Diversity and Distributions, (Diversity Distrib.) (2008) **14**, 808–817

BIODIVERSITY RESEARCH

Modelling invasion for a habitat generalist and a specialist plant species

Paul H. Evangelista[1]*, Sunil Kumar[1], Thomas J. Stohlgren[2], Catherine S. Jarnevich[2], Alycia W. Crall[1], John B. Norman III[1] and David T. Barnett[1]

[1]*Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523-1499, USA,* [2]*US Geological Survey, Fort Collins Science Center, Fort Collins, Colorado 80526-8118, USA*

*Correspondence: Paul H. Evangelista, Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523-1499, USA. E-mail: paulevan@nrel.colostate.edu

ABSTRACT

Predicting suitable habitat and the potential distribution of invasive species is a high priority for resource managers and systems ecologists. Most models are designed to identify habitat characteristics that define the ecological niche of a species with little consideration to individual species' traits. We tested five commonly used modelling methods on two invasive plant species, the habitat generalist *Bromus tectorum* and habitat specialist *Tamarix chinensis*, to compare model performances, evaluate predictability, and relate results to distribution traits associated with each species. Most of the tested models performed similarly for each species; however, the generalist species proved to be more difficult to predict than the specialist species. The highest area under the receiver-operating characteristic curve values with independent validation data sets of *B. tectorum* and *T. chinensis* was 0.503 and 0.885, respectively. Similarly, a confusion matrix for *B. tectorum* had the highest overall accuracy of 55%, while the overall accuracy for *T. chinensis* was 85%. Models for the generalist species had varying performances, poor evaluations, and inconsistent results. This may be a result of a generalist's capability to persist in a wide range of environmental conditions that are not easily defined by the data, independent variables or model design. Models for the specialist species had consistently strong performances, high evaluations, and similar results among different model applications. This is likely a consequence of the specialist's requirement for explicit environmental resources and ecological barriers that are easily defined by predictive models. Although defining new invaders as generalist or specialist species can be challenging, model performances and evaluations may provide valuable information on a species' potential invasiveness.

Keywords

Biological invasions, ecological niche modelling, generalist species, invasive species, non-native species, predictive modelling, specialist species.

INTRODUCTION

Invasion of non-native plant species continues to be a high priority issue for natural resource managers throughout the USA. The impacts that these species have on the landscape include the loss of native biodiversity (Rosentreter, 1994; Randall, 1996), alteration of ecosystem processes (Vitousek *et al.*, 1997; D'Antonio *et al.*, 1999; Mack *et al.*, 2000), and reduced availability of resources (Busch & Smith, 1995; Olson, 1999; Si *et al.*, 2005). Only a small percentage of plant species introduced to the USA become invasive, and a species' success is generally attributed to species traits and the availability of resources that enable it to survive and flourish within a system (Baker, 1974; Lodge, 1993; Cronk & Fuller, 1995; Planty-Tabacchi *et al.*, 1996; Thebaud *et al.*, 1996; Williamson & Fitter, 1996; Mack *et al.*, 2000; Lee, 2001; Stohlgren *et al.*, 2003; Pyšek & Richardson, 2007). The success of an invasive species may be dependent on a single trait that facilitates a competitive advantage under specific ecosystem conditions or multiple traits that allow widespread dispersal across environmental gradients. Given the potential range of survival strategies associated with non-native species, and diverse ecosystem characteristics across North America, identifying or predicting attributes that drive or prohibit invasions can be challenging.

The diversity of survival strategies exhibited by invaders has led some researchers to categorize species as generalists or specialists, and subsets within (e.g. dominant generalist, subdominant generalist; Chong *et al.*, 2006). These categories are loosely defined by the interaction between a species and its environment. Although comparisons between generalist and

DOI: 10.1111/j.1472-4642.2008.00486.x
Journal compilation © 2008 Blackwell Publishing Ltd
No claim to original US government works www.blackwellpublishing.com/ddi

specialist species are common in biological studies (Adler, 1999; Arens, 2001; Bohn & Amundsen, 2001; Lambrinos, 2002), they are rarely considered when forecasting invasions or testing model applications (Seoane *et al.*, 2005; Chong *et al.*, 2006). Most studies use a single model to make predictions for a species of interest (Ficetola *et al.*, 2007; Zhu *et al.*, 2007), while model comparisons tend to test multiple models for a single species (Phillips *et al.*, 2006; Stockman *et al.*, 2006; Pearson *et al.*, 2007) or test individual models for multiple species (Bonn & Schoeder, 2001; Thuiller *et al.*, 2005). In any case, considerations of species traits or dispersal strategies are extremely rare (Segurando & Araújo, 2004; Elith *et al.*, 2006). This is especially common with invasives, where detailed species information is often lacking on autecology and interspecific competition over large and small spatial scales.

Alternatively, models are generally designed to identify ecosystem characteristics that can predict habitat suitability on the landscape, and broadly applied to different species (Phillips *et al.*, 2004; Payne & Stockwell, 2006). The 'one size fits all' approach may be applicable for species that have evolved within restricted or native ranges, but may not always meet the challenges presented by invasive species (Lee, 2001; Doyle *et al.*, 2003; Saura-Mas & Lloret, 2005; Stohlgren *et al.*, 2005; Ellstrand & Schierenbeck, 2006). Most invasive species are new arrivals lacking sufficient residence time to define their ecological niche, reveal biotic and abiotic interactions, or demonstrate species traits that would characterize them as generalist or specialist species (Chong *et al.*, 2006; Stayer *et al.*, 2006). As a result, predictive models for new invaders are often believed to underestimate the potential distribution and habitats at risk (Wilson *et al.*, 2007).

We hypothesize that model performance can be influenced by species-specific traits and that the ecological niche of habitat specialists can be better predicted than those that are characterized as habitat generalists. As a result, models do not perform equally with all species; however, model performance, whether weak or strong, may provide resource managers and researchers with clues on how a particular invader will interact with potential habitats. To test our hypotheses, we selected five common modelling methods to predict the potential distribution of two highly invasive species; a habitat generalist (*Bromus tectorum*, cheatgrass; Mack, 1981; Knapp, 1996) and a habitat specialist (*Tamarix chinensis*, tamarisk; Horton, 1964; Everitt, 1980).

