model performance.

Performance of predictive models

All predictive plant models performed better than a random prediction (P < 0.001). Average test AUC values (0.91 to 0.99) and their associated 95% confidence intervals indicated that plant predictive models had high discriminative ability to differentiate suitable environments for invasive plant species from random background locations (Figure 1a). The average testing omission rate ranged from <1.0 to 18% (Figure 1b). Despite small sample sizes for horehound and giant reed, the 95% confidence intervals for test omission rates were low, between 0 and 2% for horehound, and 1v and 7% for giant reed (Figure 1b).

In addition to the high test AUC values (0.99), the horehound and giant reed models yielded high model gain values (≥3.0) but moderate errors of training omission, ranging from 11 to 14% (Table 3). Average test AUC values for the buffelgrass (0.94 ± 0.01) and Lehmann lovegrass (0.91 ± 0.01) were high, although, these species models yielded low (≤1.4) training model gain, indicating that the average sample likelihood was ≤ 4.1 times higher than that of a random background pixel (Table 3). Test omission rate for buffelgrass was 15% (P < 0.001), and 18% (P < 0.001) for Lehmann lovegrass.

Test AUC value, regularized training gain, and test omission rate were chosen to evaluate predictive model performance. Analyses indicated that these three metrics were highly correlated. On average, as the test AUC value increased, the model gain metric increased (r = 0.95), and test omission rate decreased (r = -0.99; Figure 2).

Environmental variables retained

Our variable selection process reduced the number of environmental variables retained in the models to ≤7 variables (Table 4). Only 14 of the 43 model input variables were retained in the predictive models. The combination of variables and their percent contribution to the final predictive models were different for each species, although Band 4 in the Summer 2000 dataset contributed to all species models. All plant species models retained at least three years of data, across either two or three seasons, which emphasizes the importance of temporal datasets to detect plant phenological changes.

