Reg Environ Change
DOI 10.1007/s10113-012-0389-z

ORIGINAL ARTICLE

Climate downscaling effects on predictive ecological models: a case study for threatened and endangered vertebrates in the southeastern United States

**David N. Bucklin · James I. Watling ·
Carolina Speroterra · Laura A. Brandt ·
Frank J. Mazzotti · Stephanie S. Romañach**

Received: 1 August 2012 / Accepted: 9 December 2012

Abstract High-resolution (downscaled) projections of future climate conditions are critical inputs to a wide variety of ecological and socioeconomic models and are created using numerous different approaches. Here, we conduct a sensitivity analysis of spatial predictions from climate envelope models for threatened and endangered vertebrates in the southeastern United States to determine whether two different downscaling approaches (with and without the use of a regional climate model) affect climate envelope model predictions when all other sources of variation are held constant. We found that prediction maps differed spatially between downscaling approaches and that the variation attributable to downscaling technique was comparable to variation between maps generated using different general circulation models (GCMs). Precipitation variables tended to show greater discrepancies between downscaling techniques than temperature variables, and for one GCM, there was evidence that more poorly resolved precipitation variables contributed relatively more to model uncertainty than more well-resolved variables. Our work suggests that ecological modelers requiring high-resolution climate projections should carefully consider the type of downscaling applied to the climate projections prior to their use in predictive ecological modeling. The uncertainty associated with alternative downscaling methods may rival that of other, more widely appreciated sources of variation, such as the general circulation model or emissions scenario with which future climate projections are created.

Keywords Climate change · Climate envelope model · Downscaling · Species distribution model · Florida · Endangered species

Introduction

Given the scale of climate change and sea-level rise threats to human and natural systems (IPCC 2007), forecasting the timing and intensity of climate change phenomena with high spatial precision is of paramount importance. In order to make global, coarse-scale general circulation model (GCM) projections useful for impact assessments, downscaling methods are often applied to produce future climate projections at relatively fine spatial and temporal scales (Wilby and Fowler 2010). Downscaling is achieved using either *statistical* or *dynamical* methods. Statistical methods translate GCM output to a finer spatial resolution by establishing statistical linkages between large (GCM) scale and local scale climate; this approach covers a wide breadth of techniques that can vary from simple to highly sophisticated (Fowler et al. 2007; Maraun et al. 2010). Dynamically downscaled approaches produce finer-scale output by embedding regional climate models (RCMs) within GCMs, allowing for better spatial resolution of atmospheric physics within the region (Maraun et al. 2010). Regional climate models are able to incorporate regional-scale processes, such as convective rainfall,

---

D. N. Bucklin (✉) · J. I. Watling · C. Speroterra ·
F. J. Mazzotti
Fort Lauderdale Research and Education Center, University of Florida, 3205 College Avenue, Fort Lauderdale, FL 33314, USA
e-mail: dbucklin@ufl.edu

L. A. Brandt
U.S. Fish and Wildlife Service, 3205 College Avenue,
Fort Lauderdale, FL 33314, USA

S. S. Romañach
U.S. Geological Survey, Southeast Ecological Science Center,
3205 College Avenue, Fort Lauderdale, FL 33314, USA

extreme weather events, surface-atmosphere interactions, orographic precipitation, and land use (Fowler et al. 2007; Stefanova et al. 2012). Also, they commonly produce output at short time steps (~30 min; Tabor and Williams 2010). As such, dynamical approaches are significantly more complex, time-consuming, and costly to implement than statistical approaches, but in most cases are thought to provide more realistic climate change scenarios (Fowler et al. 2007). Because of the substantial differences between approaches, downscaled projections can be highly variable, and precipitation has proven to be more difficult to model than temperature (Maraun et al. 2010).