METHODS

Study area

Our study site is the Grand Staircase-Escalante National Monument (GSENM), located in south-central Utah, USA. The Monument covers an area of approximately 769,000 ha with elevations ranging from 1160 m to 2620 m. High plateaus and deeply incised canyons are characteristic of the landscape. The climate of the region is generally temperate and arid with the average annual precipitation approaching 250 mm, mostly occurring during winter months. Mean summer temperatures range from 16 to 32 °C, and winter temperatures range from –9 to 4 °C (National Climatic Data Center, 2003). Despite its extreme climate, the Monument is rich in floral diversity and vegetation types. Over 984 plant species have been recorded within the Monument's boundaries, of which 174 (19%) are regionally endemic and 98 (10%) are non-native (W. Fertig, unpublished data). Only a small proportion of the non-native species exhibit invasive characteristics, while most do not appear to pose a significant threat to the Monument's diverse ecosystems and vegetation types (Waters *et al.*, 2004).

Field data

For this study, we used four independent data sets collected by the Bureau of Land Management (BLM), the Natural Resource Ecology Laboratory (NREL) at Colorado State University, and the US Geological Survey (USGS). BLM data sets included the results from a Noxious Weeds Inventory ($N \geq 24{,}000$) conducted by the BLM Utah State Office in 1997 (M. Benson, unpublished data) and from the Rangeland Health Monitoring Program ($N = 285$) conducted by resource managers at the GSENM between 1999 and 2003 (Pellant *et al.*, 2000). Data sets collected by the NREL and USGS include a 6-year landscape-scale vegetation assessment of GSENM ($N = 380$) conducted between 1998 and 2004 (Waters *et al.*, 2004), and a research project titled Fingerprinting Biodiversity ($N = 380$) conducted between 2005 and 2007 (T.J. Stohlgren & P.H. Evangelista, unpublished data). We randomly chose a selected number of samples from each data set to achieve balance between sample sizes and limit samples to within the Monument's boundaries (see Appendix S1 in Supplementary Material). For *B. tectorum*, we used a total of 366 samples (246 present and 120 absent) for training, and 450 samples (218 present and 232 absent) for validation of the models (see Appendix S2). For *T. chinensis,* a total of 449 samples (226 present and 223 absent) were used for training and 464 samples (227 present and 237 absent) were used for validation of the models (see Appendix S3).

Environmental variables

We generated six different environmental variables from a 10-m Digital Elevation Model (DEM) in ArcGIS 9.1, Arc Toolbox (ESRI, 2004) to represent resource availability and topographical features on the landscape that may facilitate invasion (Davis *et al.*, 2000; Stohlgren *et al.*, 2003). Each variable had a 10-m pixel size and was generated in a grid format that extended over the entire study area. All six candidate variables were considered for each model and for both species. These variables were overland distance to water, slope (degrees), solar insolation, soil wetness index, eastness, and northness. A raster data set for overland distance from water was generated using FloWS Geoprocessing tools (version 1.0; Theobald *et al.*, 2006) and slope was generated using Spatial Analyst in ArcGIS 9.1 (ESRI, 2004). A solar insolation grid was generated using the Shortwave program developed by Kumar *et al.* (1997). Soil wetness index was calculated using the formula $[ln(A/\tan\beta)]$, where $ln(.)$ is the natural logarithm, A is the area drained per unit contour or

specific area, and β is the topographical slope (Moore *et al.*, 1991; Wolock, 1993). Eastness and northness were generated in ArcGIS 9.1 (ESRI, 2004) using the formulas *Sin(A)* and *Cos(A)*, respectively, where *A* is aspect in degrees (Guisan *et al.*, 1999; Gutierrez *et al.*, 2005; Kumar *et al.*, 2006). Previous studies have demonstrated that these landscape characteristics often play significant roles in species distributions and patterns of invasions (Lambrinos, 2002; Stohlgren *et al.*, 2003; Kumar *et al.*, 2006).

Statistical analyses and modelling

The first model tested was a simple envelope design that has no statistical basis, but simply defines the parameters of the independent variables for presence-only data (Envelope; http://www.niiss.org/cwis438/gather/Envelope.php). For each presence point, the value of each independent variable is extracted, with the minimum and maximum values of all presence points defining the probability range of occurrence. The highest probability value for the Envelope model cannot exceed the number of independent variables tested. We also tested maximum entropy (Maxent, version 2.3.18; http://www.cs.princeton.edu/~schapire/maxent/; Phillips *et al.*, 2004, 2006) and genetic algorithm rule-set prediction models (Desktop GARP, version 1.1.6., http://nhm.ku.edu/desktopgarp/; Stockwell & Nobel, 1992; Anderson *et al.*, 2003) which have gained popularity because of their ability to predict species occurrence with presence-only data and small data sets. Finally, we tested two commonly employed regression models; logistic regression (Evans *et al.*, 2000; Pearce & Ferrier, 2000) and classification tree analysis (Breiman *et al.*, 1984; Lewis, 2000), both of which require presence and absence data. Both regression models were analysed using SYSTAT software (version 10.0, Systat Software Inc., San Jose, CA, USA). For logistic regression models, all variables were assessed for multicollinearity and normality. The variables were transformed using $log_{10}+1$ where appropriate. Significance of predictors in the logistic regression model was assessed at alpha 0.05.

The performance of all the models was evaluated using threshold-dependent and threshold-independent measures. Threshold-dependent evaluation was measured by specificity and sensitivity, and Cohen's maximized Kappa (Cohen, 1960). There are two possible errors that may occur in prediction models: false-negatives (under-prediction) and false-positives (over-prediction; Fielding & Bell, 1997). Using the independent validation data for each species, we present the relative proportions of these errors in a confusion matrix. Specificity, the proportion of true-positive and false-positive absences, and sensitivity, the proportion of true-positives and false- positive presences, are reported for the best two models of each species as defined by the area under the curve values described below. Overall accuracy was calculated using the formula $[(a + b)/N]$, where a is the number of correctly classified absences and b is the number of correctly classified presences, and N is the total number of samples (Fielding & Bell, 1997).