Predicted suitable environments

Our models reveal that giant reed may have the most limited

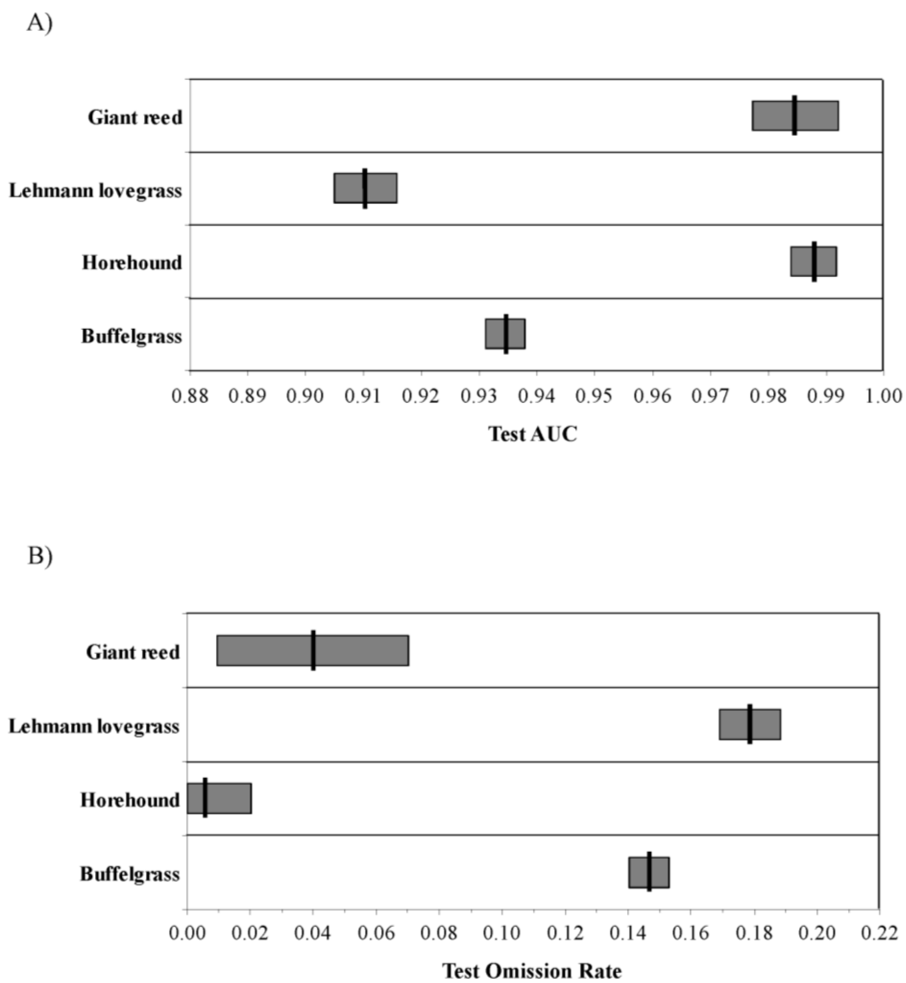

Figure 1. Performance metrics for invasive plant species predictive models in Big Bend National Park. A) Average test AUC (vertical bars) and 95% confidence intervals (shaded boxes). Discrimination ability is high when values range from 0.90 TO 0.99. B) Average test omission rate (vertical bars) and 95% confidence intervals (shaded boxes).

Table 3. Average performance measurements for invasive species predicted habitat models in Big Bend National Park. Averages were calculated from 30 replicate models created per plant species.

Common Name	Threshold independent			Threshold dependent			
	Training	Test		Equal test sensitivity and specificity			
	Gain[a]	AUC[b]	SD	Logistic Threshold	Training omission rate	Test omission rate	Test omission P-value
Giant reed	3.397	0.985	0.011	0.321	0.135	0.040	1.1E-03
Lehmann lovegrass	1.332	0.910	0.014	0.343	0.141	0.179	5.5E-18
Horehound	3.008	0.988	0.007	0.320	0.113	0.007	7.7E-05
Buffelgrass	1.357	0.935	0.009	0.383	0.116	0.147	6.3E-46

[a]Regularized model training gain; [b]AUC = Area under the curve derived from receiver operating characteristic (ROC) curves for each plant species.

potential distribution in the Park, with approximately 3,776 ha (1%) of the Park surface modeled as potential habitat. Giant reed's modeled distribution was primarily along the Rio Grande corridor where large stands are well established (Figure 3). The predicted potential distribution of horehound was also limited to approximately 16,028 ha (5%) of the Park, and was predicted to occur in developed areas near the Chisos Mountains and small patches in the

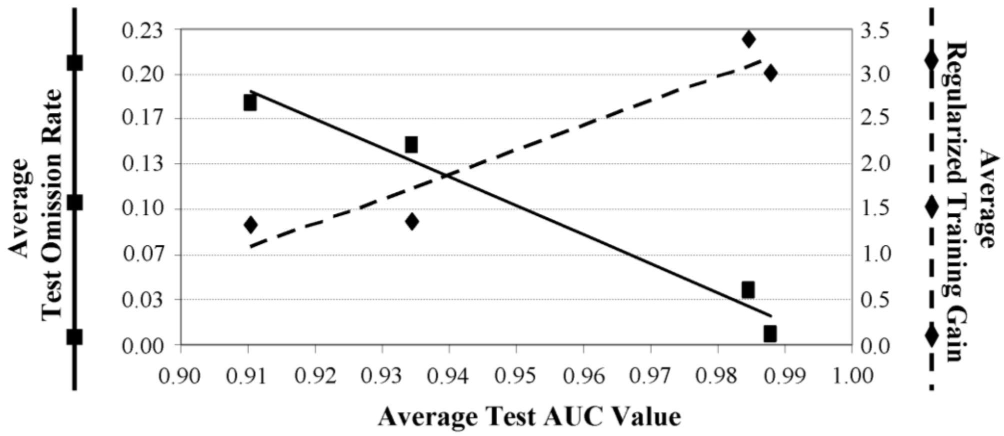

Figure 2. Relationship between average test AUC value, test omission rate and the regularized training gain for predicted habitat models of four invasive plant species in Big Bend National Park.

