Risk Analysis, Vol. 30, No. 2, 2010 DOI: 10.1111/j.1539-6924.2009.01343.x

Ensemble Habitat Mapping of Invasive Plant Species

Thomas J. Stohlgren,[1,*] Peter Ma,[2] Sunil Kumar,[3] Monique Rocca,[4] Jeffrey T. Morisette,[1] Catherine S. Jarnevich,[1] and Nate Benson[5]

Ensemble species distribution models combine the strengths of several species environmental matching models, while minimizing the weakness of any one model. Ensemble models may be particularly useful in risk analysis of recently arrived, harmful invasive species because species may not yet have spread to all suitable habitats, leaving species-environment relationships difficult to determine. We tested five individual models (logistic regression, boosted regression trees, random forest, multivariate adaptive regression splines (MARS), and maximum entropy model or Maxent) and ensemble modeling for selected nonnative plant species in Yellowstone and Grand Teton National Parks, Wyoming; Sequoia and Kings Canyon National Parks, California, and areas of interior Alaska. The models are based on field data provided by the park staffs, combined with topographic, climatic, and vegetation predictors derived from satellite data. For the four invasive plant species tested, ensemble models were the only models that ranked in the top three models for both field validation and test data. Ensemble models may be more robust than individual species-environment matching models for risk analysis.

KEY WORDS: Boosted regression trees; invasive speeies; Maxent; multivariate adaptive regression splines; random forest; species distribution modeling

1. INTRODUCTION

Harmful nonnative species continue to spread globally[(1)] and into natural areas including national parks and wildlife refuges, where natural resource policies demand containment to protect native species.[(2)] In response to increasing invasions, spatial modeling and species-environment matching models are becoming commonplace for natural resources managers, agencies, and nongovernment organizations who need accurate maps of species distributions and abundance for risk analysis.[(3)] Selection of modeling approaches may be based on specific study objectives; species autecology and syncology; species occurrence data (i.e., presence-absence versus presence-only or abundance); and the availability and resolution of predictor variables such as climatic, topographic, edaphic, or vegetation community information observed with remotely sensed data, which can be related to species distributions and abundance. However, the selection among different modeling approaches is sometimes based on cost or convenience (e.g., software availability, previous training). Problems with model interpretation arise when individual models behave differently with

[1]U.S. Geological Survey, Fort Collins Science Center, National Institute of Invasive Species Science, Fort Collins, CO, USA.
[2]NASA Goddard Space Flight Center/Sigma Spaee, Greenbelt, MD, USA.
[3]Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO, USA.
[4]Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Fort Collins, CO, USA.
[5]National Interagency Fire Center, National Park Service, Boise, ID, USA.
*Address correspondence to Thomas J. Stohlgren, U.S. Geological Survey, Fort Collins Science Center, National Institute of Invasive Species Science, Fort Collins, Colorado, USA; tel: 001-970-491-1980; fax: 001-970-491-1965; tom_stohlgren@usgs.gov.

0272-4332/10/0100-0224$22.00/1 © 2010 Society for Risk Analysis

the same set of field observations and predictor layers. Discrepancies among the results from different modeling approaches may be due to differences in model assumptions and algorithms, or different specific data requirements (e.g., some models use presence-only data[4] whereas others use presence-absence or abundance data[5]). These common situations led Araújo and New[5] to suggest using ensembles of models for species distribution modeling, rather than relying on model outputs from a single model to guide eradication efforts.

We compared five individual models: logistic regression, boosted regression trees (BRT), random forest, multivariate adaptive regression splines (MARS), and maximum entropy modeling (Maxent), and an ensemble model for selected nonnative plant species in Yellowstone and Grand Teton National Parks, Wyoming; Sequoia and Kings Canyon National Parks, California, and areas of interior Alaska. Ensemble modeling (sometimes called consensus modeling[6]), in its simplest terms, involves combining model outputs from different models (in this case, binary maps of suitable and unsuitable habitat) into a single map. In this way, areas of high model agreement can be differentiated from areas of low model agreement. From a scientific standpoint, it is instructive to learn whether individual models or ensemble models perform better for a broad array of species. From a practical standpoint, resource managers need to know which model or groups of models provide the most accurate predictions in space and time. This is the essence of risk analysis for biotic invaders.[7,8]

The use of ensemble models for mapping harmful invasive species is not yet widespread. Crossman and Bass[9] used ensemble modeling to map *Olea europaea* L. (European olive) in southern Australia, reporting that the ensemble model outperformed individual models. Marmion *et al.*[6] created ensemble models for 28 threatened plants in one study area in northeastern Finland. Roura-Pascual *et al.*[10] used consensual predictions for identifying areas of uncertainty regarding the Argentine ants' invasive potential in the Iberian Peninsula. However, we found no studies that compared models of multiple species in multiple biomes to support the notion that ensemble models consistently outperform individual models. Our objectives were to: (1) test five individual modeling approaches and an ensemble model for four invasive plant species (one species in each of four study areas); and (2) provide resource managers of natural areas with recommendations on the benefits and drawbacks of an ensemble approach for risk assessments of mapping harmful invasive species.