Next, we used Cohen's maximized Kappa (Cohen, 1960) to measure the proportion of correctly classified points (i.e. presence, absence) after accounting for the probability of chance agreement. Kappa statistic values range from –1 to +1, where +1 would be perfect agreement and any values less than 0 would indicate a performance no better than random (Cohen, 1960; Allouche *et al.*, 2006). Landis & Koch (1977) ranked analysis performances as poor when kappa values are < 0.40, good when the kappa values range from 0.40 to 0.75, and excellent when kappa values are > 0.75.

The threshold-independent evaluation required a receiver-operating characteristic curve (ROC), where sensitivity is plotted against 1 – specificity for all possible thresholds (Pearce & Ferrier, 2000). From the ROC analysis, we calculated the area under the ROC curves (AUC) using true presence and absence observations to measure the probability that a random, positive point falls within the predicted range of occurrence, and a random negative point falls outside (Fielding & Bell, 1997). The Maxent model generates AUC values using pseudo-random absence points on its own (Phillips *et al.*, 2006), but we do not report these because we chose to maintain consistency with all models tested.

The independent variables used in the model comparisons play different roles of importance in predicting the potential habitat and distribution of each species. The Maxent, logistic regression and classification tree models provide the contributions of each independent variable. The evaluations cannot be directly compared to each other, but are reported to serve as a general guide for ranking the importance of each environmental variable in model predictions.

RESULTS

Generalist species

Models tested for the habitat generalist *B. tectorum* generally did not perform as well as those tested for the habitat specialist *T. chinensis* (Table 1). Evaluation of model performances for the *B. tectorum* training data showed that logistic regression was the best with an AUC value of 0.590. The GARP and Maxent models had slightly weaker predictive strength, while the Envelope model had the lowest performance. The classification tree analysis was unable to make any predictions using the independent variables tested. Evaluation of model performances for the *B. tectorum* validation data rank GARP the best with an AUC value of 0.503 (Table 1) which is no better than random, and all the models were within a range of 0.01 from each other. The Maxent, Envelope, and logistic regression models had only slightly weaker performances, respectively. For *B. tectorum*, the GARP model had an overall accuracy of 54.7%, while Maxent was 54.0% (Table 2).

The predictive contributions of each independent variable for *B. tectorum* were ranked differently for each model. The analysis of variable importance provided by the Maxent program ranked solar insolation as the most important contributor (55.9%); slope as second (24.2%); and overland distance to water as third (12.4%). For logistic regression, soil wetness index was the only significant variable at $P < 0.05$.

No claim to original US government works

Table 1 Performances of different models for *Bromus tectorum* and *Tamarix chinensis* based on training and validation data evaluated by area under the curve (AUC) values and Cohen's maximized kappa.

	B. tectorum				*T. chinensis*			
	Training data		Validation data		Training data		Validation data	
Model	AUC	Kappa	AUC	Kappa	AUC	Kappa	AUC	Kappa
Envelope	0.513	0.034	0.499	–0.002	0.542	0.096	0.458	–0.073
Maxent	0.559	0.055	0.502	0.055	0.773	0.452	0.828	0.612
GARP	0.582	0.212	0.503	0.077	0.721	0.421	0.830	0.63
Classification tree	n.a.	n.a.	n.a.	n.a.	0.741	0.393	0.829	0.66
Logistic regression	0.590	0.151	0.495	0.029	0.744	0.416	0.885	0.694

n.a., not available.

Table 2 Confusion matrix tables for the two best models of *Bromus tectorum* and *Tamarix chinensis* with cut-off at maximized kappa value.

	GARP (*B. tectorum*)			Maxent (*B. tectorum*)			Logistic regression (*T. chinensis*)			Classification tree (*T. chinensis*)		
	Predicted			Predicted			Predicted			Predicted		
	0	1	Correct	0	1	Correct	0	1	Correct	0	1	Correct
0	192	40	82.8%	219	13	94.4%	200	37	85%	187	50	78.2%
1	164	54	24.8%	194	24	11%	34	193	84.4%	29	198	87.2%
Overall			54.7%			54.0%			84.7%			87.2%
Cut-off thresholds			0.85			0.84			0.50			0.55

Specialist species

For the specialist species *T. chinensis*, the Maxent model performed best with the training data, having an AUC value of 0.773 (Table 1). The logistic regression, classification tree, and GARP models also performed strongly while the Envelope model had considerably lower predictive performances than the other models. Model evaluations using the validation data had higher performances overall (Table 1). Logistic regression performed the best with an AUC of 0.885 (Table 1). The GARP, classification tree, and Maxent models performed equally well, while the Envelope model performed poorly. The overall accuracy for the two best models for *T. chinensis*, logistic regression and classification tree, was 84.7% and 87.2%, respectively (Table 2).

The predictive contributions of the independent variables for the models of *T. chinensis* favoured overland distance to water for the Maxent, logistic regression, and classification tree models. Maxent ranked overland distance to water as the most important contributing variable (82.1%), slope as second (8.3%), and soil wetness index as third (2.9%). Logistic regression found overland distance to water as the only significant predictor at $P < 0.05$. Similarly, the classification tree analysis found overland distance to water as the only predictor for *T. chinensis* and split the tree at two different distances. Seventy percent of the presence points (mean = 0.70, standard deviation = 0.45) occurred when overland distance to water was < 314 m, and *T. chinensis* was not found in the training data when the overland distance to water was > 1731 m.