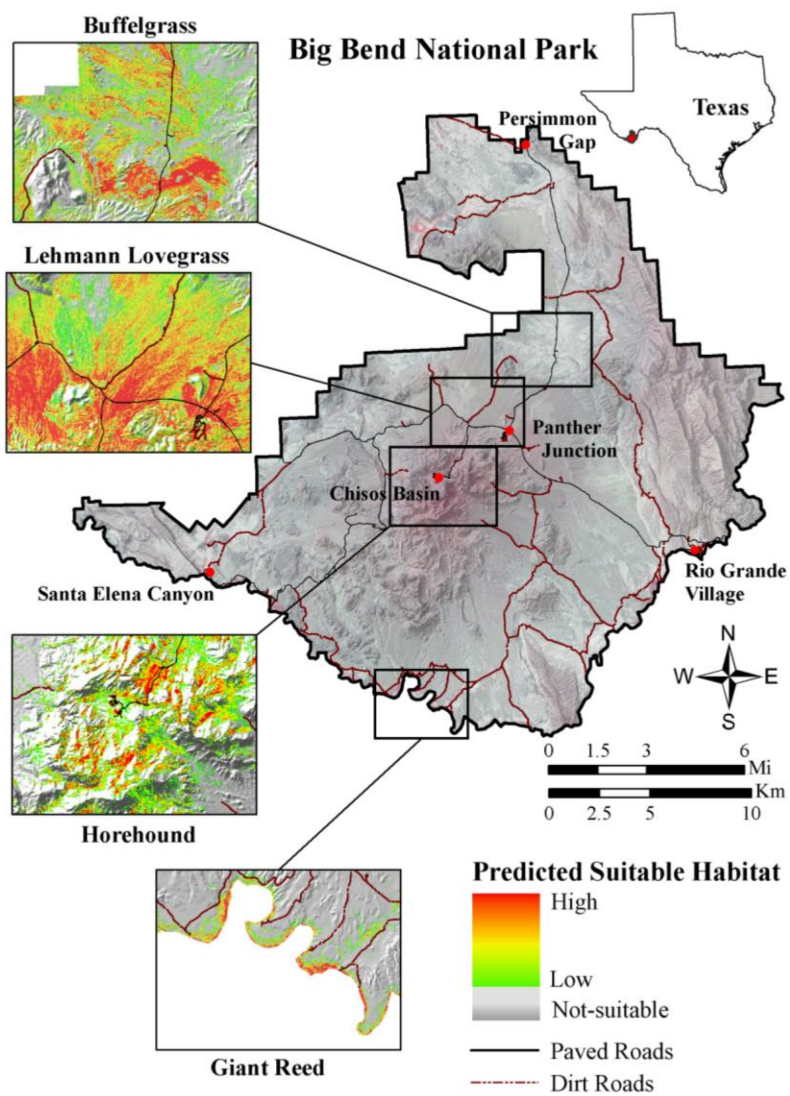

Figure 3. Landsat 7 ETM+ imagery of Big Bend National Park with key predicted suitable habitat areas (in call out boxes) that warrant monitoring efforts for giant reed, Lehmann lovegrass, horehound and buffelgrass in Big Bend National Park.

Table 4. Average percent contribution of each environmental variable retained in each invasive plant species predictive habitat model. Spaces with a "-" indicate the variable was not retained in the final predictive model.

Date	Model variable	Giant reed	Lehmann lovegrass	Horehound	Buffelgrass
Fall 1999	DEM	57.5	-	-	-
	Band 1	-	10.0	38.7	-
	Band 5	9.0	-	-	-
Summer 2000	Band 4	3.2	27.9	13.5	35.2
Fall 2000	Band 2	-	-	-	5.7
	Band 4	-	-	29.9	12.2
	Band 5	-	46.5	-	-
	Band 6	5.8	-	-	-
Spring 2001	Band 4	-	-	5.7	4.3
	Band 5	-	15.6	-	14.8
	Band 6	18.5	-	-	-
Fall 2001	Band 4	6.1	-	-	-
	Band 6	-	-	12.2	10.5
Spring 2003	Band 5	-	-	-	17.2

the northern part of the park near the Rosillos Mountains (Figure 3).

Buffelgrass presents greater potential to spread throughout Big Bend National Park as evidenced in our model which described 103,474 ha (32%) of the Park as suitable for this species. Lehmann lovegrass exhibited similar potential distributions in the Park, with 127,807 ha (40%) of suitable environments.

DISCUSSION

Performance of predictive models

The evaluation of multiple models is often warranted because natural landscapes and associated species habitat relationships are typically too complex to estimate in a single model (McComb et al., 2002). Creating replicate models by bootstrapping species occurrence locations and background points allowed for a greater level of confidence in the modeling results, especially for species models created with small sample sizes (e.g., giant reed and horehound).