2. METHODS

2.1. Study Sites and Field Data

This study was conducted at four different sites across the United States. Two of these sites, Yellowstone (44°07′44″N to 45°06′17″N; 109°50′28″W to 111°09′12″W; area: 1,112,652 ha) and Grand Teton (43°32′15″N to 44°08′48″N; 110°25′28″W to 110°56′50″W; area: 281,812 ha) National Parks (N.P.) are located in northwest Wyoming, USA. Data sets on *Linaria dalmatica* (Dalmation toadflax) in Yellowstone N.P. and *Carduus nutans* (musk thistle) in Grand Teton N.P. were collected for more than a 20-year period. Sequoia and Kings Canyon National Parks (SEKI; 36°15′49″N to 37°14′21″N; 118°07′35″W to 118°59′51″W; area: 829,463 ha) are located in the southern Sierra Nevada of California, USA. Locations of *Bromus tectorum* (cheatgrass) have been collected in a variety of surveys and monitoring projects in the parks. The interior Alaska area (62°39′26″N to 67°08′41″N; 143°30′22″W to 153°58′0.26″W; area: 18,603,882 ha) included lands around Fairbanks and east of Denali National Park and had over 10 years of data collection from various surveys recording *Melilotus officianalis* (white sweet clover) and other invasive species. All species occurrence data were collected using global positioning systems by park personnel.

The four study sites selected incorporate large geographic ranges, which cross ecological zones, political boundaries, and are current working extents of existing noxious plant coordinating committees that include local, state, and national land managers. While the extents of each study site are represented by political boundaries in the graphics, models used to generate these outputs considered the entire extent of the study site, which corresponds to the coordinates listed above.

2.2. Environmental Data

Environmental data layers that present biophysical characteristics play an important role in our ability to shed light on underlying species-environment processes and relationships.[11] For these reasons, it is common practice that climate layers derived from historic records are temporally cultivated and coupled with topographic and remotely sensed

parameters to derive species distribution models. However, one of the main challenges in model construct is the fundamental representation of the physical environment with the appropriate scale of digital proxies, that is, geospatial data layers. As Guisan and Thuiller[3] describe, the resolution and extent of a study site are important in understanding the limitation of results. Furthermore, the inconsistencies in resolution across sources of biophysical parameters are common and in a perfect scenario all environmental layers would be generated at the same native spatial and temporal resolution. Nevertheless, accounting for coherence limitations we chose to resample all environmental predictor layers to a 250 m^2 spatial resolution. This scale was chosen because it provided the highest native spatial resolution of the remotely sensed MODIS phenology data product, while additionally supplying an appropriate resolution map for park service personnel to execute general treatments and accuracy assessment analysis.

In our study, we examined similar climate layers in conjunction with a recently developed remotely sensed, seasonal interannual vegetation measure called the MODIS phenology product. NASA's MODIS land surface phenology product defines seasonal patterns of variation in vegetated land surfaces from satellite observations.[12–14] The product derives 14 temporally smoothed and spatially gap filled annual phenological parameters from common vegetation indices (NDVI/EVI) at a spatial resolution of 250 m^2, a close proximity in spatial resolution to the 1 km^2 climatic layers.

In addition to the remotely sensed MODIS vegetation phenological parameter layers, a spatial data set consisting of over 70 candidate habitat predictor layers was assembled for each study site from a variety of sources. Environmental variables included 19 bioclimatic layers[15] derived from monthly temperature and precipitation data (Daymet climate data set; www.daymet.org/; 1-km^2 spatial resolution; 1980–1997) and multiple types of land surface data sets such as elevation, percent tree cover, and incoming solar radiation. The Appendix further describes the entire list of habitat predictor layers considered in our analysis.

All predictors were compared in a pairwise fashion to evaluate their correlation for each species at each site. For any highly correlated ($r > 0.8$) pairs, we removed one of the variables from consideration as habitat predictors to reduce problems associated with multicollinearity. Our choice of which variable to remove was somewhat subjective. We kept the variables we felt were more directly interpretable by the modeling team and the National Park Service personnel. For example, if the two phenology metrics: "length of growing season" and "rate of greenup"[12] were highly correlated we would keep "length of growing season" assuming that metric was fairly well understood. Since the variables were highly correlated, they were essentially providing the same information. So, for the purpose of this study, our main objective was to reduce numerical problems with multicollinearity and our somewhat subjective selection of which variable to remove does not impact the model intercomparison as the same variables were available for the different model techniques.

2.3. Statistical Modeling Methods

Based on the findings of recent model comparison studies[16–20] and the applications in several ecological studies, we selected five of the top performing species distribution models: boosted regression trees or BRT,[20–24] logistic regression,[25,26] Maxent,[27–29] multivariate adaptive regression splines or MARS,[16,30,31] and random forest.[32,33] BRT and random forest are ensemble techniques within themselves, producing a final model that is a combination of many different trees. MARS is a nonparametric technique that models nonlinearities and interactions by combining regression and recursive partitioning methods.[16] Maxent is a presence-only technique while the others require both presence and absence data.[27]

For each species in each location, we randomly divided species location data into two groups, with 50% of our available presence-absence data (i.e., split sample approach)[34] for use in model testing (the "test" data) and the other subset (50%) reserved for model fitting or training. Thus, both the test and training subsets had a prevalence of 0.5 (i.e., the proportion of presence points to absence points in the data set).

In the predictive phase of modeling, each of the five modeling methods produced an output map of continuous values between zero and one corresponding to the probability that a pixel is suitable habitat for the species. To convert these continuous maps into binary (habitat versus nonhabitat) maps, we needed to select a threshold such that probability values above that threshold were designated as habitat, while values at or below that threshold were designated as nonhabitat. Using the "test" data set, we selected a threshold at which the specificity was

equal to the sensitivity.[35,36] That is, we selected a threshold at which there were as many false positives as there were false negatives. Once a threshold was selected, the model output for each individual model was converted to a binary map.

Ensemble maps were produced by combining the binary habitat maps from the five individual models. The resulting ensemble map displayed integer values from zero to five, with the value of each pixel corresponding to the number of individual models that predict that pixel to be habitat (a *frequency histogram ensemble approach*).[5] This output can be considered a "vote" from the five modeling techniques. A score of zero indicates that none of the modeling techniques assigned that area as suitable habitat, while a value of five indicates that all modeling techniques assigned that area as suitable habitat.