Downscaled climate projections are an essential product for a variety of climate change impact modelers. In ecology, they are widely used in climate envelope models (CEMs), a type of predictive ecological model that relates species presences to a set of climatic variables, with the goal of capturing meaningful relationships between the species and its environment (Elith and Graham 2009). These relationships are often used to create predictions of species' climatic space or "niche" in novel areas or climates, under the assumption that the species–climate relationship remains consistent (Franklin 2009). (Note that throughout the manuscript, we refer to the output from CEMs as *predictions*, whereas outputs from climate models are referred to as *projections*.) Climate envelope models have been shown to be relatively good predictors of species distributions in historical and contemporary climate (Araújo et al. 2005; Mitikka et al. 2008; Rapacciuolo et al. 2012). However, their utility in predicting species' responses to future climate regimes, which is of particular interest for conservation management (Guisan and Thuiller 2005), is an area of concern and debate in the field (Austin 2002; Araújo and Guisan 2006; Chapman 2010; Real et al. 2010; Graham et al. 2011, Araújo and Peterson 2012, Rapacciuolo et al. 2012). CEM predictions are dependent on the particular climate scenarios to which they are applied, which depend on a set of conditions that (broadly) includes GCM, emissions scenario, geographic area, and time period. Because CEMs require high-resolution data sets if they are to be useful for site-specific conservation planning under climate change (Kremen et al. 2008; Tabor and Williams 2010), it is also important that we understand the effect that projections from different downscaling methods can have on CEMs (i.e., the degree of discrepancy between CEM predictions) when all other conditions are held constant.

To this end, we created CEM predictions for 2050 using both dynamically and statistically downscaled data sources for 14 species or subspecies, which have part (or all) of their current range in the southeastern United States (Table 1). Species selection was based on a two important considerations: (1) We wanted to select a group of species with varied biological traits (e.g., taxonomies, range sizes, physiologies, habitat associations) to thoroughly analyze the effect of climate model on CEMs, and (2) we wanted the group to be composed of species relevant to conservation planning and management in the southeastern United States (also the spatial domain for the RCM we used in this study). To achieve this, we selected a group of 25 terrestrial vertebrates on the U.S. federally threatened or endangered lists occurring in the state of Florida. Among this group of species, some are expected to be directly impacted by climate change because of physiological tolerances (e.g., ectothermic reptiles living near range limits, wading birds dependent on wet-season rainfall), while for others the impact is less certain, potentially involving interacting abiotic and biotic factors (e.g., see Araújo and Luoto 2007; Cahill et al. 2012). Despite this uncertainty, the conservation statuses of threatened and endangered species make them high-priority candidates for the use of climate envelope models as an initial assessment of vulnerability, helping aid in the development and prioritization of policy responses to climate change (Povilitis and Suckling 2010; Rowland et al. 2011). As such, our modeling exercise provides a real-life application of climate envelope modeling for biological planning. Finally, because endangered, rare species are more likely to have smaller ranges and be specialized to their habitats (Lomba et al. 2010), they are likely to occur in areas with unique climate conditions, and the processes creating these conditions may be more strongly represented by a regional climate model.

In the process of creating predictions from CEMs, ecological modelers are faced with an abundance of downscaled data choices for depicting future climate. Statistically downscaled products (hereafter referred to as "non-RCM") are widely available, can be created relatively quickly for multiple GCM and emissions scenarios (including by the end user using a variety of software packages), and are not specifically limited in geographic scope. Dynamically downscaled products employing RCMs are more resource-intensive and are created in order to provide a realistic portrayal of fine-scale atmospheric physical processes for a specific region. If regional patterns and processes are effectively modeled, RCMs may also lead to more realistic future climate projections. As such, we find it useful to compare outputs between CEM predictions produced with non-RCM versus RCM climate projections. We know that future predictions cannot be tested in a way which would allow us to determine whether certain climate inputs lead to better performing models (Graham et al. 2011). Instead of a performance test, we test for the magnitude of discrepancy between CEM predictions informed by non-RCM versus RCM projections, using two different GCMs in the A2 emissions scenario for

