Biol Invasions (2009) 11:1373–1379
DOI 10.1007/s10530-008-9345-8

ORIGINAL PAPER

Near term climate projections for invasive species distributions

Catherine S. Jarnevich · Thomas J. Stohlgren

Received: 17 June 2008 / Accepted: 15 August 2008 / Published online: 3 September 2008
© Springer Science+Business Media B.V. 2008

Abstract Climate change and invasive species pose important conservation issues separately, and should be examined together. We used existing long term climate datasets for the US to project potential climate change into the future at a finer spatial and temporal resolution than the climate change scenarios generally available. These fine scale projections, along with new species distribution modeling techniques to forecast the potential extent of invasive species, can provide useful information to aide conservation and invasive species management efforts. We created habitat suitability maps for *Pueraria montana* (kudzu) under current climatic conditions and potential average conditions up to 30 years in the future. We examined how the potential distribution of this species will be affected by changing climate, and the management implications associated with these changes. Our models indicated that *P. montana* may increase its distribution particularly in the Northeast with climate change and may decrease in other areas.

Keywords Invasive species · Climate change · GARP · Modeling · Species distributions · Climate variation

C. S. Jarnevich (✉) · T. J. Stohlgren
Fort Collins Science Center, U.S. Geological Survey, 2150 Centre Avenue, Building C, Fort Collins, CO 80526, USA
e-mail: jarnevichc@usgs.gov

Introduction

Climate change and invasive species are two of the most important issues facing the conservation of biodiversity today (Thomas et al. 2004). Recently, there have been many attempts to forecast the impact of climate change on the potential distribution of both native and non-native species. Researchers have used species environment models that capture the niche of a species using current climate data and projected this niche onto future climate scenarios (for examples, see Hijmans and Graham 2006, Levinsky et al. 2007). These climate scenarios, however, may have greater variability within a single scenario's different realizations than between different scenarios, leading to great uncertainty when applying these models to species distributions (Beaumont et al. 2007).

Most available climate scenarios exist at relatively coarse spatial resolutions (e.g., a minimum of 0.5°), and most vegetation distribution models for future climates have been generated at these coarse resolutions (for example, see Levinsky et al. 2007). These resolutions have utility for determining patterns of change and their effects at national, continental and global scales. However, while these may be useful for predictions of climate effects such as inundation of coastal area by rising sea level, they are not particularly useful for projecting impacts to local scales, like the relatively small natural areas scattered across the United States. Additionally, the readily

available climate change projections, even those down-scaled to finer resolutions, still have a coarse temporal scale as they project changes for 100 years into the future (e.g., Govindasamy et al. 2003). Resource managers need short term, fine resolution projections of how climate may affect the areas they are trying to manage.

In this paper, we examine the utility of future climate predictions derived from climatic trends in the twentieth century as surrogates for the coarse resolution, long range projections of global climate change scenarios. We use current climate along with the projected climates to map the distribution of *Pueraria montana* (kudzu), an invasive vine in the Southeastern United States, and examine the potential change in its distribution with the change in climate. We also examine the inter-annual variability and coefficient of variation for the climatic variables in the past to pinpoint locations potentially more susceptible to climate variability.

Methods

Datasets

Climate data consisted of 4 km monthly data for the years 1895–2006 and 800 m data averaged over 1971–2000 for precipitation, minimum temperature, and maximum temperature (PRISM data available from http://www.prism.oregonstate.edu/, Daly et al. 2000). The 30 year average dataset was used as input for current distributions, while the 112 year dataset was used to analyze climatic trends and to predict future climate.

Field data locations for our example model of *P. montana* were obtained from the National Institute of Invasive Species Science website (http://www.niiss.org) and consisted of data points contributed from the Southern Appalachian Information Node of the National Biological Information Infrastructure, the Southeast Exotic Pest Council early detection and distribution mapping system, and the Invasive Plants Atlas of New England. Additional data points were obtained from the newly developed Invasive Plant Atlas of the MidSouth (http://www.gri.msstate.edu/research/ipams/ipams.php), the Illinois Department of Natural Resources and TEXASINVASIVES.ORG. These locations totaled 993 presence points scattered across 11 states in the Eastern US. These data do not capture all known locations, but are representative of the species' extent.