DISCUSSION

Predictability of a generalist versus a specialist

As a generalist species, *B. tectorum* has become widely distributed throughout the GSENM and persists in a broad range of habitat types and environmental conditions (Evangelista *et al.*, 2004; Guenther *et al.*, 2004; Chong *et al.*, 2006; Crall *et al.*, 2006). Field data collected during the landscape-scale vegetation assessment found *B. tectorum* in 93% of the plots, making it the second most common species recorded in GSENM (Waters *et al.*, 2004). Therefore, it was not surprising to see that the models tested in this study had difficulty in defining *B. tectorum*'s ecological niche or identifying specific landscape characteristics that may inhibit the species' distribution (Fig. 1). All the models had consistently weaker performances for the habitat generalist *B. tectorum* than for the habitat specialist *T. chinensis* (Table 1). Other modelling efforts for *B. tectorum* (Alley *et al.* 2004; Crall *et al.*, 2006) and *Tamarix* species (Davern, 2006; Morissette *et al.*, 2006) have had similar results. Some of the models, such as Envelope, appear to have over-predicted the range of *B. tectorum*, which has been a

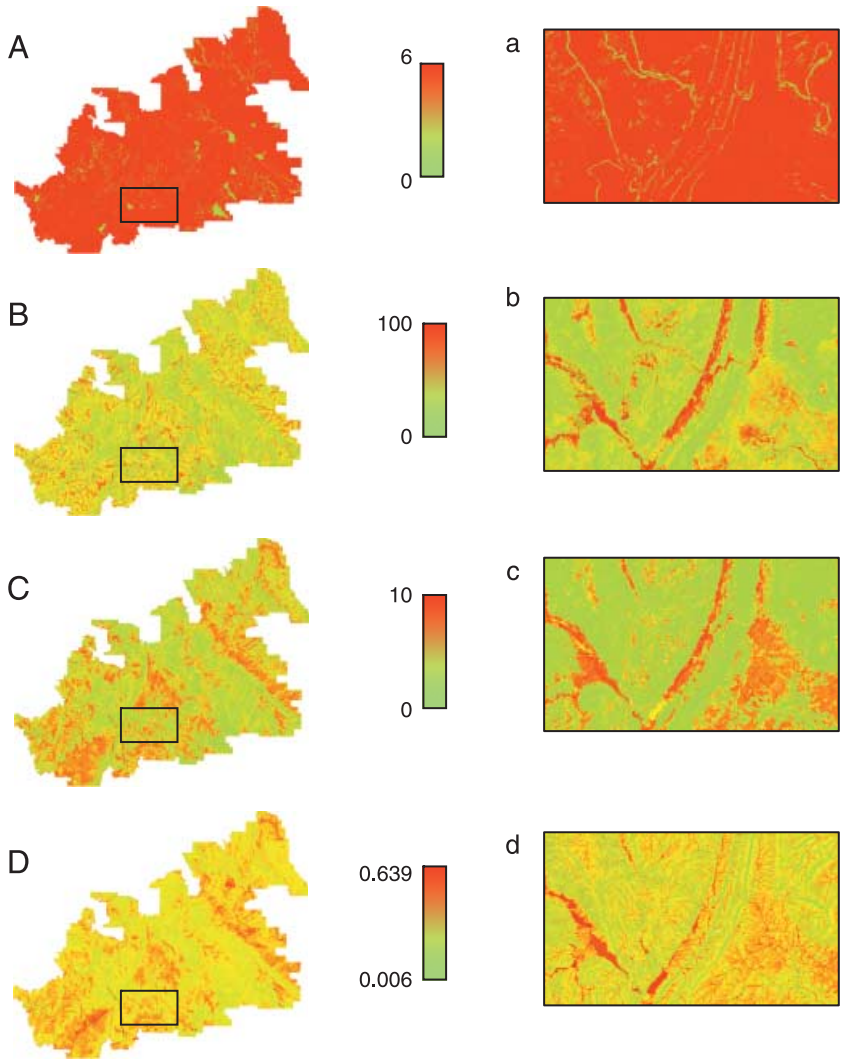

Figure 1 Predicted habitat and potential distribution of *Bromus tectorum* in the Grand Staircase-Escalante National Monument (A–D) and at a finer scale along the Hackberry Creek, Cottonwood Creek, and Paria River confluences (a–d; see inset). Models compared are Envelope (A, a), Maxent (B, b), GARP (C, c), and logistic regression (D, d).

concern with other generalist species (Dettmers *et al.*, 2002; Hepinstall *et al.*, 2002; Seoane *et al.*, 2005). However, given the pervasiveness of *B. tectorum*, the predicted surface from Envelope appears to match field observations better than the more conservative model surfaces (Fig. 1).

In contrast, most of the models for the habitat specialist *T. chinensis* had stronger predictive performances. This trend might be expected because specialist species tend to have sharply defined niches and environmental barriers that are easier to distinguish by model analyses (Bohn & Amundsen, 2001; Caley & Munday, 2003). In our model comparisons, overland distance to water was the most significant independent variable in the three models that provided predictive contributions of independent variables, which further highlights the specialist nature of *T. chinensis* that largely confines it to riparian systems.

Model performance and interpretation

Despite varying performances, we believe that all the models tested still have promising applications in predicting suitable habitat and the potential distribution of generalist and specialist invasive species. The results from this study suggest that no single model is superior in all circumstances (Table 1) supporting reviews from other studies (Elith *et al.*, 2006). Most of the models we tested performed similarly with each species, suggesting that the varying results between the generalist and specialist species are likely correlated to species traits or the independent variables tested. On average, the models for *B. tectorum* performed no better than random, while the predictive contributions of each independent variable tested for *B. tectorum* fluctuated considerably between the Maxent and logistic regression models. In contrast, most of the models for *T. chinensis* performed strongly. The Maxent, logistic regression, and classification tree models agreed that one independent variable (i.e. overland distance to water) was highly significant in predicting the habitat specialist *T. chinensis*.