Table 1 List of species, number of occurrences, and variables selected for model construction

Common name	Scientific name	Number of occurrences (occupied cells)	Variables used in model
Mammals			
Florida panther	*Puma concolor coryi*	109	Temperature: July, October Precipitation: February, April, May, August, September, October
Birds			
Cape Sable seaside sparrow	*Ammodramus maritimus mirabilis*	17	Temperature: November Precipitation: February, May, June, July, September, October
Florida grasshopper sparrow	*Ammodramus savannarum floridanus*	25	Temperature: September Precipitation: February, May, June, October
Florida scrub jay	*Aphelocoma coerulescens*	192	Temperature: July, November Precipitation: January, March, May, Jun, August, October
Piping plover	*Charadrius melodus*	782	Temperature: July, October Precipitation: February, May, September, October
Wood stork	*Mycteria americana*	1,550	Temperature: July, October Precipitation: March, May, July, September, October
Audubon's crested caracara	*Polyborus plancus audubonii*	159	Temperature: August, November Precipitation: April, May, September, December
Everglade snail kite	*Rostrhamus sociabilis plumbeus*	92	Temperature: July, September, November Precipitation: February, April, May, September, October
Red-cockaded woodpecker	*Picoides borealis*	515	Temperature: August Precipitation: March, May, June, July, September
Reptiles			
American crocodile	*Crocodylus acutus*	112	Temperature: May, December Precipitation: January, April, May, July, September, October
Bluetail mole skink	*Eumeces egregius lividus*	20	Temperature: October Precipitation: January, April, May, June, August, September, October, November
Sand skink	*Neoseps reynoldsi*	18	Temperature: September, December Precipitation: February, April, May, June, August
Eastern indigo snake	*Drymarchon corais couperi*	278	Temperature: July, September Precipitation: February, April, May, August, September, October
Amphibians			
Flatwoods salamander	*Ambystoma cingulatum*	31	Temperature: October Precipitation: February, April, May, June, July, August, September, October, November

a group of vertebrate species. For ecological modelers and conservation planners with limited funds that are faced with assessing and planning for the impacts of climate change on biodiversity, we hope to provide specific insight into whether the choice between downscaled climate inputs is significant. If there is a negligible difference between

CEM predictions using the two different downscaling techniques, there may not be strong justification for the adoption of RCM projections for ecological modeling, especially when non-RCM alternatives with fewer limitations are abundant (e.g., Ramirez and Jarvis 2008; Tabor and Williams 2010).

Materials and methods

Downscaled climate projections

For our "non-RCM" downscaled climate projections, we acquired climate projections from the University of Wisconsin–Madison Center for Climatic Research (Tabor and Williams 2010). Downscaling for these data sets was performed using the *change-factor* technique, a simple, fast method which was implemented by calculating the original-resolution monthly anomalies from the GCM outputs (future projection—present simulation), downscaling these anomalies using a spline interpolation technique, and adding the anomalies to the Climate Research Unit (CRU; New et al. 2002) 1961–1990 climatology (Tabor and Williams 2010). They produced projections at the 10 arc-minute resolution ($\sim$17 km^2 at the equator) for 24 GCMs, three IPCC emissions scenarios, and two future time periods (Tabor and Williams 2010). In our study, projections from two GCMs for the A2 scenario are used—the National Center for Atmospheric Research's Community Climate System Model, version 3.0 (hereafter CCSM3), and the UK Met Office's Hadley Centre for Climate Prediction and Research's HadCM3. We acquired projections of mean monthly temperature and precipitation for these two GCMs for the 2050 time period (representing 2041–2060).