Although, the sample sizes used to construct and test horehound and giant reed models were small, model AUC and gain values indicated a high likelihood that a random positive occurrence and a random negative location would be accurately predicted. The average sample likelihood was ≥20 times higher than that of a random background pixel. Likewise, the ranges of test omission rates for these species were low. Horehound is primarily restricted to small isolated patches in the Park, whereas giant reed co-occurs with saltcedar along the Rio Grande. The co-occurrence of species may have increased the variability associated with spectral values at giant reed locations and the reduced precision of the model.

Elith et al. (2006) compared 16 modeling methods over 226 species from six regions of the world and were unable to relate sample size to modeling success. Phillips et al. (2004) recommended 50 to 100 samples for optimal Maxent models. Peterson et al. (2007) expressed caution with models that were evaluated with <5 occurrences, and suggest a jackknife validation approach when sample sizes are <25. Wisz et al. (2008) noted that Maxent was less sensitive to sample size considerations as compared to other modeling procedures. Conversely, Hernandez et al. (2006) evaluated four modeling methods, including Maxent, using different sample sizes of 18 species representative of different levels of ecological specialization. The authors concluded that Maxent was the most capable of the four modeling approaches in producing models with sample sizes between 5 and 25 occurrences. Our results were similar to those of Hernandez et al. (2006) and clearly demonstrate the utility of invasive species predictive habitat models created with a small number of sample locations and a maxi-mum entropy analytical approach. Approaches that are robust to small sample size are critical to early detection needs.

Although, the buffelgrass and Lehmann lovegrass predictive models yielded high discrimination ability, they also yielded the highest omission rates (15% to 18%). Buffelgrass can be found in mesic environments throughout Big Bend National Park in large-sized patches that are best distinguished from the surrounding landscape during the appropriate phenological stage. Lehmann lovegrass colonizes disturbed areas and establishes easily in roadside areas. Most of the Lehmann lovegrass occurrence records were associated with roadside areas which support the general observation that the species colonizes disturbed areas and establishes easily in roadside areas.

The spatial resolution (30 m) of Landsat 7 ETM+ likely contributed to high omission rates for these species, as small mesic environments and linear features (small roads and drainages) would not have been well represented on the imagery. From a conservation perspective, errors of omission are less acceptable than errors of commission (Shrader-Frechette and McCoy, 1993). If the predictive model concluded an area was suitable for an invasive plant species, and the species was not present (commission), then the area would be a good candidate to monitor for potential expansion of the species. Of the three Maxent metrics chosen to evaluate predictive model performance (test AUC value, regularized training gain, test omission rate), our analyses indicated that model performance was adequately addressed using the AUC and omission metrics. These two metrics were correlated in this study. However, it is possible to achieve high AUC values and high omission rates (Peterson et al., 2008). Thus, considering the model omission rate in addition to the AUC metric may help in detecting a poorly performing predictive model.

Environmental variables retained

A model that incorporates a large number of environmental variables is considered to have the highest flexibility in fitting the observed data (White, 2001). However, a large number of input variables increases the amount of variability and decreases the precision of the model, effectively decreasing the ability to accurately predict suitable habitat.

Conversely, models with fewer parameters are more precise because all the data are being used to estimate the parameters; however, these models may have more bias associated with them. In other words, as the number of variables increases, bias is reduced but precision is lost (Franklin et al., 2001). Phillips et al. (2004) noted that Maxent models which incorporate a larger number of variables tend to overfit small training sets, but they provide a more accurate prediction for large training sets. Likewise, VanDerWal et al. (2009) found that as the analytical area increased, predictive models retained fewer environmental variables. Thus, the size of the analytical area also affects precision and bias.

In this study, the regularization parameter variable used in Maxent is considered to be effective in selecting important variables and disregarding unimportant ones. However, when models were created utilizing the default regularization parameter, the resulting models retained a large number of variables with low percent contributions. As such, our study removed variables that yielded <3% contribution to the models and removed correlated variables to obtain a more parsimonious model. Another approach to variable selection would be to increase the regularization parameter value in Maxent (Elith et al., 2010).

The combination of variables and their percent contribution to the final predictive models were different for each species, indicating that there was no single best combination of remotely sensed data that would adequately describe environmental conditions of occupied areas. However, Band 4 wavelength may be important in temporal analyses. Landsat 7 ETM+ Band 4 represents the near infrared (NIR) spectrum (wavelength 0.75 to 0.90 µm) and is well suited for vegetation discrimination as reflectance values are much higher than in the visible bands due to leaf cellular structure. At least two seasons of Band 4 were retained for most of the predictive habitat models, with the exception of Lehmann lovegrass.