Closer examination of the predicted surfaces for *T. chinensis* suggests that Maxent performed slightly better at a finer scale than the other models. For example, the lower portion of Hackberry Creek is cradled by steep canyon walls that are

No claim to original US government works

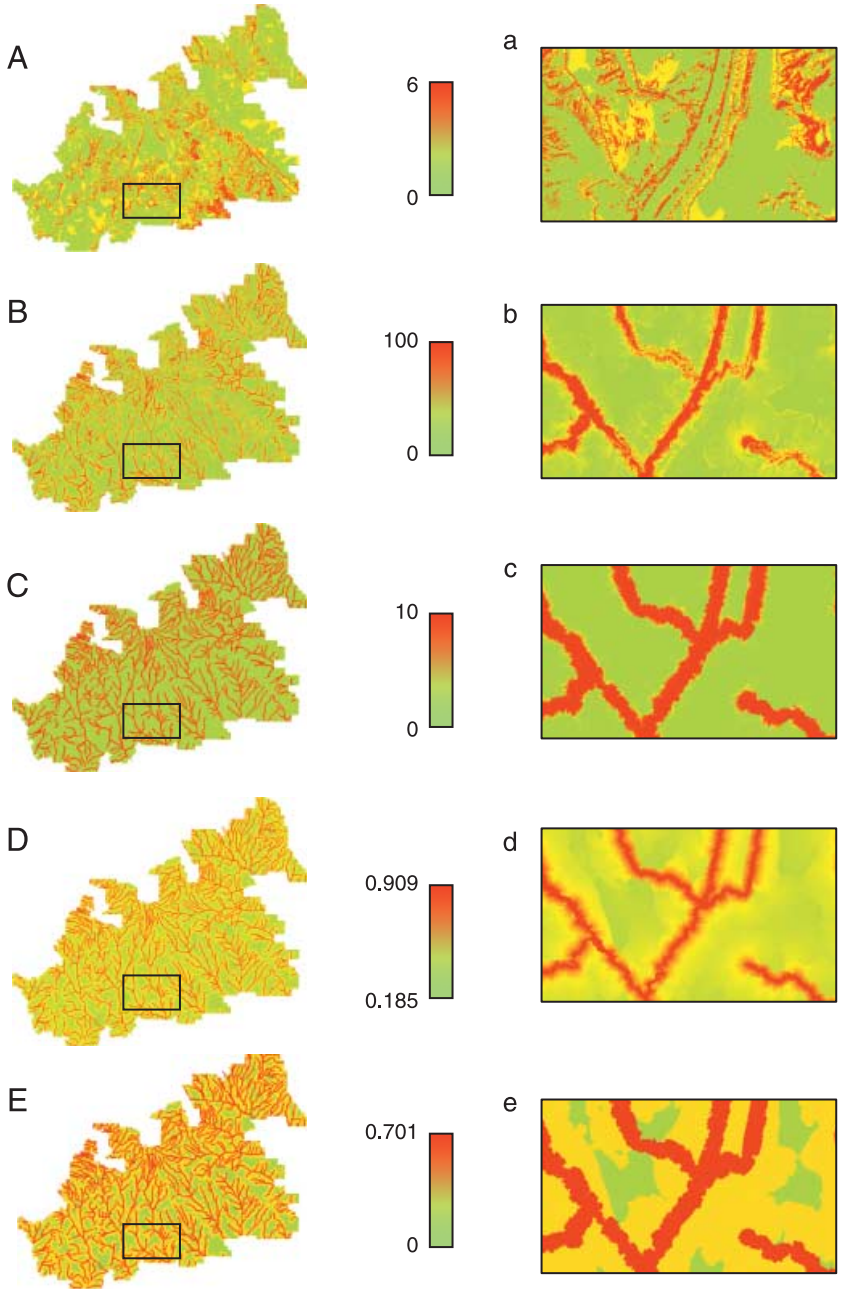

Figure 2 Predicted habitat and potential distribution of *Tamarix chinensis* in the Grand Staircase-Escalante National Monument (A–E) and at a finer scale along the Hackberry Creek, Cottonwood Creek, and Paria River confluences (a–d; see inset). Models compared are Envelope (A, a), Maxent (B, b), GARP (C, c), logistic regression (D, d), and (E, e) classification tree.

prohibitive to most plant establishment (Figs 1a–d and 2a–e). Above the canyons are flat xeric landscapes dominated by *Pinus edulus* (pinyon pine), *Juniperus osteosperma* (Utah juniper), and *Artemisia tridentata* (big sage). *Tamarix chinensis* is absent from these communities, and these communities are highly unlikely to support future establishment of the specialist species (P.H.E., pers. obs.). The sensitivity of Maxent was able to distinguish canyon walls from the riparian system, but incorrectly predicted potential *T. chinensis* habitat on the xeric plateau above Hackberry Creek. The logistic regression, classification tree, and GARP models made coarser predictions and did not account for the steep slopes of the canyon walls while predicting *T. chinensis* to occur on the xeric plateaus.

Improved predicted surfaces for both species may have been likely with inclusion of additional independent variables. For example, *T. chinensis* models may be improved by including variables related to stream flow, native species richness, or climate (Horton, 1977; Brotherson & Field, 1987; Morissette *et al.*, 2006). Similarly, the performance of models for *B. tectorum* may have been improved had variables for fire history, livestock grazing, and recreational use been available (Evangelista *et al.*, 2004; Crall *et al.*, 2006). Many studies have found that model performance greatly relies on the particular independent variables tested, spatial resolution, or inadequate field data (Suarez-Seoane *et al.*, 2002; Gibson *et al.*, 2004; Barry & Elith, 2006). We do not suggest that the independent variables selected for this

study are completely representative of ecological niches for each species; however, they are known to have significant influences on invasive plant species in general (Stohlgren *et al.*, 2003) and were selected with consideration of our objectives in comparing models and species.

When models perform strongly, it is generally indicative that key variables associated with habitat suitability and species traits were successfully identified for the analyses (Boyce *et al.*, 2002; McKenney *et al.*, 2003; Gibson *et al.*, 2004). It is also reasoned that when models perform poorly, the analyses were lacking significant variables to predict suitable habitats (Waters *et al.*, 2004) or relied on inconclusive field data (Hernandez *et al.*, 2006; Real *et al.*, 2006). Although these are rational interpretations in the model process, they are not easily corrected when assessing risk or modelling distribution of invasive species. If model performances are influenced by species traits (e.g. habitat generalist or specialist), then a broader interpretation of results may provide important insight into species traits, thus gauging the potential distribution of new invaders. We suggest that predictive models that perform poorly may not always be a shortcoming in methodology, the independent variables selected, or the completeness of occurrence data. An alternative interpretation might consider weak model performances as a clue to species traits and an early warning of a generalist's positive response to new habitats. Although this is demonstrated by the poor performances of all the models tested with *B. tectorum*, it is highlighted by the inability of the classification tree analysis to produce any results. Overall results of our analyses of *B. tectorum* further highlight the pervasive nature of the species and the models' inability to define any significant barriers to invasion among the independent variables tested. Our results were consistent with those from other studies on *B. tectorum* (Evangelista *et al.*, 2004; Waters *et al.*, 2004), but a thorough review of the literature examining different species and model responses was inconclusive since poor model performances are rarely published.