Climate analyses

Simple linear regressions were performed for individual grid cells for each climate variable including precipitation, minimum temperature, and maximum temperature for the average annual data and the average monthly data (Daly et al. 2000). All calculations were carried out in ArcGIS raster calculator (ESRI) using the raster (grid) layer for each of the variables. The slope (m) of the trend was calculated as:

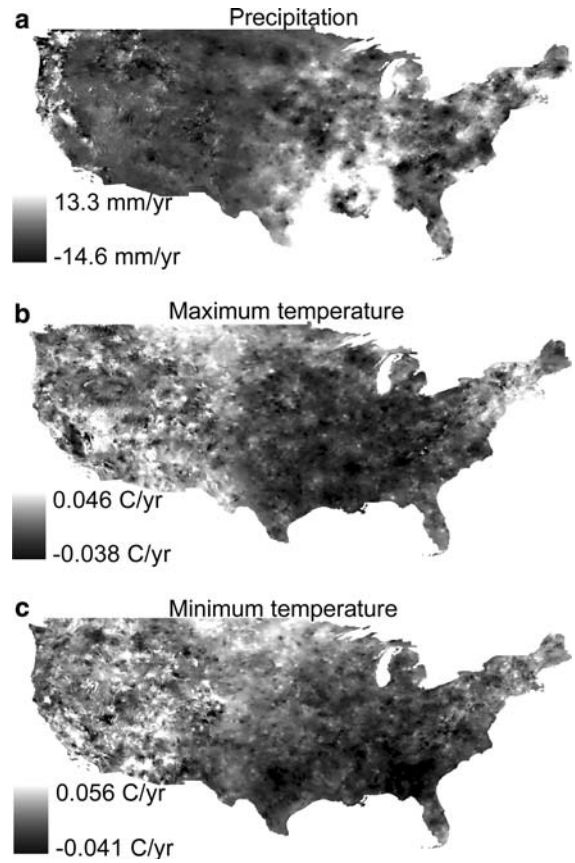

Fig. 1 Slope of the line indicating the 112 year annual trend (1895–2006) for the United States at each grid cell location for (**a**) precipitation (mm/year), (**b**) maximum temperature (°C/year) and (**c**) minimum temperature (°C/year) with positive values indicating an increase over the time period and negative values indicating a decrease

$$m = \frac{\sum_{i=1}^{112}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{112}(x_i - \overline{x})^2}$$

where x is the year and y is the value for precipitation or minimum or maximum temperature in year x_i. The intercept of the line was then calculated as:

$$b = \overline{y} - m * \overline{x}$$

where y is the average precipitation or minimum or maximum temperature and x is the average year.

Using these results, we could then calculate potential climate for any future year using the equation for a line, $y = mx + b$, substituting the derived slope raster for m, the desired year for x, and the derived intercept raster for b.

We examined patterns of climate variability in the country by calculating average inter-annual variation for precipitation, minimum temperature, and maximum temperature using the 112 year data as:

$$\text{var} = \frac{\sum_{i=1}^{111}(y_i - y_{i+1})}{111}$$

where y_i is the precipitation or temperature in year i and y_{i+1} is the precipitation or temperature in the subsequent year. Thus, the variation describes the average variability experienced in each grid cell over the last 112 years. Along with inter-annual variation we calculated the coefficient of variation for the climate variables as:

$$\text{cv} = \frac{\sqrt{\left(\frac{1}{n} * \sum_{i=1}^{n}(x_i - \overline{x})\right)}}{\overline{x}}$$

where x is the temperature or precipitation for year i and n is the total number of years (here, 112). For the coefficient of variation calculations, temperature was transformed to F to ensure that values were positive.

Species analysis

For both current and future climate, we generated 19 bioclimatic variables to better represent growing conditions compared to simple precipitation and temperature data following Hijmans et al. (2005). These 19 variables represent not only the annual variation in climate, but also capture seasonal variation and environmental extremes that can facilitate or inhibit plant growth and establishment of populations. They include annual mean temperature, mean diurnal range, isothermality, temperature seasonality, maximum temperatures of the warmest month, minimum temperatures of the coldest month, temperature annual range, mean temperature of the wettest, driest, warmest and coldest quarters, annual precipitation, precipitation of the wettest and driest quarters, precipitation of seasonality, and precipitation of the wettest, driest, warmest, and coldest quarters. We derived bioclimatic variables reflecting current climate from the monthly 30 year average data and bioclimatic variables reflecting potential future climate from the 2035 predicted monthly climate data layers derived from the long-term trend analysis. The variables were calculated using the freely available ESRI ArcInfo AML program (available at http://www.worldclim.org/bioclim.htm). These 19 bioclimatic variables served as inputs to a Genetic Algorithm for Rule-set Prediction model (GARP; openModeller v 1.0.6; Anderson et al. 2003).