Dynamically downscaled climate projections (RCM) were acquired from the Center for Ocean–Atmospheric Prediction Studies (COAPS) at Florida State University, part of the COAPS Land–Atmosphere Regional Ensemble Climate Change Experiment for the Southeast United States at a 10-km resolution (http://coaps.fsu.edu/CLARReS10/index.shtml). They employed the FSU-FCI RSM dynamical model (Kanamitsu et al. 2010) to create projections for three GCMs using the A2 emissions scenario, with three different vegetation scenarios used to inform the models (Stefanova et al. 2012). For our study, projections created using the "current" vegetation scenario were acquired for the CCSM3 and HadCM3 GCMs. Projections were acquired as daily values of temperature (minimum and maximum) and precipitation for future (2039–2070) and historical (1968–2000) simulations. Values from 2041 to 2060 were extracted to match the non-RCM data temporal period, and monthly averages were calculated. Mean monthly temperature was then calculated as the average of maximum and minimum monthly temperature. Following the creation of these 24 monthly average temperature and precipitation projections for the two GCMs, we needed to ensure that they were de-biased (i.e., climate model bias was removed). Daily precipitation values were already bias-corrected by the COAPS group using a quantile-matching approach (Wood et al. 2002). This process creates a cumulative distribution function (CDF) of future daily precipitation values on a month-by-month basis, which is then adjusted to have the same mean and variance as an observation (twentieth century) data set. At this point, the future data set is linearly interpolated to the CDF of the twentieth-century data to create the final bias-corrected values (Stefanova et al. 2012). Temperature values were not available in a bias-corrected format at the time of acquisition. To achieve this, we performed a *delta-method* bias correction, where monthly mean temperature in the historical simulation was subtracted from monthly mean temperature for the future time period to produce anomalies. These anomalies were then added to mean monthly temperatures of a contemporary "consensus" data set, which we created by averaging the CRU and Worldclim (Hijmans et al. 2005) contemporary climatologies.

A geographic projection was applied to both the RCM and non-RCM data using the WGS 1984 datum, and non-RCM data were resampled to a $\sim$10-km resolution to match the RCM projections. Since the RCM data were produced for the southeastern United States only (24°N to 36°N south to north, -90°E to -76°E west to east), the non-RCM data were clipped to match this domain. To describe spatial variation between climate projections made using different downscaling techniques, we calculated spatial correlations between RCM and non-RCM projections for each climate variable (average monthly temperature and precipitation). To calculate a spatial correlation, each cell in a prediction map is paired with the corresponding cell in another, and Pearson's correlation is calculated across all cells in the two maps (Syphard and Franklin 2009).

Species models

Occurrences for the original group of 25 terrestrial vertebrates were obtained from several online databases and the literature sources and compiled for either the species or subspecies taxa, depending on the designation on the U.S. federally threatened and endangered species list for Florida. From the group of 25, we removed species occurring only in the Florida Keys ($n = 6$), as this area was not represented in the RCM projections. We also removed species with fewer than 15 occurrences ($n = 5$), producing a final group of 14 "focal" species across four taxa—one mammal, eight birds, four reptiles, and one amphibian (Table 1). Determining an appropriate geographic background (spatial domain) from which to draw environmental data is an active area of research in species modeling, and it has been shown that this factor can significantly impact model parameterization and performance (VanDerWal et al. 2009; Elith et al. 2010; Barve et al. 2011; Acevedo et al. 2012). In order to tailor a biologically informed model domain for each focal species, we employed an approach that utilized "target-group" species occurrences (Phillips et al. 2009). Briefly, for each focal species, the occurrences for at least three phylogenetically related and ecologically similar species (the target-group) were gathered from an online database (www.gbif.org). The 100 % minimum convex polygon circumscribing these occurrences was then generated and represented the model domain for the focal species. As we lacked true absence data, we generated 10,000 "pseudo-absences" within the model domain, not allowing absences in cells where the focal species was present.