CONCLUSIONS

Although most of the models we tested generally performed similarly with both species, there were some discrete differences noted. The models tested each have unique analytical approaches that may function better or worse with different species and data sets (Segurando & Araújo, 2004; Pearson *et al.*, 2007; Ward, 2007). For this reason, we recommend that users select several models to test their data. The scientific literature offers numerous model comparisons suggesting that some models are superior to others. We do not discount the results of any individual study; however, collectively they reveal that model performances can exhibit different degrees of variability.

Our study also demonstrates the importance of model functions and evaluations in interpreting results. We recommend the selection of models that have built-in functions that evaluate the performance of results (e.g. R^2, AUC). Although model evaluations can be conducted independently of models, they require access to different software and a comprehensive understanding of statistical methodology. Additionally, models that measure the predictive contributions of the independent variables (e.g. standardized regression coefficients, jackknife) will provide users with valuable clues regarding environmental conditions that may prohibit or facilitate species dispersal and establishment. Maxent and classification tree models, for example, provide user-friendly and comprehensive operations that quantify the predictive strength of each independent variable tested. Users should consider the results from all the available functions when evaluating model performances or making predictions on a species' potential distribution, while considering expert knowledge whenever possible.

Because this study only examined two species, we cannot conclude that model performance alone can be used to gauge species traits or potential distribution risks of invasive species. However, this study offers several insights into model performances and their relations to species traits. Only a few studies on these relationships have been conducted in any depth (Chong *et al.*, 2006; McPherson & Jetz, 2007), leaving a large gap in our understanding of the effects that autecology has on predictive modelling. Our results also demonstrate the need for further studies that include a wide array of species (native and non-native) that exhibit different traits and dispersal strategies. Specifically, increased research may prove to be especially useful in assessing the risk of new invasive species.

ACKNOWLEDGEMENTS

The authors would like to thank the Grand Staircase-Escalante National Monument, Utah Bureau of Land Management, US Geological Survey Fort Collins Science Center, NASA Goddard Space Flight Center, and Natural Resource Ecology Laboratory at Colorado State University for collaborative support and data sharing. Geneva Chong, Tracy Davern, and anonymous reviewers provided valuable comments and discussions on model and manuscript preparation. We would also like to thank Matt Benson for GIS contributions, Steven Phillips for Maxent support and consulting, and Nate Alley, Amber Hughes, and Deb Guenther for data collection.

REFERENCES

Adler, L.S. (1999) Size-mediated performance of a generalist herbivore feeding on mixed diets. *The Southwestern Naturalist*, **49**, 189–196.

Alley, N., Stolhgren, T.J., Evangelista, P. & Guenther, D. (2004) Iterative model development for natural resource managers: a case example in Utah's Grand Staircase-Escalante national Monument. *Geographic Information Sciences*, **10**, 1–9.

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, Kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Anderson, R.P., Lew, D. & Peterson, A.T. (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modeling*, **162**, 211–232.

Arens, N.C. (2001) Variation in performance of the tree fern *Cyathea caracasana* (*Cyatheaceae*) across a successional mosaic in an Andean Cloud Forest. *American Journal of Botany*, **88**, 545–551.

Baker, H.G. (1974) The evolution of weeds. *Annual Review of Ecological Systems*, **5**, 1–24.

Barry, S. & Elith, J. (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, **43**, 413–423.

Bohn, T. & Amundsen, P. (2001) The competitive edge of an invading specialist. *Ecology*, **82**, 2150–2163.

Bonn, A. & Schroder, B. (2001) Habitat models and their transfer for single and multi species groups: a case study of carabids in an alluvial forest. *Ecography*, **24**, 483–496.

Boyce, M., Vernier, P.R., Nielsen, S.E. & Schmiegelow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modeling*, **157**, 281–300.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.G. (1984) *Classification and regression trees*. Wadsworth International Group, Belmont, California.

Brotherson, J.D. & Field. D. (1987) *Tamarix*: impacts of a successful weed. *Rangelands*, **9**, 110–112.

Busch, D.E. & Smith, S.D. (1995) Mechanisms associated with decline of woody species in riparian ecosystems of the southwestern US. *Ecological Monographs*, **65**, 347–370.

Caley, M.J. & Munday, P.L. (2003) Growth trade off with habitat specialization. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **270**, S175–S177.

Chong, G.W., Otsuki, Y., Stohlgren, T.J., Guenther, D., Evangelista, P., Villa, C. & Waters, A. (2006) Evaluating plant invasions from both habitat and species perspectives. *Western North American Naturalist*, **66**, 92–105.

Cohen, J. (1960) A coefficient of agreement of nominal scales. *Educational and Psychological Measurement*, **20**, 37–46.

Crall, A.W., Newman, G.J., Stohlgren, T.J., Jarnevich, C.S., Evangelista, P. & Guenther, D. (2006) Evaluating dominance as a component of non-native species invasions. *Diversity and Distributions*, **12**, 195–204.

Cronk, C.B. & Fuller, J.L. (1995) *Plant invaders: the threat to natural ecosystems*. Chapman & Hall, London.

D'Antonio, C.M., Dudley, T.L. & Mack, M. (1999) Disturbance and biological invasions: direct effects and feedbacks. *Ecosystems of disturbed ground* (ed. by L. Walker), pp. 143–452. Elsevier, Amsterdam.

Davern, T. (2006) *A predictive model: Tamarisk habitat in California and Colorado*. Thesis. Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Fort Collins, Colorado.

Davis, M.A., Grime, J.P. & Thompson, K. (2000) Fluctuating resources in plant communities: a general theory of invisibility. *Journal of Ecology*, **88**, 528–534.

Dettmers, R., Buehler, D. & Bartlett, J. (2002) A test and comparison of wildlife-habitat modeling techniques for predicting bird occurrence at a regional scale. *Predicting species occurrences. Issues of scale and accuracy* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 607–615. Island Press, Washington, D.C.