We compared model accuracy of individual and ensemble models using two validation data sets. First, we used the reserved test data set to validate the model. Second, our National Park Service partners conducted new field surveys in the summer of 2008 (the "field validation" data), with sampling locations selected by stratifying across the six (zero through five) classes in the ensemble map. Thirty points were sampled in each stratum for each study area.

We used the area under the receiver-operating characteristic (ROC) curve (AUC)[37] to assess model accuracy with each validation data set. AUC, a measure of model fit, has been widely used in several model comparison studies.[16–19,38] A model performs well when the AUC is large. Usually, AUC values of >0.9 indicate high accuracy, values of 0.7–0.9 indicate good accuracy, and values 0.5 (random) to 0.7 indicate low accuracy.[39] We minimized the potential problems associated with the reliance on AUC values (see Lobo *et al.*[40]) by: (1) predicting the habitat suitability for only one species at a time over the same extent for all models; (2) selecting only models that performed well in previous model comparison studies;[18,19,41] (3) eliminating highly cross-correlated predictor variables prior to testing; and (4) and restricting the spatial distributions of all the models to similar environments (e.g., the Greater Yellowstone Ecosystem).

3. RESULTS

In this section, we cover the performance of individual models, the performance of ensemble models, and finally, the effectiveness of the ensemble models to guide field validation efforts. The performance of individual models varied by park and by species, and varied for field validation and test data sets (Table I). For example, for the field validation data for *Carduus nutans* in Grand Tetons N.P., AUC scores were highest for the MARS model, followed by Maxent and the ensemble model. For the test data for the same species and park, AUC scores were highest for the random forest model, followed by boosted regression trees, and the ensemble model. The random forest model tended to be the best performing individual model for most species, but it performed poorly for *Carduus nutans* in Grand Tetons N.P. Ensemble models were the only models to rank among the top three models for field validation and test data in all study areas (Table I).

Habitat suitability maps varied by species and by models used. For example, all the models of *Linaria dalmatica* in the Yellowstone area predicted suitable habitat in the north-central part of the park (Fig. 1), overlapping the majority of reported presence locations. The random forest model accentuated habitat suitability in the west-central part of the park, relative to the other models. The boosted regression trees model accentuated the road corridors. The MARS model accentuated habitat suitability in the park's interior (Fig. 1).

The ensemble model for *Linaria dalmatica* in the Yellowstone N.P. and adjacent areas (Fig. 2) showed high agreement among models in the north-central, west-central, and along many of the major roads. The greatest habitat suitability corresponds to development zones in the park, including the north entrance to Yellowstone (and Mammoth area), the West Yellowstone entrance, and the road corridor at the east entrance to the park. An additional "potential hotspot" for *Linaria dalmatica* can be seen about 30 km east of the west entrance, near a large picnic area along the Gibbon River, where no presence points are reported.

Field validation data confirmed the modeling results. In Yellowstone National Park, 2008 validation data collection focused on the northern half of Yellowstone Park, enabling staff to survey over 300 km of off-trail backcountry routes that had not been previously surveyed, covering about 240 ha. In Grand Teton National Park, the 2008 validation data set consisted of 65 stratified random plots. In Sequoia and Kings Canyon National Parks, the 2008 validation data set included sampling between June and September, including an additional 200 plots (0.06-ha circular plots) while traversing over 360 km of trails.

Table I. The Area Under the Curve (AUC) Values for the Different Models and Parks, Calculated Using the ROC-AUC Program[53]

Model	*Carduus Nutans* in Grand Tetons	*Bromus Tectorum* in Sequoia and Kings Canyon	*Linaria Dalmatica* in Yellowstone	*Melilotus Officianalis* in Alaska
Field Validation Data				
Ensemble	0.714	0.474	0.940	0.657
Logistic reg.	0.615	0.406	0.746	0.613
Random forest	0.615	0.453	0.902	0.654
Boosted regression trees	0.516	0.425	0.830	0.646
Maxent	0.802	0.442	0.822	0.444
MARS	0.859	0.345	0.587	0.660
No. observations (pres. vs. abs.)	2 to 96	8 to 191	3 to 93	113 to 212
Test Data				
Ensemble	0.894	0.969	0.981	0.938
Logistic reg.	0.796	0.896	0.932	0.906
Random forest	0.975	1.000	1.000	1.000
Boosted regression trees	0.931	0.969	0.996	1.000
Maxent	0.742	0.499	0.938	0.313
MARS	0.821	0.948	0.755	0.875
No. observations (pres. vs. abs.)	526 to 526	55 to 55	228 to 229	8 to 8

Field teams recorded coordinates of not-yet-reported locations ($n = 8$) of *Bromus tectorum* (cheatgrass), suggesting that field validation based on model results was effective.

4. DISCUSSION

4.1. The Value of Ensemble Modeling

Individual models varied in their performance (Table I). This is true for all model comparison papers published to date (e.g., Elith *et al.*[16] Kumar *et al.*[19]). Geographic differences (Table I) or the species-specific traits of habitat generalists or habitat specialists[18] may take precedence in local model selection. However, the ensemble map approach adds substantial robustness and consistency of performance across the four parks and species considered (Table I; Fig. 2). Testing multiple models and combining their predictions to produce ensemble models may help avoid the pitfall of trusting any one convenient model. For example, the MARS model performed reasonably well for *Carduus nutans* in Grand Tetons N.P. and for *Melilotus officianalis* in Alaska, but then it performed poorly for *Bromus tectorum* in Sequoia and Kings Canyon N.P. and *Linaria dalmatica* in Yellowstone N.P.