Model covariates were selected individually for each species from an original pool of 24 monthly mean temperature and precipitation variables, using the CRU contemporary data set. Using ecological-niche factor analysis (ENFA) in the Biomapper program (Hirzel et al. 2002), we analyzed the cluster diagram of variable correlations for each species and removed all but one of the highly correlated ($r > 0.85$) variables from each cluster, retaining the variable which was most associated with species' presence (based on its influence on marginality; see Hirzel et al. 2002). Using this method produced sets of moderately correlated variables including 1–3 mean monthly temperature and 4–9 mean monthly precipitation covariates (Table 1). Precipitation variables were more commonly selected due to precipitation's inherent spatiotemporal variability and nonlinear nature (resulting in lower correlation coefficients between monthly averages). Contemporary climate data for model construction were drawn from the consensus data set (described above). Values for each of the selected variables were extracted for the focal species' occurrences and pseudo-absences, only allowing one species occurrence per 10 arc-min grid cell ($\sim$17 km^2).

Models were created for all species using the random forests algorithm implemented in R 2.15.0 (R Development Core Team 2012) using the randomForest package (Liaw and Wiener 2002). Random forests is a machine-learning algorithm that can handle a large number of covariates and has been shown to be resistant to overfitting; as such, it has also shown good performance in species distribution models, even for species with limited observations (Cutler et al. 2007). Assessing model performance with multiple metrics is an important part of the modeling process (Elith and Graham 2009). In order to test the performance of our contemporary models, we used a cross-validation procedure to assess the predictive performance of the models in which we randomly sampled the training and testing subsets (using a 75–25 % split, respectively) for 100 models. Model performance was measured during the cross-validation using three metrics—area under the receiver–operator curve (AUC), Cohen's kappa (hereafter kappa), and the true skill statistic (TSS). The AUC is a threshold-independent metric that measures the tendency of a randomly occupied cell to have a higher predicted probability of suitable climate than a pseudo-absence (Fielding and Bell 1997). Kappa is a measure of agreement between predicted and observed presence–absence, which corrects for agreement resulting from random chance (Fielding and Bell 1997). It requires that a threshold value be applied to the model's probabilistic output. We applied two thresholds to model output in this study. Our primary threshold is the value which maximized kappa (hereafter "max kappa"; Freeman and Moisen 2008). This value was determined by initial model runs for each species where we applied the threshold at each 0.01 unit change for five model randomizations. After consideration of a prevalence-based threshold as an alternate, we instead decided to use a fixed threshold of 0.1, as prevalence was very low for most of our study species (<0.1 for 13 out of 14). In all cases, the fixed threshold was lower than the max kappa threshold. The max kappa threshold was also applied for the calculation of TSS, which is a measure of [fraction correctly classified presences (*sensitivity*) + fraction correctly classified absences (*specificity*) − 1] (Allouche et al. 2006). One model was created using 100 % of data ("baseline") to provide the spatial prediction for the contemporary model, the classification for the future models, and a measure of permutation importance for each variable in model construction.

Testing the effects of different downscaling approaches on species models

The baseline random forest models for each species were applied to the climate projections to produce four future CEM predictions (non-RCM CCSM3, non-RCM HadCM3, RCM CCSM3, and RCM HadCM3). To test for discrepancies between prediction maps attributable to different downscaling techniques, we calculated spatial correlations between the RCM and non-RCM prediction maps for each species, with separate analyses for each GCM. Because abundant evidence suggests that using different GCMs introduces variation into predictions from ecological models (Real et al. 2010), we compared discrepancies in CEM predictions attributable to the different downscaling techniques to discrepancies associated with different GCMs by calculating spatial correlations between CCSM3 and HadCM3 prediction maps for each species, with separate analyses for the two downscaling approaches.