Doyle, R.D., Francis, M.D. & Smart, R.M. (2003) Interference competition between *Ludwigia repens* and *Hygrophila polysperma*: two morphologically similar aquatic plant species. *Aquatic Botany*, **77**, 223–234.

Elith, J., Graham, C.J., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R., Huettmann, F., Leathwick, J., Lehmann, A., Li, J., Lohmann, L., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Peterson, A.T., Phillips, S., Richardson, K., Scachetti-Pereia, R., Schapire, R., Soberon, J., Williams, S., Wisz, M. & Zimmermann. N. (2006) Novel methods improve prediction of species distributions from occurrence data. *Ecography*, **29**, 129–151.

Ellstrand, N.C. & Schierenbeck, K.A. (2006) Hybridization as a stimulus for the evolution of invasiveness in plants? *Euphytica*, **148**, 35–46.

ESRI. (2004) *ArcGIS 9.1*. ESRI, Redlands, California. http://www.esri.com/index.html.

Evangelista, P., Guenther, D., Stohlgren, T.J. & Stewart, S. (2004) Vegetation response to fire and post-burn seeding treatments in juniper woodlands of the Grand Staircase-Escalante National Monument, Utah. *Western North American Naturalist*, **64**, 293–305.

Evans, M., Hastings, N. & Peacock, B. (2000) Bernoulli distribution. *Statistical distributions*, 3rd edn. pp. 31–33. Wiley, New York.

Everitt, B.L. (1980) Ecology of saltcedar – a plea for research. *Environmental Geology*, **3**, 77–84.

Ficetola, G.F., Thuiller, W. & Miaud, C. (2007) Prediction and validation of the potential global distribution of problematic alien invasive species – the American bullfrog. *Diversity and Distributions*, **13**, 476–485.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Gibson, L.A., Wilson, B.A., Cahill, D.M. & Hill, J. (2004) Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *Journal of Applied Ecology*, **41**, 231–223.

Guenther, D.A., Stohlgren, T.J. & Evangelista, P. (2004) Relict sites compared to grazed landscapes in the Grand Staircase-Escalante National Monument, Utah. *The Colorado plateau: cultural, biology, and physical research* (ed. by C. van Riper III and K.L.Cole), pp. 153–162. The University of Arizona Press, Tucson, Arizona.

Guisan, A., Weiss, S.B. & Weiss, A.D. (1999) GLM versus CCA spatial modelling of plant species distribution. *Plant Ecology*, **143**, 107–122.

Gutierrez, D., Fernandez, P., Seymour, A.S. & Jordano, D. (2005) Habitat distribution models: are mutualist distributions good predictors of their associates? *Ecological Applications*, **15**, 3–18.

Hepinstall, J., Krohn, W. & Sadler, S. (2002) Effects of niche width on the performance and agreement of avian habitat models. *Predicting species occurrences. Issues of scale and accuracy* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 593–606. Island Press, Washington, D.C.

Hernandez, P., Graham, C., Master, L. & Albert, D. (2006) The effect of sample size and species characteristics on performance of different species distribution modelling methods. *Ecography*, **29**, 773–785.

Horton, J.S. (1964) *Notes on the introduction of deciduous Tamarix*. US Department of Agriculture. Forest Service, Research Note RM-16. 8 pp.

Horton, J.S. (1977) The development and perpetuation of the permanent tamarisk type in the phreatophyte zone of the Southwest. Department of Agriculture, Forest Service, General Technical Report RN 43: 124–127.

Knapp, P. (1996) Cheatgrass (*Bromus tectorum* L.) dominance in the Great Basin Desert. *Global Environmental Change*, **6**, 37–52.

Kumar, L., Skidmore, A.K. & Knowles, E. (1997) Modeling topographic variation in solar radiation in a GIS environment. *International Journal of Geographical Information Science*, **11**, 475–497.

Kumar, S., Stohlgren, T.J. & Chong, G. (2006) Spatial heterogeneity influences native and nonnative plant species richness. *Ecology*, **87**, 3186–3199.

Lambrinos, J.G. (2002) The variable invasive success of *Cortaderia* species in a complex landscape. *Ecology*, **83**, 518–529.

Landis, J.R. & Koch, G.C. (1977) The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159–174.

Lee, M. (2001) *Non-native plant invasions in Rocky Mountain National Park: linking species traits and habitat characteristics*. Thesis, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado.

Lewis, R.J. (2000) *An introduction to classification and regression (CART) analysis*. Presented at the 2000 Annual Meeting of the Society of Academic Emergency Medicine. San Francisco, California. http://www.saem.org/download/lewis1.pdf (last accessed 7 July 2007).

Lodge, D.M. (1993) Biological invasions: lessons for ecology. *Trends in Ecology and Evolution*, **8**, 133–137.

Mack, R.N. (1981) Invasion of *Bromus tectorum* L. into Western North America: an ecological chronicle. *Agro-Ecosystems*, **7**, 145–165.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M. & Bazzaz, F. (2000) Biotic Invasions: causes, epidemiology, global consequences and control. *Issues in Ecology*, **5**, 1–19.

McKenney, D.W., Hopkin, A.A., Campbell, K.L., Mackey, B.G. & Foottit, R. (2003) Opportunities for improved risk assessments of exotic species in Canada using bioclimatic modeling. *Environmental Monitoring and Assessment*, **88**, 445–461.

McPherson, J.M. & Jetz, W. (2007) Effects of species' ecology on the accuracy of distribution models. *Ecography*, **30**, 135–151.

Moore, I.D., Grayson, R.B. & Ladson, A.R. (1991) Terrain based catchment partitioning and runoff prediction using vector elevation data. *Water Resource Research*, **27**, 1177–1191.

Morissette, J.T., Jarnevich, C.S., Ullah, A., Cai, W., Pedelty, J.A., Gentle, J.E., Stohlgren, T.J. & Schnase, J.L. (2006) A tamarisk habitat suitability map for the continental United States. *Frontiers in Ecology and the Environment*, **4**, 11–17.