GARP is a member of a suite of modeling techniques known as species environmental matching models which have been reviewed and tested in recent literature (see, Elith et al. 2006, Pearson et al. 2006 for reviews). It is one technique requiring only presence locations for a species as dependent data rather than presence and absence locations. The free program (available at http://openmodeller.sourceforge.net) uses a genetic algorithm to select a set of rules that best describe a species' distribution using the presence locations and independent variables (here, the 19 bioclimatic variables). GARP also allows projection of this model onto other geospatial layers, such as future climate scenarios.

In GARP, we specified that 80% of the data be used for training the model, with a convergence limit of 0.01 and the maximum number of iterations set as 1,000. We employed the 'best subsets' procedure available in OpenModeller, specifying that 100 models be run. The final model resulted from a summation of the top 10 of the 100 model runs.

We used a threshold independent measure, the receiver operating characteristic (ROC), which evaluates the performance of a model at all possible thresholds using the area under the curve (AUC; Fielding and Bell 1997), to validate the model. The

ROC values calculated by OpenModeller tend to be inflated because they require both presence and absence data, requiring the use of pseudo-absence locations. We used a separate county level dataset from 2007 that is based on herbarium records and generated by the Biota of North America Program to further validate the model, using Zonal Statistics in Arc to capture the maximum suitability value for each county in the continental United States. We calculated the statistics using ROC_AUC version 1.3 available at http://brandenburg.geoecology.uni-potsdam.de/users/schroeder/download.html. For each county we extracted the maximum suitability value from the GARP model for that county. We also calculated specificity, which is the probability that a presence location is predicted as present (Fielding and Bell 1997).

We compared the model layer produced using current climate to the one produced using potential climate in the year 2035 and to the current distribution reflected by the BONAP county data. For the two GARP models, we used the 'any model predicts' criterion to define locations as present or absent following Raxworthy et al. (2003). We merged the future and current layers together and identified locations classified as (1) presence locations currently but not in the future; (2) locations now and in the future; (3) presence locations only in the future, and; (4) absence locations now and in the future.

Results

The slope of the regression, indicating the rate of change in each variable, varied across the United States (Fig. 1). For precipitation, the average annual rate of the change ranged from a decrease of 14.6 mm/year to an increase of 13.3 mm/year (Fig. 1a). The trend in precipitation over the time period showed the greatest increases in the Midsouth to the Midwest with other smaller spots along both coasts, while greatest decreases occurred in the central to western United States. Average rates of change were similar for both maximum (Fig. 1b) and minimum (Fig. 1c) temperature, and the pattern of change was similar with the southeast generally decreasing and the western United States generally increasing in both temperatures. Trends ranged from an increase of 0.05 or 0.06°C/year for maximum and minimum temperature, respectively, to a decrease −0.04°C/year.

Inter-annual variance in precipitation and temperature also varied across the country (Fig. 2a–c). Precipitation seemed to vary the most in the wettest areas of the country, including the southeastern US and the West Coast. Maximum and minimum temperatures differed in their pattern. Both exhibited the greatest fluctuations between years in the north-central US, though greatest variance in maximum temperature extended farther south, while greatest variance in minimum temperature extended east in New England.

Coefficient of variations had similar trends for minimum and maximum temperature, but differed for precipitation. Highest values for temperature were in the north-central US and the Rocky Mountains, while precipitation had the greatest values in the Southwest.