We also wanted to describe how CEM prediction maps using RCM versus non-RCM data varied when a threshold was applied to the probabilistic value, a common practice in species' distribution modeling that produces a binary "suitable/unsuitable" climate map (Guisan and Thuiller 2005) We did this by calculating a "spatial coincidence" metric for each species' future spatial prediction maps for the two GCMs. It is a measurement of the area (A) where non-RCM and RCM predictions coincide (A_{overlap}), where the overlap area is divided by the total area predicted as suitable in either of the maps [$A_{\text{overlap}}/([A_{\text{non-RCM}} + A_{\text{RCM}}] - A_{\text{overlap}})$]. This provided us with a test of the spatial agreement of RCM and non-RCM prediction maps which was independent of the spatial correlation metric. For 10 species, we applied the threshold at max kappa. For the other six (all with very limited ranges in the study area), the max kappa threshold was higher than any pixel's predicted value for at least one of the prediction maps; for these species, the fixed threshold of 0.1 was used.

Finally, we hypothesized that differences between RCM and non-RCM prediction maps (the effect of alternative downscaling techniques on species models) may be particularly great when the models are strongly influenced by monthly climate variables that are poorly resolved (i.e., described very differently using the two downscaling approaches). Although it would be ideal to test this idea on a species-by-species basis, some climate variables were either not included in any of the species' models or only included in a few models. Thus, the differential representation of variables in the species' models precludes a robust assessment of variable contribution to model predictions. Instead, we ask a more general question: Do more poorly resolved climate variables contribute relatively more to model predictions? If so, this would suggest that prediction maps may be particularly influenced by the most poorly resolved climate variables. We identified the most poorly resolved climate variables using the spatial correlation between RCM and non-RCM climate projections as described above. We estimated the importance of all climate variables in each species' random forest model by calculating the relative mean decrease in accuracy attributed to that variable (Cutler et al. 2007) in the baseline model run and averaging this metric of variable importance across all species. If more poorly resolved climate variables contribute relatively more to model predictions, there should be a negative relationship between spatial correlation and mean variable importance. We tested for that relationship using linear regression, with a separate model for each GCM.

Results

Spatial correlations between projected average monthly temperatures for RCM versus non-RCM variables (mean = 0.974, range = 0.896–0.998) were higher and less variable than those for average monthly precipitation projections (mean = 0.740, range = 0.160–0.920; Table 2). Correlations between projections for July (Fig. 1) and October precipitation from the CCSM3 GCM were particularly low ($r = 0.160$ and $r = 0.254$, respectively); correlations for all other CCSM3 variables (temperature and precipitation) were greater than 0.500. For the HadCM3 model, only correlation for May precipitation ($r = 0.461$) was below 0.500.

Results from the cross-validation for contemporary CEMs showed good performance of all models according

Table 2 Spatial correlation between climate projections made using two different downscaling techniques (with and without the use of a regional climate model) for the HadCM3 and CCSM3 general circulation models

Month	Spatial correlation (RCM vs. non-RCM)			
	HadCM3		CCSM3	
	Temperature	Precipitation	Temperature	Precipitation
January	0.996	0.821	0.997	0.799
February	0.998	0.903	0.996	0.877
March	0.996	0.837	0.995	0.920
April	0.991	0.793	0.992	0.898
May	0.987	0.461	0.987	0.727
June	0.905	0.862	0.934	0.687
July	0.897	0.655	0.896	0.160
August	0.943	0.856	0.949	0.651
September	0.973	0.749	0.978	0.795
October	0.987	0.682	0.993	0.254
November	0.993	0.855	0.996	0.764
December	0.998	0.909	0.998	0.852
Mean ± 1 SD	0.972 ± 0.037	0.782 ± 0.129	0.976 ± 0.033	0.699 ± 0.245