Predicted suitable environments

Maxent outputs include a logistic surface with values ranging from 0 to 1, representing the probability of the pixel hosting suitable environmental conditions for the species. This allows resource managers to prioritize their monitoring, control or conservation efforts, focusing on areas with high probabilities of suitable environments, followed by lower probability values. Comparisons of logistic values between species may be inappropriate because probability of presence is only defined relative to a given level of sampling effort (Elith et al., 2010), which may not be equal for different species.

Additionally, threshold values can be chosen based on specific management objectives, and allow for differentiating between suitable and unsuitable environments. Habitat suitability thresholds are often selected subjectively (Hirzel et al., 2006), and there has been little research to suggest appropriate thresholds for presence-only modeling endeavors (Hirzel et al., 2006; Phillips et al., 2006). In general, appropriate thresholds yield low omission rates (Peterson et al., 2007; Phillips, 2008). VanDerWal et al. (2009) used a "balance" threshold, while Yost et al. (2008) reported results from three different thresholds when evaluating a predictive model for sage grouse (*Centrocercus urophasianus*). We chose a threshold that provided an equal tradeoff between sensitivity and specificity, which would balance commission and omission errors.

Our modeled distributions of suitable environments were consistent with the published literature. For exam-ple, the model for giant reed revealed that potential suitable environments were primarily along the Rio Grande corridor where large stands are well established and is consistent with its preference for riparian and wetland habitat types, floodplains, plains and arroyos (Dick-Peddie, 1993; Tracy and DeLoach, 1998).

Horehound occurs in disturbed sites along roadsides, stockyards, fields, pastures, near dwellings and dry riverbanks in scattered patches (Sievers, 1930) which are consistent with our modeled results. Buffelgrass is actively spreading throughout Big Bend National Park and is found along roadside runoff areas, developed areas and previously disturbed sites. It is common and dense in Rio Grande Village, Boquillas Canyon and paved roads, and has been found at over 1,525 m (Big Bend National Park, 1998), although this is considered to be outside its normal elevation range. Further, Arriaga et al. (2004) noted that buffelgrass in northern Mexico have been known to invade desert scrub and mesquite woodlands.

Lehmann lovegrass is well established in disturbed roadside run off areas and developed areas (Big Bend National Park, 1998). This plant has not been found in lower elevation areas along the Rio Grande. It does not appear that this species is spreading into undisturbed communities of Big Bend; however, areas adjacent to established plants are susceptible if moisture regimes change or disturbances occur.

Conservation implications

Belovsky et al. (2004) noted that modeling activities that are removed from the underlying ecology of the organism may not be effective. While remotely sensed data may not fully represent all ecological parameters, remotely sensed spectral values may provide adequate surrogates to the location of existing populations and landscape features that promote or enhance invasions. We demonstrated that Landsat 7 ETM+ spectral values, in concert with accurate field data, can successfully be used to create reliable spatial models. Multi-temporal datasets captured a series of unique phenological characteristics and were able to differentiate invasive plant populations or their habitats from the surrounding landscape. Further, the use of Landsat 7 ETM+ removed any bias associated with human perception of suitable habitat, and provided an empirical estimate of areas where the target species may become established.

No doubt, detecting small or sparse plant populations is still hampered by spatial and spectral resolution, and by our ability to analyze large datasets. The optimal remote sensing data, or combination of data, would have characteristics of hyperspectral sensors and high spatial resolution sensors. While hyperspectral data facilitate detection of individual plants, hyperspectral data have approximately 75 times higher data volume than an equivalent area using Landsat 7 ETM+ (Thenkabail et al., 2004).

Likewise, multispectral, high spatial resolution sensors (for example, IKONOS or QuickBird) also show promise in detecting invasive plants with spatial resolutions <5 m. These sensors, however, are also encumbered by large data volumes when used on large areas. The new challenge will be to develop methods that integrate the required spectral resolution with the ideal spatial resolution, yet are efficient with the high-dimensional datasets for large area analyses.

We demonstrated that remotely sensed analyses can aid in the development of spatially explicit predictive models over large areas and can provide land managers with early detection tools, a means to evaluate current and future control needs, and a means to prioritize conservation efforts. Early detection methods increase our ability to eradicate invasive plants and ultimately reduce control costs (Rejmanek and Pitcairn, 2002). While predictive modeling is not a substitute for detailed collection of field data, Hernandez et al. (2006) demonstrated that reasonable models are appropriate for rare species. The results of this study are similar with their findings, and should encourage land managers to add predictive modeling to their toolbox.