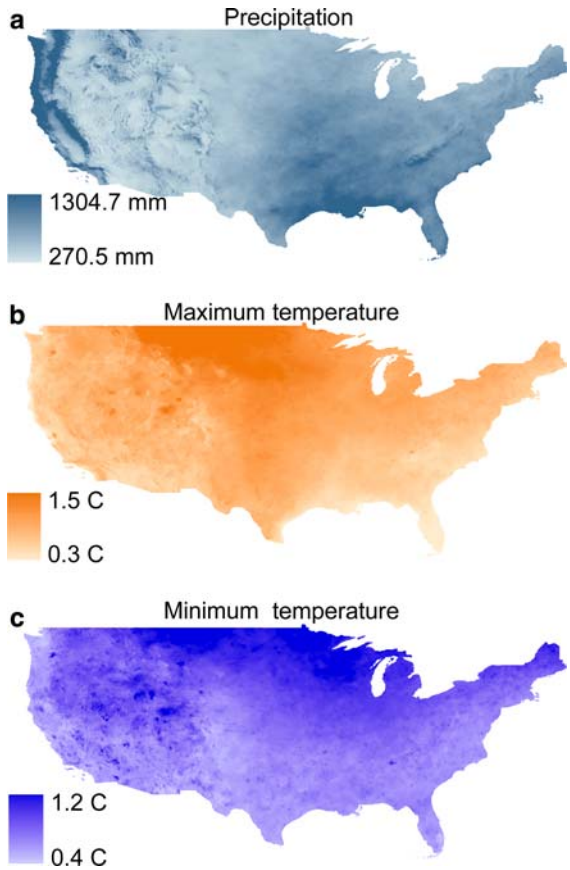

Fig. 2 Average inter-annual variation between 1895 and 2006 for (**a**) precipitation (mm), (**b**) maximum temperature (°C) and (**c**) minimum temperature (°C) for the United States

For temperature, the trends were also similar to annual variance whereas precipitation had an opposite trend.

The GARP model for *P. montana* had an AUC value of 0.93 as calculated by the program, indicating good model performance. Using the county BONAP dataset to validate the model, the AUC was slightly lower at 0.86 and sensitivity was 0.85. The locations missed by the model were at the extremes of the distribution. The most notable differences in the current and 2035 climate models for *P. montana* include changes in habitat suitability for Florida and increases in suitability for the Pacific Northwest, where it has recently appeared, and New England (Fig. 3a, b).

The differences between distributions under current climate and a potential future climate are highlighted by the maps comparing the future scenario to the current known county distribution and the model of current conditions (Fig. 4a, b). The current distribution model comparison highlights much more area as stable. The geographic area indicating the largest amount of land to target for early detection according to both figures is the northwest United States. Differences in the Southeast may be due to lack of reporting, and should probably be targeted for surveys.

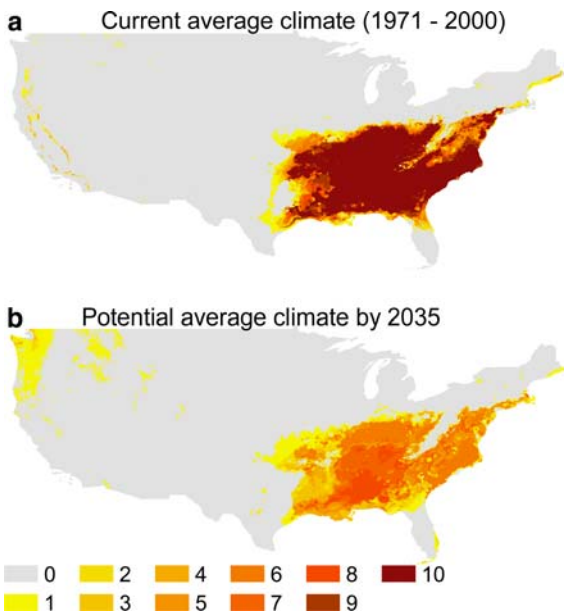

Fig. 3 Potential distribution of *P. montana* in the United States with (**a**) current climatic conditions (average climate from 1971–2000) and (**b**) potential 2035 climate based on the 112 year trend (Fig. 1)

Discussion

Spatial and temporal heterogeneity

The methods described in this paper captured heterogeneity both spatially and temporally. Using data on climatic trends over the last 112 years allowed us to examine spatial and temporal heterogeneity at a finer resolution than that available using global circulation models. Our trend model predictions of climate change are probably conservative estimates, given that they predict a linear change in climate compared to most global circulation model projections which indicate a steeper change. Thus, these conditions may occur sooner than the projected date (2035). The various existing models of climate change (e.g., Hadley, CCM3, etc.) all differ greatly in their predictions of future climate. The climate scenarios