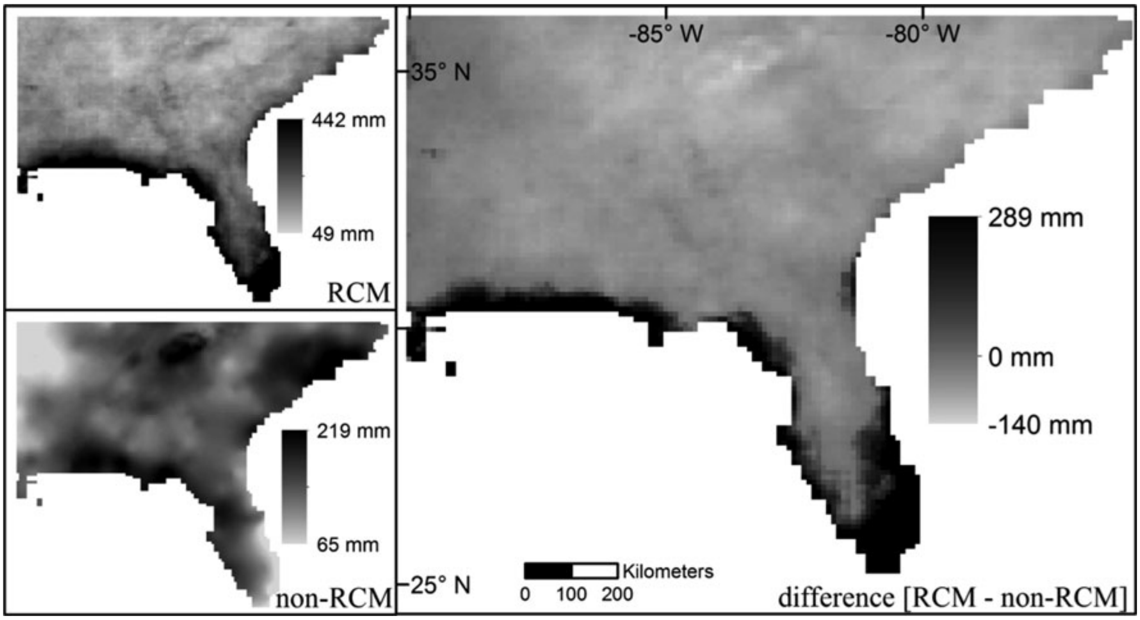


Fig. 1 Mean total July precipitation for 2041–2060 according to RCM and non-RCM projections for the NCAR CCSM3 GCM, illustrating the significant discrepancies that can arise with different downscaling approaches

to the AUC statistic (mean = 0.988), indicating that models were able to discriminate among occupied sites and pseudo-absences (Manel et al. 2001). Classification performance metrics using the max kappa threshold were more variable (Table 3), with several models showing decreased performance due to low sensitivity (ability to classify test presences). In general, however, the good performance results using random forests were broadly congruent with previous models run for a similar set of species using different climate data and variables (Watling et al. 2012).

Spatial correlations (r) between CEM prediction maps using the two downscaling techniques (holding GCM constant) were moderate (Table 4; mean = 0.632, range = 0.074–0.865 for the HadCM3 GCM and mean = 0.530, range = 0.347–0.844 for the CCSM3 GCM; see also Fig. 2 for the predictions using HadCM3). By way of comparison, spatial correlations between maps created using the different downscaling techniques were very similar to those between GCM projections, where downscaling method was held constant (Table 4; mean = 0.667, range = 0.384–0.923 for the non-RCM data; and mean = 0.519, range = −0.051–0.786 for the RCM data).

Mean spatial coincidence between RCM and non-RCM predictions for the 28 combinations (14 species × 2 GCMs) was low (mean = 0.365) and varied slightly by GCM (CCSM3 mean = 0.333; HadCM3 mean = 0.397). As our objective is not to test the variability between GCM projections, the spatial coincidence metric was only calculated for the comparison of RCM versus non-RCM data sets and is presented as the species' mean value for the two GCMs (Table 4).