National Climatic Data Center (2003) http://www.ncdc.noaa.gov/oa/ncdc.html

Olson, B.E. (1999) Impacts of noxious weeds on ecologic and economic systems. *Biology and management of noxious rangeland weeds* (ed. by R.L. Sheley and J.K. Petroff), pp. 4–18. OSU Press, Corvallis, Oregon.

Payne, K. & Stockwell, D.R.B. (2006) *GARP modeling system user's guide and technical reference*. http://www.landshape.org/enm/

Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modeling*, **133**, 225–245.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distribution from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.

Pellant, M., Shaver, P., Pyke, D. & Herrick, J. (2000) *Interpreting indicators of rangeland health*, Version 3, pp. 118. United States Department of Interior, Bureau of Land Management, Technical Reference 1734–6. Denver, Colorado.

Phillips, S.J., Dudik, M. & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on Machine Learning*, pp. 655–662. ACM Press, New York.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modeling*, **190**, 231–259.

Planty-Tabacchi, A., Tabacchi, E., Naiman, R.J., Deferrari, C. & Decamps, H. (1996) Invasibility of species-rich communities in riparian zones. *Conservation Biology*, **10**, 598–607.

Pyšek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological invasions* (ed. by W. Nentwig), pp. Ecological Studies 193. Springer, Berlin.

Randall, J.M. (1996) Weed control for the preservation of biodiversity. *Weed Technology*, **10**, 370–383.

Real, R., Barbosa, A.M. & Vargas, J.M. (2006) Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, **13**, 237–245.

Rosentreter, R. (1994) Displacement of rare plants by exotic grasses. *Proceedings, ecology and management of annual rangelands* (ed. by S.B. Monsen and S.G. Kitchen), pp. 170–175. USDA-USFS, General Technical Report INT-GTR-313, Intermountain Research Station, Ogeden, Utah.

Saura-Mas, S. & Lloret, F. (2005) Wind effects on dispersal patterns of the invasive alien *Cortaderia selloana*. Mediterranean wetlands. *Acta Oecologica*, **27**, 129–133.

Segurando, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

Seoane, J., Carrascal, L., Luis Alonso, C. & Palomino, D. (2005) Species-specific traits associated to prediction errors in bird habitat suitability modeling. *Ecological Modeling*, **185**, 299–308.

Si, J.H., Feng, Q., Zhang, X., Liu, W., Su, Y. & Zhang, Y. (2005) Growing season evapotranspiration from *Tamarix ramosissima* stands under extreme arid conditions in northwest China. *Environmental Geology*, **48**, 861–870.

No claim to original US government works

Stayer, D.L., Eviner, V.T., Jescheke, J.M. & Pace, M.L. (2006) Understanding the long-term effects of species invasions. *Trends in Ecology and Evolution*, **21**, 645–651.

Stockman, A., Beamer, D. & Bond, J. (2006) Am evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Diversity and Distributions*, **12**, 81–89.

Stockwell, D. & Nobel, I.R. (1992) Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematics and Computer in Simulation*, **33**, 385–390.

Stohlgren, T.J., Barnett, D. & Kartez, J.T. (2003) The rich get richer: patterns of plant invasions in the United States. *Frontiers in Ecology and the Environment*, **1**, 11–14.

Stohlgren, T.J., Crosier, C., Chong, G., Guenther, D. & Evangelista, P. (2005) Life-history habitat matching in invading non-native plant species. *Plant and Soil*, **277**, 7–18.

Suarez-Seoane, S., Osborne, P.E. & Carlos Alonso, J. (2002) Large-scale habitat selection by agriculture steppe birds in Spain: identifying species-habitat responses using generalized additive models. *Journal of Applied Ecology*, **39**, 755–771.

Thebaud, C.A., Finzi, C., Affre, L., Debusscche, M. & Escarre, J. (1996) Assessing why two introduced *Conyza* differ in their ability to invade Mediterranean old fields. *Ecology*, **77**, 791–804.

Theobald, D.M., Norman, J.B., Peterson, E., Ferraz, S., Wade, A. & Sherburne, M.R. (2006) *Functional linkage of water basins and streams (FLoWS) v1 user's guide: ArcGIS tools for network-based analysis of freshwater ecosystems*. pp. 43. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado.

Thuiller, W., Lavorel, S. & Araújo, M.B. (2005) Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, **14**, 347–357.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Rejmánek, M. & Westbrooks, R. (1997) Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, **21**, 1–16.

Ward, D. (2007) Modeling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions*, **9**, 1387–3547.

Waters, M.A., Stohlgren, T.J., Evangelista, P.H., Guenther, D.A., Alley, N. & Newman, G.J. (2004) *Landscape-scale assessment of Grand Staircase-Escalante National Monument: technical report 1998–2004*. pp. 205. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado.

Williamson, M. & Fitter, A. (1996) The varying success of invaders. *Ecology*, **77**, 1661–1666.

Wilson, J.R.U., Richardson, D.M., Rouget, M., Procheş, Ş., Amis, M.A., Henderson, L. & Thuiller, W. (2007) Residence time and potential range: crucial considerations in modelling plant invasions. *Diversity and Distributions*, **13**, 11–22.

Wolock, D.M. (1993) TOPMODEL, a topography based watershed model. *Proceedings of the Federal Interagency Workshop on Hydrologic Modeling Demands for the 90s*. pp. 8-50-8-57. USGS Water-Resources Investigations Report 93-4018.

Zhu, L., Sun, O., Sang, W., Li, Z. & Ma, K. (2007) Predicting the spatial distribution of an invasive plant species (*Eupatorium adenophorum*) in China. *Landscape Ecology*, **22**, 1143–1154.

Editor: Mark Robertson

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Data sources and sample sizes used in models and validation of *Bromus tectorum* and *Tamarix chinensis*.

Appendix S2 Sample locations of *Bromus tectorum* in Grand Staircase-Escalante National Monument. Map A shows the points used for training models and Map B shows the points used for validating models.

Appendix S3 Sample locations of *Tamarix chinensis* in Grand Staircase-Escalante National Monument. Map A shows the points used for training models and Map B shows the points used for validating models.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1472-4642.2008.00486.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.