Many individual model comparison studies recommend the best-performing individual model without testing ensemble models. We showed, as did Marmion *et al.*,[6] that ensemble models do not necessarily always improve the predictive accuracy of the single models, primarily because ensemble model accuracy is always dependent on the accuracy of individual models.[5] However, our results indicate that ensemble modeling should not be ignored. Ensemble techniques may be particularly helpful in an automated habitat mapping system where robustness can be the most important model selection criteria. Ensemble models may also be particularly useful in modeling recently arrived, harmful invasive species because species may not yet have spread to all suitable habitats, leaving species-environment relationships difficult to determine.

4.2. Caveats

There are several caveats for these types of investigations related to data completeness and model selection and performance. In three of the four parks, there were very few presence points relative to the number of absence points in the field validation data sets. While these surveys were guided by the ensemble models, the lack of presence locations may not be because the locations are unsuitable but rather because the species have not yet spread to those locations. The low field validation AUC values may be suspect due to the high bias in prevalence toward negative locations and lack of knowledge about "true" absence locations, and the high test data AUC values (e.g., 1.0 for three of the four random

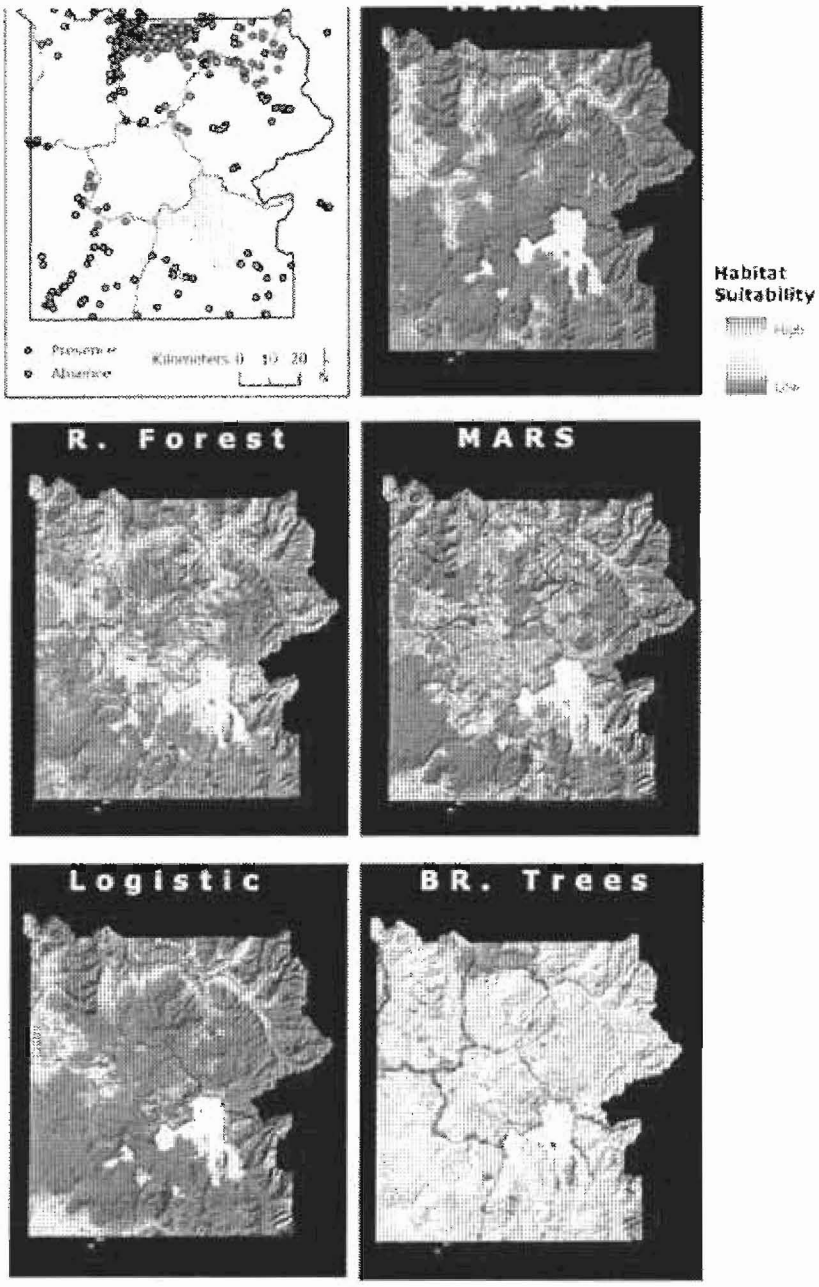

Fig. 1. Point distribution of *Linaria dalmatica* (Dalmation toadflax) in Yellowstone National Park, Wyoming, and modeled distribution from five common species distribution models: Maxent, random forest (R. Forest), multivariate adaptive regression splines (MARS), logistic regression (Logistic), and boosted regression trees (BR. Trees).

forest models) is also suspect. Spatial autocorrelation between training and test data can lead to inflated AUC values.[42] Spatial models are affected by sample size and the contagion of samples relative to environmental gradients across the study extent.[29] In addition, all the species tested here occur outside our area of interest, and restricting the models to a portion of the overall species ranges may bias the models.[43] It was not our objective to produce larger regional-scale models. We realize that models performed for different "areas of interest" will produce different results, not just due to the environmental layers used for prediction, or the additional occurrence points at the larger extents. Our objective was to capture the local extent of selected plant species and genotypes specific to the particular ecosystems.