ACKNOWLEDGEMENTS

The research was funded by the U.S. Geological Survey Status & Trends of Biological Resources Program and was conducted through close cooperation with the National Park Service. Additional support was provided by the New Mexico State University Agricultural Experiment Station, the Department of Fish, Wildlife and Conservation Ecology, and the U.S. Geological Survey, New Mexico Cooperative Fish and Wildlife Research Unit. We thank J. Sirotnak and B. Alex (Big Bend National Park) for their cooperation and assistance in conducting this study. We also thank K. Boykin, Q. Young and A. Ernst for their technical assistance and advice. Any use of trade names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

REFERENCES

Alpert P, Bone E, Holzapfel C (2000). Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants. Perspect. Plant Ecol. 3: 52-66.

Anable ME, McClaran MP, Ruyle GB (1992). Spread of introduced Lehmann lovegrass (*Eragrostis lehmanniana* Nees) in southern Arizona, USA. Biol. Conserv. 61: 181-188.

Anderson GL, Hanson JD, Hart GF (1993). Using satellite data to map false broomweed (*Ericameria austrotexana*) infestations on south Texas rangelands. Weed Technol. 7: 865-871.

Argaez JA, Christen JA, Nakamura M, Soberon J (2005). Prediction of potential areas of species distributions based on presence-only data. Environ. Ecol. Stat. 12: 27-44.

Arriaga LA, Castellanos E, Moreno E, Alarcon J (2004). Potential ecological distribution of alien invasive species and risk assessment: a case study of buffelgrass in arid regions of Mexico. Conserv. Biol. 18: 1504-1514.

Bell GP (1997). Ecology and management of *Arundo donax*, and approaches to riparian habitat restoration in southern California. In:

Brock JH, Wade M, Pysek P, Green D (eds.), Plant invasions: studies from North America and Europe. The Netherlands: Backhuys Publishers Leiden. pp. 103-113.

Belovsky GE, Botkin DB, Crowl TA, Cummins KW, Franklin JF, Hunter ML, Joern A, Lindenmayer DB, MacMahon JA, Margules CR, Scott JM (2004). Ten suggestions to strengthen the science of ecology. BioScience 54: 345-351.

Big Bend National Park (1998). Evaluation of the potential and current impacts of exotic plant species in Big Bend National Park. Texas: United States Department of the Interior, National Park Service.

Big Bend National Park (2004). General Management Plan/ Environmental Impact Statement. Texas: United States Department of the Interior, National Park Service, p. 324.

Bradley BA, Mustard JF (2005). Identifying land cover variability distinct from land cover change: cheatgrass in the Great Basin. Remote Sens. Environ. 94: 204-213.

Burnham KP, Anderson DR (1992). Data-based selection of an appropriate biological model: the key to modern data analysis. In: McCullough DR, Barrett RH (eds.), Wildlife 2001: populations. London, England: Elsevier Appl. Sci., pp. 16-30.

Cochran RA, Rives JL (1985). Soil survey of Big Bend National Park, part of Brewster County, Texas. Washington D.C.: U.S. Department of Agriculture, Soil Conservation Service, pp. 119.

Cox JR, Martin MH, Ibarra FA, Fourie JH, Rethman NFG, Wilcox DG (1988). The influence of climates and soils in the distribution of four African grasses. J. Range Manage. 41: 127-139.

Dewey SA, Price KP, Ramsey D (1991). Satellite remote sensing to predict potential distribution of dyers woad (*Isatis tinctoria*). Weed Technol. 5: 479-484.

Dick-Peddie WA (1993). New Mexico vegetation: past, present, and future. Albuquerque, New Mexico: University of New Mexico Press, pp 244.

Dudík M, Phillips SJ, Schapire RE (2007). Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. J. Mach. Learn. Res. 8: 1217-1260.

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

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-

Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006). Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129–151.

Elith J, Phillips SJ, Haste T, Dudík M, Chee YE, Yates CJ (2010). A statistical explanation of MaxEnt for ecologists. Diversity Distrib. 17: 43-57.

Everitt JH, DeLoach CJ (1990). Remote sensing of Chinese tamarisk (*Tamarix chinensis*) and associated vegetation. Weed Sci. 38: 273-278.