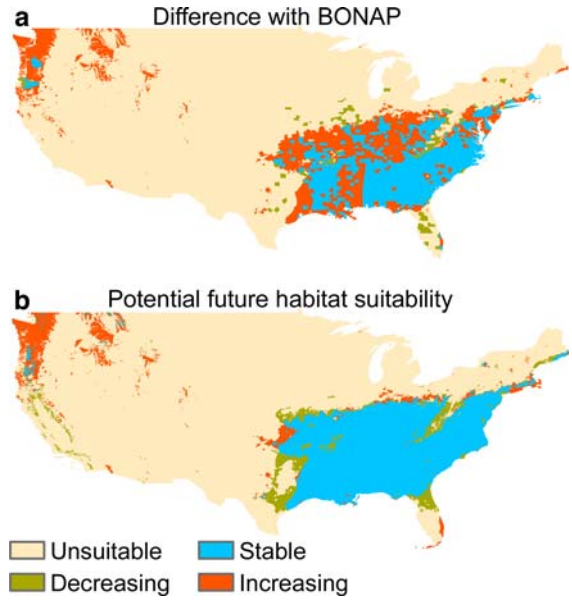

Fig. 4 Potential changes in *P. montana* distribution using the 'any model predicts' criterion for presence between the potential 2035 distribution (Fig. 3b) and (**a**) the BONAP 2007 county distribution and (**b**) the predicted current distribution (Fig. 3a) with unsuitable indicating absent in both time scenarios, decreasing indicating present only under current conditions, stable indicating present in both time scenarios, and increasing indicating present only in the future scenario

based on climatic trends may be a better representation of conditions that may soon occur. These will be transient conditions, and may occur sooner than our year estimate of 2035. However, they still provide resource managers with projections of species' distribution changes due to climate that are likely to occur in the course of their careers. These predictions may show what locations may become unsuitable in the future especially well.

Additionally, examination of our maps of inter-annual variation should inform managers which locations may be most susceptible to changes in climate. Plant species or animal taxa with limited mobility and multi-year life spans would be required to survive changes in climate to persist in locations with high inter-annual variability. These organisms have already experienced temporal heterogeneity in climatic conditions.

Near-term climate change modeling

The *P. montana* example provides a predicted future distribution that incorporates potential trends in climate change at a spatial resolution useful at local management levels. It also provides this information on a short time scale, providing data for current management based on potential spread over the next few years while considering the changing climate.

The difficulty in obtaining the field location data for *P. montana* supports advocating a system to share invasive species data (such as Ricciardi et al. 2000, Simpson et al. 2006, Graham et al. 2007). Lack of data across the entire range of a species decreases the viability of predictive models as a subset may not encompass the entire range of climatic conditions where the species occurs. The complete range of climatic conditions of a species distribution is necessary to develop an accurate bioclimatic model (Thuiller et al. 2004).

The comparison of current and future predictions for *P. montana* provides useful information to resource managers. Because *P. montana* is an invasive species that continues to spread in the US, it may not yet be established in all predicted presence locations in the current map. Areas where the species' distribution is classified as increasing should be added to watch lists as early detection sites to monitor and prevent the future spread of the species. Areas where the species' distribution is stable should be targeted as control locations or monitored as early detection sites if the species has not yet arrived there. Areas the species' distribution is predicted as decreasing may have a lower priority, though if the species is currently established there, it may persist for decades (Stohlgren et al. 2008). These sites may also act as source populations for the spread of the species to more suitable areas. If the suitability of the climate is poor, the species may prove to be easier to locally eradicate than in other locations.

One criticism of climate driven changes in distribution models is the lack of parameters related to migration to determine how a species will move between its current and future range (Thuiller et al. 2008). However, these parameters are often difficult to obtain, and given the erratic pattern of long distance dispersal, fairly unpredictable (Clark et al. 2003). For invasive species, though, this issue seems less important given that an invasive species has demonstrated an ability to spread (i.e., migrate) by establishing a distribution in the invaded range. While the mechanisms of spread may not be understood, mobility is a requirement for a species to be a successful invader, and generally is not limited to biotic means that may also alter with climate change. Species adapt to new conditions.

These short-term predictions of climate change provide information to resource managers about how a species' potential distribution may alter over the span of their career. Many other invasive species can be modeled in this way. These models provide information at a temporal and a spatial scale that is useful to resource managers.

Caveats and planned improvements

The uncertainty associated with future climate projections is difficult to quantify. However, climate projections can be easily modified by adding new climate data yearly. Species distribution modeling also is an iterative process, requiring new models to be developed as new locations are added to the dataset. Lag effects in the invasion process (Crooks 2005) and persistence in suboptimal habitat can also affect species distribution models. There is no substitute for accurate and current field data.