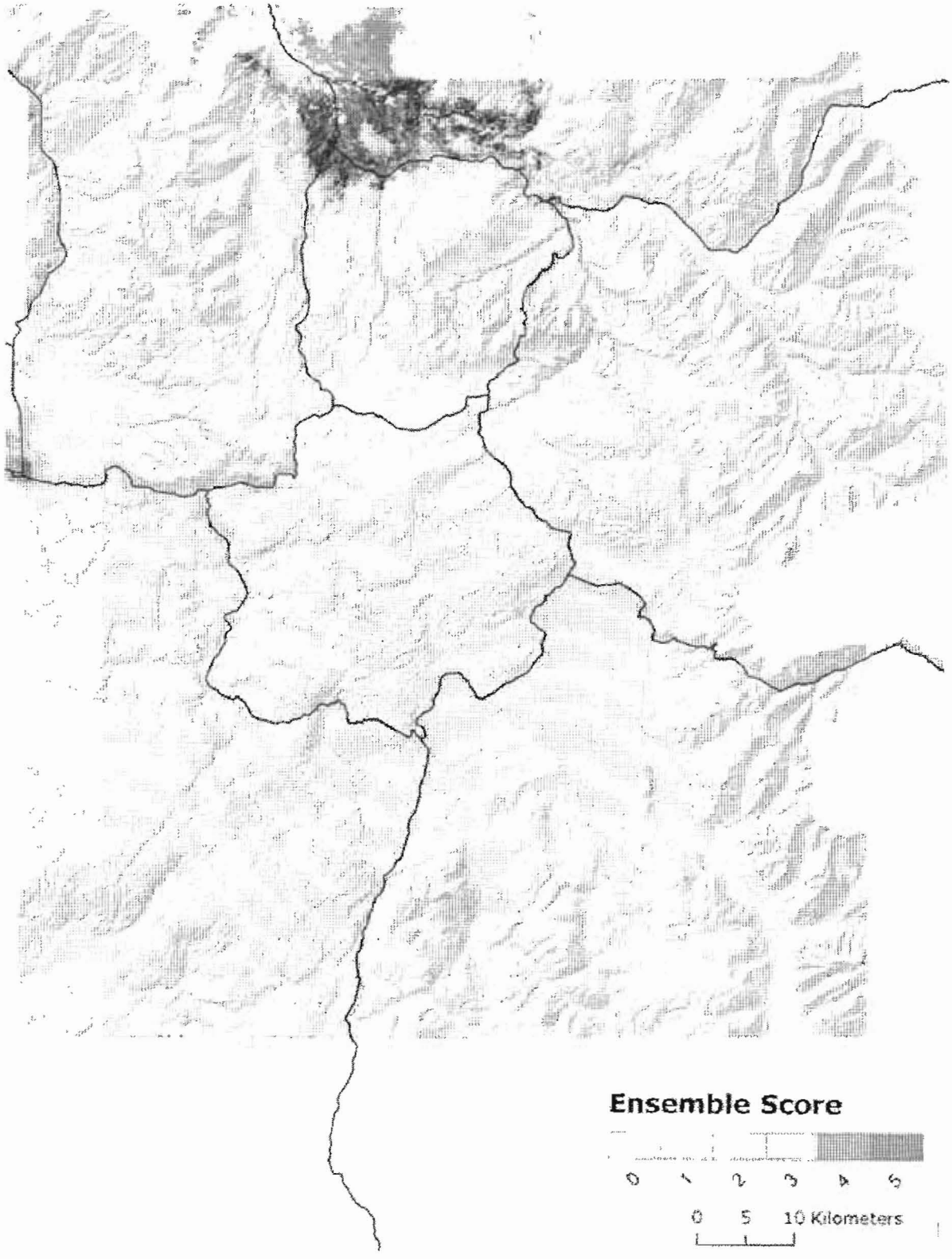

Fig. 2. Ensemble model of *Linaria dalmatica* (Dalmation toadflax) in Yellowstone National Park, Wyoming, and adjacent lands.

Regional models often combine different genotypes in a wide variety of habitats unavailable in our area of interest.

Model selection and performance affects results. Although we tested models that were highly successful in previous model comparison studies,[16,44] model performance in those studies may have been affected by data completeness, the extent to which true absence location could be verified, the extent of the study area, and the strength of species-environment relationships. Furthermore, we did not calculate the confidence intervals for our predictions of habitat.[45] Uncertainties exist in all statistical models;[27,29] however, some models (e.g., GARP) have consistently underperformed in several model comparison studies,[19,41] and inclusion of weak models in a model ensemble would likely weaken overall results. It was not our intent to compare individual projections with each other following Thuiller,[46] although we recognize this may be of considerable value when deciding which models to retain or eliminate in the ensemble process.

Pearson *et al.*[35] argued that suitability of habitats cannot be appropriately distinguished from bioclimatic layers alone, and that remotely sensed land cover data significantly improved exclusively climate-driven predictions. In their work, the improvement of model results is based on a hierarchical approach that downsizes and compartmentalizes climate layers to help account for differences in spatial resolution with land cover data sources. However, their research also suggests that the nonhierarchical approach performed similarly well, with small statistical differences in the results. In related studies, improved results from contributions of remotely sensed land cover data have been mixed.[43] On a coarse scale, Thuiller *et al.*[43] explain that the inclusion of land cover data improved explanatory power, while having no effect on the predictive power of the models. Their discussion also suggested that at a smaller spatial extent and finer resolution data may produce superior results. Zimmerman *et al.*[47] explain that when 30 m remotely sensed Landsat predictors depicting vegetation structure, biomass, and productivity were included with 1 km climate variables at a finer spatial scale (90 m), climate data continued to outperform the remotely sensed predictors. Zimmerman *et al.*[47] also emphasized the need for continuous gradient remotely sensed predictors as apposed to filtered classifications and discussed the growing research interest and importance in the subtle characteristics and differences of phenological events. Like Zimmerman *et al.*,[47] recent research has focused on the use of vegetation indices with climatic layers in species distribution models to help exploit the spectral response from vegetation and link remote sensing to underlying ecological relationships of species.