Everitt JH, Alaniz MA, Escobar DE (1992). Using remote sensing to distinguish common goldenweed (*Isocoma coronopifolia*) and Drummon goldenweed (*Isocoma drummondii*). Weed Sci. 40: 621-628.

Fieldings AH, Bell JF (1997). A review of methods for the assessment of prediction errors in conservation presence absence models. Environ. Conserv. 24: 38-49.

Franklin AB, Shenk TM, Anderson DR, Burnham KP (2001). Statistical model selection: an alternative to null hypothesis testing. In: Shenk TM, Franklin AB (eds), Modeling in natural resource management: development, interpretation, and application. Washington D.C.: Island Press, Pp. 75-90.

Franklin J (2009). Mapping species distributions: spatial inference and prediction. Cambridge, UK: Cambridge University Press, Pp 320.

Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT (2004). New developments in museum-based informatics and applications in biodiversity analysis. Trends Ecol. Evol. 19: 497–503.

Guisan A, Zimmermann NE (2000). Predictive habitat distribution models in ecology. Ecol. Modell. 135: 147-186.

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.

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006). Evaluating the ability of habitat suitability models to predict species presences. Ecol. Modell. 199: 142-152.

Holt EC (1985). Buffelgrass - A brief history. In: Runge ECA, Schuster JL (eds), Buffelgrass: adaptation, management and forage quality symposium. College Station, Texas, Texas Agricultural Experiment Station MP-1.575, pp. 1-5.

Hoshovsky M (1987). Element stewardship abstract for *Arundo donax, g*iant reed. Arlington, Virginia: The Nature Conservancy, Pp 10.

Ibarra-F FA, Cox JR, Hartin-R MH, Crowl TA, Call CA (1995). Predicting buffelgrass survival across a geographical and environmental gradient. J. Range Manage. 48: 53-59.

Joshi C, de Leeuw J, van Duren IC (2004). Remote sensing and GIS applications for mapping and spatial modeling of invasive species. Proceedings of ISPRS, Istanbul, Turkey, 35: 669-677.

Lass LW, Prather TS, Glenn NF, Weber KT, Mundt JT, Pettingill J (2005). A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (*Centaurea maculosa*) and baby breath (*Gypsophila paniculata*) with a hyperspectral sensor. Weed Sci. 53: 242-251.

Lawrence RL, Wood SD, Sheley RL (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). Remote Sens. Environ. 100: 356-362.

Mackenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA (2002). Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248-2255.

Manel S, Williams HC, Ormerdo SJ (2001). Evaluating presence-absence models in ecology: the need to account for prevalence. J. Appl. Ecol. 38: 921-931.

Marcus WA, Lefleiter CJ, Aspinall RJ, Boardman JW, Crabtree RL (2003). High spatial resolution hyperspectral mapping of in-stream habitats, depths, and woody debris in mountain streams. Geomorphol. 55: 363-380.

McClaran MP, Anable ME (1992). Spread of introduced Lehmann lovegrass along a grazing intensity gradient. J. Appl. Ecol. 29: 92-98.

McComb WC, McGrath MT, Spies TA, Vesely D (2002). Models for mapping potential habitat at landscape scales: an example using northern spotted owls. For. Sci. 48: 203-216.

National Park Service (1983). Big Bend: official National Park handbook. Division of Publications, National Park Service. Washington, D.C.: U.S. Department of the Interior, pp. 127.

National Parks Conservation Association (2003). State of the Parks: Big Bend National Park. Washington, D.C.: U.S. Department of Interior, National Park Service, Pp 24.

National Park Service (2005). Exotic plant management team; 2005 Annual Report. Washington, D.C.: U.S. Department of Interior, National Park Service, Pp 78.

Oliver F, Wotherspoon SJ (2006). Modelling habitat selection using presence-only data: case study of a colonial hollow nesting bird, the snow petrel. Ecol. Modell. 195: 187-204.

Parker-Williams AE, Hunt ER (2002). Estimation of leafy spurge cover from hyperspectral imagery using mixture tuned matched filtering. Remote Sens. Environ. 82: 446-456.

Pearce J, Ferrier S (2000). Evaluating the predictive performance of habitat models developed using logistic regression. Ecol. Modell. 133:225-245.

Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeogr. 34: 102-117.

Peterson AT, Papes M, Eaton M (2007). Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography 30: 550-560.