We emphasized one species and one modeling approach here, but we advocate testing multiple species and multiple models (Evangelista et al.

2008). The choice of species distributions models is affected by species attributes (e.g., habitat generalists versus habitat specialists) and the clustering of reported presence points, absence points, and poorly surveyed areas.

Assessing current and future species distributions at local scales will demand fine resolution data and climate projections. Down-scaling and verifying climate data are in their infancy, but hey, you have to start somewhere!

Acknowledgements This work was partially supported by funding the United States Geological Survey through their research opportunities in global change science program. We would like to thank two reviewers for their comments. We would also like to thank the data contributors to the sites where we obtained field data for their willingness to share. To all we are grateful.

References

Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Model 162:211–232

Beaumont LJ, Pitman AJ, Poulsen M, Hughes L (2007) Where will species go? Incorporating new advances in climate modelling into projections of species distributions. Glob Change Biol 13:1368–1385. doi:10.1111/j.1365-2486.2007.01357.x

Clark JS, Lewis M, McLachlan JS, HilleRisLambers J (2003) Estimating population spread: what can we forecast and how well? Ecology 84:1979–1988. doi:10.1890/01-0618

Crooks JA (2005) Lag times and exotic species: the ecology and management of biological invasions in slow-motion. Ecoscience 12:316–329. doi:10.2980/i1195-6860-12-3-316.1

Daly C, Taylor GH, Gibson WP, Parzybok TW, Johnson GL, Pasteris PA (2000) High-quality spatial climate data sets for the United States and beyond. Trans ASAE 43:1957–1962

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Evangelista P, Kumar S, Stohlgren TJ, Jarnevich CS, Crall AW, Norman JB III, et al. (2008) Model selection for predicting a habitat generalist (Bromus tectorum) and a specialist (Tamarix chinensis) invasive plant species in Grand Staircase Escalante National Monument, Utah, USA. Divers Distrib 14:808–871

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49. doi:10.1017/S0376892997000088

Govindasamy B, Duffy PB, Coquard J (2003) High-resolution simulations of global climate, part 2: effects of increased greenhouse cases. Clim Dyn 21:391–404. doi:10.1007/s00382-003-0340-6

Graham J, Newman G, Jarnevich C, Shory R, Stohlgren TJ (2007) A global organism detection and monitoring system for non-native species. Ecol Inform 2:177–183

Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. Glob Change Biol 12:2272–2281

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978. doi:10.1002/joc.1276

Levinsky I, Skov F, Svenning JC, Rahbek C (2007) Potential impacts of climate change on the distributions and diversity patterns of European mammals. Biodivers Conserv 16:3803–3816. doi:10.1007/s10531-007-9181-7

Pearson RG, Thuiller W, Araujo MB, Martinez-Meyer E, Brotons L, McClean C, Miles L, Segurado P, Dawson TP, Lees DC (2006) Model-based uncertainty in species range prediction. J Biogeogr 33:1704–1711

Raxworthy CJ, Martinez-Meyer E, Horning N, Nussbaum RA, Schneider GE, Ortega-Huerta MA, Peterson AT (2003) Predicting distributions of known and unknown reptile species in Madagascar. Nature 426:837–841

Ricciardi A, Steiner WWM, Mack RN, Simberloff D (2000) Toward a global information system for invasive species. Bioscience 50:239–244

Simpson A, Sellers E, Grosse A, Xie Y (2006) Essential elements of online information networks on invasive alien species. Biol Invasions 8:1579–1587

Stohlgren TJ, Barnett DT, Jarnevich CS, Flather C, Kartesz J (2008) The myth of plant species saturation. Ecol Lett 11:313–326. doi:10.1111/j.1461-0248.2008.01153.x

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, de Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, van Jaarsveld AS, Midgley GF, Miles L, Ortega-Huerta MA, Peterson AT, Phillips OL, Williams SE (2004) Extinction risk from climate change. Nature 427:145–148

Thuiller W, Albert C, Araújo MB, Berry PM, Guisan G, Hickler T, et al. (2008) Predicting global change impacts on plant species distributions: future challenges. Perspect Plant Ecol Evol Syst 9:137–152

Thuiller W, Brotons L, Araujo MB, Lavorel S (2004) Effects of restricting environmental range of data to project current and future species distributions. Ecography 27:165–172. doi:10.1111/j.0906-7590.2004.03673.x