4.3. Alternative Approaches to Ensemble Modeling

The ensemble approach used here is a simple way to compare various species distribution models. The way in which model results are combined also affects results. Marmion *et al.*[6] reported that weighted average techniques and mean values from repeated individual model runs consistently outperformed individual models for 28 threatened species. Several other approaches to ensemble modeling (or consensus methods) are reviewed by Marmion *et al.*[6] It is possible that there would have been some differences in our results had we used weighted averages rather than simple summation of binary results. However, we believe the general trend in the numbers would have been similar and the ensemble technique we used here is the most straightforward to explain and interpret.

4.4. Iterative Mapping and Modeling of Invasive Species

Environmental conditions will change with respect to climate and land-use change, and many other factors. Thus, mapping and modeling of invasive species must be considered an iterative process.[8,48] Invasive species will spread naturally by wind, animals, and other means of dispersal. The leading edge of a population distribution may behave differently than the initial infestations.[49] Containment efforts affect current and future populations. Resource managers may be in a constant cycle of collecting new field information, new remote sensing information, and performing new model runs to track invasive species and to alter strategies for containment or gathering new information. Ensemble models provide a simple technique to provide robust estimates of suitable habitat for a given invasive species at a given time.

4.5. Relevance to Risk Assessment

Ensemble models and maps are an essential tool in risk analysis of invasive species.[7,50,51] However, new field data to validate and improve the models

are also essential. Especially with a rapidly invading species, the associated risks change equally quickly. Species adapt, climates and environments change, and new techniques are developed. Adding economic costs of effects and control strategies are equally important.[52]

ACKNOWLEDGMENTS

We wish to thank and fully credit the National Park Service collaborators from Yellowstone: Craig McClure, Ann Rodman, and Eric Miller. Sequoia and Kings Canyon: Athena Demetry, Tony Caprio, and Pat Lineback. Alaska: Jeff Heyes, Jennifer Allen, and Brian Sorbel. Grand Teton: Kelly McCloskey and Diane Abendroth. We would also like to credit the professionals, volunteers, plant ecologists, fire ecologists, biological science technicians, GIS specialists, and the many seasonal field technicians for their contributions of expertise, time, effort, and data. More specifically, thanks to Christi Sorrell, Mike Stefancic, Roy Renkin, Gil Crain, Steve Miller, Bob Crabtree, Josh Harmsen, Dan Tyers, Susan Klein, Carolyn Richard, Carrie Guiles, Gary Muerle, Richard Thiel, Erik Frenzel, Karen Webster, Thomas McGinnis, Sylvia Haultain, Karen Folger, Jody Lyle, Thomas Liebscher, Trisha Wurtz, Katie Villano, Matt Macander, Blaine Spellman, Michele Young, Eric "Doc" Janssen, Klara Varga, and the Greater Yellowstone Coordinating Committee. Logistical support was provided by the USGS Fort Collins Science Center and the Natural Resource Ecology Laboratory at Colorado State University, Fort Collins, Colorado. Funding was provided by NASA Decision Support through Earth Science Results, Grant NN-H-04-Z-YO-010 (Woody Turner, program manager). Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. government.

APPENDIX

Environmental predictor variables considered in the five predictive models including their native spatial resolution and the sources of data.

Predictor Variable	Native Spatial Resolution	Data Source
Climatic variables		
Annual mean temperature (bio1)	1 km	Daymet[1]
Mean diurnal range in temperature (bio2)	1 km	Daymet[1]
Isothermality (bio3)	1 km	Daymet[1]
Temperature seasonality (SD × 100) (bio4)	1 km	Daymet[1]
Maximum temperature of warmest month (bio5)	1 km	Daymet[1]
Minimum temperature of coldest month (bio6)	1 km	Daymet[1]
Temperature annual range (bio7)	1 km	Daymet[1]
Mean temperature of wettest quarter (bio8)	1 km	Daymet[1]
Mean temperature of driest quarter (bio9)	1 km	Daymet[1]
Mean temperature of warmest quarter (bio10)	1 km	Daymet[1]
Mean temperature of coldest quarter (bio11)	1 km	Daymet[1]
Mean annual precipitation (bio12)	1 km	Daymet[1]
Precipitation of wettest month (bio13)	1 km	Daymet[1]
Precipitation of driest month (bio14)	1 km	Daymet[1]
Precipitation seasonality (CV) (bio15)	1 km	Daymet[1]
Precipitation of wettest quarter (bio16)	1 km	Daymet[1]
Precipitation of driest quarter (bio17)	1 km	Daymet[1]
Precipitation of warmest quarter (bio18)	1 km	Daymet[1]
Precipitation of coldest quarter (bio19)	1 km	Daymet[1]
Remotely sensed variables		
Enhanced Vegetation Index (EVI) green up	250 m	MODIS[2]
EVI brown down	250 m	MODIS[2]
EVI season length	250 m	MODIS[2]
EVI base level	250 m	MODIS[2]
EVI peak date	250 m	MODIS[2]
EVI maximum peak date	250 m	MODIS[2]
EVI amplitude	250 m	MODIS[2]
EVI green up rate	250 m	MODIS[2]

(*Continued*)