Peterson AT, Papes M, Soberon J (2008). Rethinking receiver operating characteristics analysis applications in ecological niche modeling. Ecol. Modell. 213: 63-72.

Phillips SJ (2008). Transferability, sample selection bias and background data in presence-only modeling: a response to Peterson et al. (2007). Ecography 31: 272-278.

Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy

modeling of species geographic distributions. Ecol. Modell. 190: 231-259.
Phillips SJ, Dudík M (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31: 161-175.
Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann AA, Leathwick J, Ferrier S (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol. Appl. 19: 181–197.
Phillips SJ, Dudík M, Schapire RE (2004). A maximum entropy approach to species distribution modeling. Proceedings of the Twenty-First International Conference on Machine Learning, 69: 655-662.
Plumb GA (1991). Vegetation classification of Big Bend National Park, Texas. Tex. J. Sci. 44: 375-386.
Price KP, Guo X, Stiles JM (2002). Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. Int. J. Remote Sens. 23: 5031-5042.
Rejmanek M, Richardson DM (1996). What attributes make some plant species more invasive? Ecology 77: 1655-1661.
Rejmanek M (2000). Invasive plants: approaches and predictions. Austral Ecol. 25: 497-506.
Rejmanek M, Pitcairn MJ (2002). When is eradication of exotic pest plants a realistic goal? In: Veitch CR, Clout MN (eds), Turning the tide: the eradication of invasive species. Gland, Switzerland and Cambridge, UK: IUCN SSC Invasive Species Specialist Group. IUCN, pp. 249-253.
Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JW, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001). The population biology of invasive species. Annu. Rev. Ecol. Syst. 32: 305-332.
Shafii B, Price WJ, Prather TS, Lass LW, Thill DC (2003). Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. Weed Sci. 51: 748-751.
Shrader-Frechette KS, McCoy ED. (1993). Method in ecology: strategies for conservation. London: Cambridge University Press.
Sievers AF (1930). The herb hunters guide. Misc. Publication. 77. Washington D.C: US Department of Agriculture.
Simon JE, Chadwick AF, Craker LE (1984). Herbs: an indexed bibliography. 1971-1980. The scientific literature on selected herbs, and aromatic and medicinal plants of the temperate zone. Hamden, Connecticut: Archon Books, p. 770.
Stohlgren TJ, Brinkley D, Chong GW, Kalkhan MA, Schell LD, Bull KA, Otsuki Y, Newman G, Bashkin M, Son Y (1999). Exotic plant species invade hot spots of native plant diversity. Ecol. Monogr. 69: 25-46.
Thenkabail PS, Enclona EA, Ashton MS, Legg C, De Dieu MJ (2004). Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. Remote Sens. Environ. 90: 23-43.
Tracy JL, DeLoach CJ (1998). Suitability of classical biological control for giant reed (*Arundo donax*) in the United States. In: Bell CE (ed), Arundo and saltcedar: the deadly duo, Proceedings of a workshop on combating the threat from arundo and saltcedar. Holtville, California, Cooperative Extension University of California, pp. 73-108.
VanDerWal J, Shoo LP, Graham C, Williams SE (2009). Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? Ecol. Modell. 220: 598-594.
Wauer RH (1996). Naturalist's Big Bend. College Station Texas: Texas A&M University Press, p. 149.
Weidong G, Swihart RK (2004). Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. Biol. Conserv. 116: 195-203.
White GC (2001). Statistical models: keys to understanding the natural world. In: Shenk TM, Franklin AB (eds), Modeling in natural resource management: development, interpretation, and application. Washington D.C.: Island Press, pp. 35-56.
Wisz MS, Hijmans RJ, Li K, Peterson AT, Graham CH, Guisan A (2008). NCWAS Predicting Species Distributions Working Group. Effects of sample size on the performance of species distribution models. Diversity Distrib. 14: 763-773.
Yost AC, Petersen SL, Gregg M, Miller R (2008). Predictive modeling and mapping sage grouse (*Centrocercus urophasianus*) nesting habitat using Maximum Entropy and a long-term dataset from Southern Oregon. Ecol. Inform. 3: 375-386.
Zaniewski AE, Lehmann A, McC. Overton J (2002). Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecol. Modell. 157: 261-280.
Zhang X, Friedl MA, Schaaf CB, Strahler AH, Hodges JCF, Gao F, Reed, BC Huete A (2003). Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84: 471-475.