Predictor Variable	Native Spatial Resolution	Data Source
EVI brown down rate	250 m	MODIS[2]
EVI large integrals—season start to end	250 m	MODIS[2]
EVI small integrals—season to base level	250 m	MODIS[2]
EVI annual maximum value	250 m	MODIS[2]
EVI annual minimum value	250 m	MODIS[2]
EVI annual mean value	250 m	MODIS[2]
EVI quality control	250 m	MODIS[2]
EVI land cover	250 m	MODIS[2]
Normalized Differenee Vegetation Index (NDVI) green up	250 m	MODIS[2]
NDVI brown down	250 m	MODIS[2]
NDVI season length	250 m	MODIS[2]
NDVI base level	250 m	MODIS[2]
NDVI peak date	250 m	MODIS[2]
NDVI maximum peak date	250 m	MODIS[2]
NDVI amplitude	250 m	MODIS[2]
NDVI green up rate	250 m	MODIS[2]
NDVI brown down rate	250 m	MODIS[2]
NDVI large integrals—season start to end	250 m	MODIS[2]
NDVI small integrals—season to base level	250 m	MODIS[2]
NDVI annual maximum value	250 m	MODIS[2]
NDVI annual minimum value	250 m	MODIS[2]
NDVI annual mean value	250 m	MODIS[2]
NDVI quality control	250 m	MODIS[2]
NDVI land cover	250 m	MODIS[2]
Percent impervious surface	30 m	MLRC[3]
Percent tree cover	30 m	MLRC[3]
Vegetation Continuous Field (VCF): percent tree cover	500 m	GLCF[4]
VCF: percent bare ground cover	500 m	GLCF[4]
VCF: percent herbaceous cover	500 m	GLCF[4]
Topographic and other environmental variables		
Elevation	30 m	SRTM[5]
Elevation standard deviation	30 m	SRTM[5]
Slope	30 m	SRTM[5]
Relative slope	30 m	SRTM[5]
Incident solar radiation	30 m	SRTM[5]
Fire frequency	250 m	Internal[6]
Most recent burn	250 m	Internal[6]
Year since last burn	250 m	Internal[6]
Forest damage	250 m	NPS[7]

Data sources:
[1]Daymet: http://www.daymet.org/.
[2]MODerate resolution Imaging Spectroradiometer (MODIS) Vegetation Indices: http://edcdaac.usgs.gov/modis/dataproducts.asp#mod13.
[3]Multiresolution Land Characteristics Consortium (MLRC): http://www.mrlc.gov/.
[4]Global Land Cover Facility (GLCF): http://www.landcover.org/index.shtml.
[5]Surface Radar Topography Mission (SRTM): http://srtm.csi.cgiar.org/.
[6]Internal: NASA GSFC/ USGS Science Center.
[7]NPS: National Park Service.

REFERENCES

1. Thuiller W, Richardson DM, Pysek P, Midgley GF, Hughes, GO, Rouget M. Niche-based modelliug as a tool for predicting the risk of alien plant invasions at a global scale. Global Change Biology, 2005; 11(12):2234–2250.
2. Allen JA, Brown CS, Stohlgren TJ. Non-native plant invasions of the United States National Parks. Biological Invasions, 2009; 11(10):2195–2207.
3. Guisan A, Thuiller W. Predicting species distribution: Offering more than simple habitat models. Ecology Letters, 2005; 8(9):993–1009.
4. Hirzel AH, Hausser J, Chessel D, Perrin N. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? Ecology, 2002; 83(7):2027–2036.
5. Araújo MB, New M. Ensemble forecasting of species distributions. Trends in Ecology & Evolution, 2007; 22(1): 42–47.

6. Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W. Evaluation of consensus methods in predictive species distribution modeling. Diversity and Distribution, 2009; 15(1):59–69.
7. Landis WG (ed). Regional Scale Ecological Risk Assessment Using the Relative Risk Model. Boca Raton, FL: CRC Press, 2005.
8. Stohlgren TJ, Schnase JL. Risk analysis for biological hazards: What we need to know about invasive species. Risk Analysis, 2006; 26(1):163–173.
9. Crossman ND, Bass DA. Application of common predictive habitat techniques for post-border weed risk management. Diversity and Distribution, 2008; 14(2):213–224.
10. Roura-Pascual N, Brotons L, Peterson AT, Thuiller W. Consensual predictions of potential distributional areas for invasive species: A case study of Argentine ants in the Iberian Peninsula. Biological Invasions, 2009; 11(4):1017–1031.
11. Turner W. Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M. Remote sensing for biodiversity science and conservation. Trends in Ecology & Evolution, 2003; 18(6):306–314.
12. Jönsson P, Eklundh L. TIMESAT—A program for analyzing time-series of satellite sensor data. Computers Geoscie, 2004; 30(8):833–845.
13. Tan B, Morisette JT, Wolfe RE, Gao F, Ederer GA, Nightingale J *et al.* Vegetation phenology metrics derived from temporally smoothed and gap-filled MODIS data. IGARSS, 2008; III:593–595.
14. Morisette JT, Richardson AD, Knapp AK, Fisher JI, Graham EA, Abatzoglou J *et al.* Tracking the rhythm of the seasons in the face of global change: Phenological research in the 21st century. Frontiers in Ecology and Environment, 2009; 7(5):253–260.
15. Nix HA. A biogeographic analysis of Australian elapid snakes. Chapter 7 in Longmore R (ed). Australian Flora and Fauna Series 8. Canberra, Australia: Australian Government Publishing Service, 1986.
16. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A *et al.* Novel methods improve prediction of species' distributions from occurrence data. Ecography, 2006; 29(2):129–151.
17. Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT. What matters for predicting the occurrences of trees: Techniques, data or species' characteristics? Ecological Monographs, 2007; 77(4):615–630.
18. Evangelista PH, Kumar S, Stohlgren TJ, Jarnevich CS, Crall AW, Norman JB *et al.* Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distribution, 2008; 14(5):808–817.
19. Kumar S, Spaulding SA, Stohlgren TJ, Herman KA, Schmidt TS, Bahls LL. Potential habitat distribution for the freshwater diatom Didymosphenia geminata in the continental US. Frontiers in Ecology and Environment, 2009; 7(8):415–420.
20. Friedman JH, Hastie T, Tibshirani R. Additive logistic regression: A statistical view of boosting. Annals of Statisrics, 2000; 28(2):337–407.
21. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann. Stat, 2001; 29(2):1189–1232.
22. Friedman JH. Stochastic gradient boosting. Computational Statistics & Data Analysis, 2002; 38(4):367–378.
23. Schapire R. The boosting approach to machine learning—An overview. In MSRI Workshop on Nonlinear Estimation and Classification, Berkeley, CA: 2003.
24. Elith J. Leathwick, JR Hastie T. A working guide to boosted regression trees. Journal of Animal Ecology, 2008; 77(4):802–813.
25. McCullagh P, Nelder JA. Generalized Linear Models. New York: Chapman and Hall, 1989.
26. Pearce J, Ferrier S. Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling, 2000; 133(3):225–245.
27. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecological Modelling, 2006; 190(3–4):231–259.
28. Phillips SJ, Dudik M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography, 2008; 31(2):161–175.
29. Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A, Leathwick J *et al.* Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecological Applications, 2009; 19(1):181–197.
30. Friedman JH. Multivariate adaptive regression splines (with discussion). Annals of Statistics, 1991; 19(1):1–141.
31. Hastie T, Tibshirani RJ. Discriminant analysis by Gaussian mixtures. Journal of Royal Statistical Society. Series B. 1996; 58(1):155–176.
32. Breiman L. Random forests. Machine Learning, 2001; 45(1):5–32.
33. Prasad AM, Iverson LR, Liaw A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems, 2006; 9(2):181–199.
34. Guisan A, Hofer U. Predicting reptile distributions at the mesoscale: Relation to climate and topography. Journal of Biogeography, 2003; 30(8)1233–43.
35. Pearson RG, Dawson TP, Liu C. Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. Ecography, 2004; 27(3):285–298.
36. Liu CR, Berry PM, Dawson TP, Pearson RG. Selecting thresholds of occurrence in the prediction of species distributions. Ecography, 2005; 28(3):85–393.
37. Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environment Conservation, 1997; 24(1):38–49.
38. Graham CH, Elith J, Hijmans RJ, Guisan A, Peterson AT, Loiselle BA *et al.* The influence of spatial errors in species occurrence data used in distribution models. Journal of Applied Ecology, 2008; 45(1):239–247.
39. Swets JA. Measuring the accuracy of diagnostic systems. Science, 1988. 240(3)1285–1293.
40. Lobo JM, Jiménez-Valverdel A, Real R. AUC: A misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography. 2008; 17(2):145–151.
41. Li M-Y, Ju Y-W, Kumar S, Stohlgren TJ. Modeling potential habitats for alien species Dreissena polymorpha (Zebra mussel) in the continental USA. Acta Ecologica Sinica, 2008; 28(9):4253–4258.
42. Segurado P, Araujo MB, Kunin WE. Consequences of spatial autocorrelation for niche-based models. Journal of Applied Ecology, 2006; 43(2):433–444.
43. Thuiller W, Brotons L, Araujo MB, Lavorel S. Effects of restricting environmental range of data to project current and future species distributions. Ecography, 2004; 27(2): 165–172.
44. Peters J, De Baets B, Verhoest NEC, Samson R, Degroeve S, De Becker P *et al.* Random forests as a tool for ecohydrological distribution modelling. Ecological Modelling, 2007; 207(2–4):304–318.
45. Schisterman EF, Faraggi D, Reiser B, Trevisan, M. Statistical inference for the area under the receiver operating characteristic curve in the presence of random measurement error. American Journal of Epidemiology, 2001; 154(2):174–179.
46. Thuiller W. Patterns and uncertainties of species' range shifts under climate change. Global Change Biology, 2004; 10(12):2020–2027.
47. Zimmermann NE, Edwards TC, Moisen GG, Frescino TS, Blackard JA. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree

species in Utah. Journal of Applied Ecology, 2007; 44(5):1057–1067.
48. Jarnevich CS, Stohlgren TJ. Near term climate projections for invasive species distributions. Biological Invasions, 2009; 11(6):1373–1379.
49. McInerny GJ, Turner JRG, Wong HY, Travis JMJ, Benton TG. How range shifts induced by climate change affect neutral evolution. Proceedings of the Royal Society-Biological Sciences, 2008; 276(1661):1527–1534.
50. Beaumont LJ, Gallagher RV, Thuiller W, Downey PO, Leishman MR, Hughes L. Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. Diversity and Distributions, 2009; 15(3):409–420.
51. Bradley BA, Oppenheimer M, Wilcove DS. Climate change and plant invasion: Restoration opportunities ahead? Global Change Biology, 2009; 15(6):1511–1521.
52. Keller RP, Lodge DM, Lewis MA, Shogren JF (eds). Bioeconomics of Invasive Species: Integrating Ecology, Economics, Policy, and Management. New York: Oxford, 2009.
53. Schroeder B. ROC/AUC-calculation, 2006. http://brandenburg.geoecology.uni-potsdam.de/users/schroeder/download.html, Accessed on July 1, 2009.