Mean variable importance scores ranged from 0.000 to 0.143. For the CCSM3 GCM, there was no evidence to suggest that poorly resolved climate variables were particularly important in species models ($F_{1,10} = 2.14$, $p = 0.17$; and $F_{1,10} = 0.36$, $p = 0.56$ for temperature and precipitation variables, respectively). For the HadCM3 GCM, there was no association between spatial correlations and variable importance for temperature ($F_{1,10} = 1.36$, $p = 0.27$), but the relationship was significant for precipitation ($F_{1,10} = 8.05$, $p = 0.02$), indicating a tendency for more poorly resolved precipitation variables to contribute relatively greatly to model predictions.

Discussion

In this study, we showed that for a set of 14 threatened and endangered vertebrates, climate envelope model predictions created with dynamically downscaled data (RCM) were only moderately spatially correlated (mean $r = 0.581$) with predictions created with non-dynamically downscaled (non-RCM) data and that the magnitude of difference was virtually the same as that between predictions created using different GCMs (mean $r = 0.593$). There were considerable differences between RCM and non-RCM projections for several climate variables, and there was evidence that the most discrepant precipitation variables contributed disproportionately to CEM predictions. Our results show that just as discrepancies can arise between large-scale GCMs (and in turn affect CEM predictions; Real et al. 2010), equally important differences

Table 3 Cross-validation results for species models based on 100 random partitions of the species' occurrences and pseudo-absences (mean value ± one SD)

Common name	AUC	Cohen's kappa	TSS	Sensitivity	Specificity
Florida panther	0.994 ± 0.009	0.918 ± 0.032	0.958 ± 0.038	0.959 ± 0.038	0.999 ± 0.001
Cape Sable seaside sparrow	1.00 ± 0.000	0.786 ± 0.173	0.732 ± 0.225	0.733 ± 0.225	1.00 ± 0.000
Florida grasshopper sparrow	0.996 ± 0.015	0.707 ± 0.144	0.676 ± 0.192	0.677 ± 0.192	0.999 ± 0.001
Florida scrub jay	0.999 ± 0.000	0.902 ± 0.028	0.966 ± 0.029	0.969 ± 0.029	0.997 ± 0.001
Piping plover	0.962 ± 0.006	0.655 ± 0.027	0.63 ± 0.034	0.651 ± 0.035	0.979 ± 0.004
Wood stork	0.964 ± 0.005	0.751 ± 0.017	0.749 ± 0.022	0.783 ± 0.022	0.967 ± 0.003
Audubon's crested caracara	0.999 ± 0.002	0.914 ± 0.03	0.972 ± 0.031	0.975 ± 0.031	0.998 ± 0.001
Everglade snail kite	0.999 ± 0.001	0.836 ± 0.046	0.909 ± 0.061	0.912 ± 0.061	0.998 ± 0.001
Red-cockaded woodpecker	0.984 ± 0.005	0.718 ± 0.027	0.782 ± 0.038	0.802 ± 0.039	0.98 ± 0.003
American crocodile	0.960 ± 0.017	0.366 ± 0.062	0.424 ± 0.092	0.434 ± 0.092	0.99 ± 0.002
Bluetail mole skink	0.998 ± 0.01	0.747 ± 0.153	0.718 ± 0.209	0.718 ± 0.209	1.00 ± 0.000
Sand skink	0.993 ± 0.025	0.472 ± 0.153	0.653 ± 0.253	0.655 ± 0.253	0.998 ± 0.001
Eastern indigo snake	0.997 ± 0.001	0.861 ± 0.023	0.95 ± 0.022	0.957 ± 0.022	0.993 ± 0.002
Flatwoods salamander	0.993 ± 0.01	0.358 ± 0.134	0.327 ± 0.139	0.329 ± 0.14	0.998 ± 0.001
Mean ± 1SD	0.988 ± 0.014	0.714 ± 0.183	0.746 ± 0.193	0.754 ± 0.192	0.993 ± 0.010

Threshold-dependent metrics (kappa, TSS, sensitivity, and specificity) are calculated using the max kappa threshold

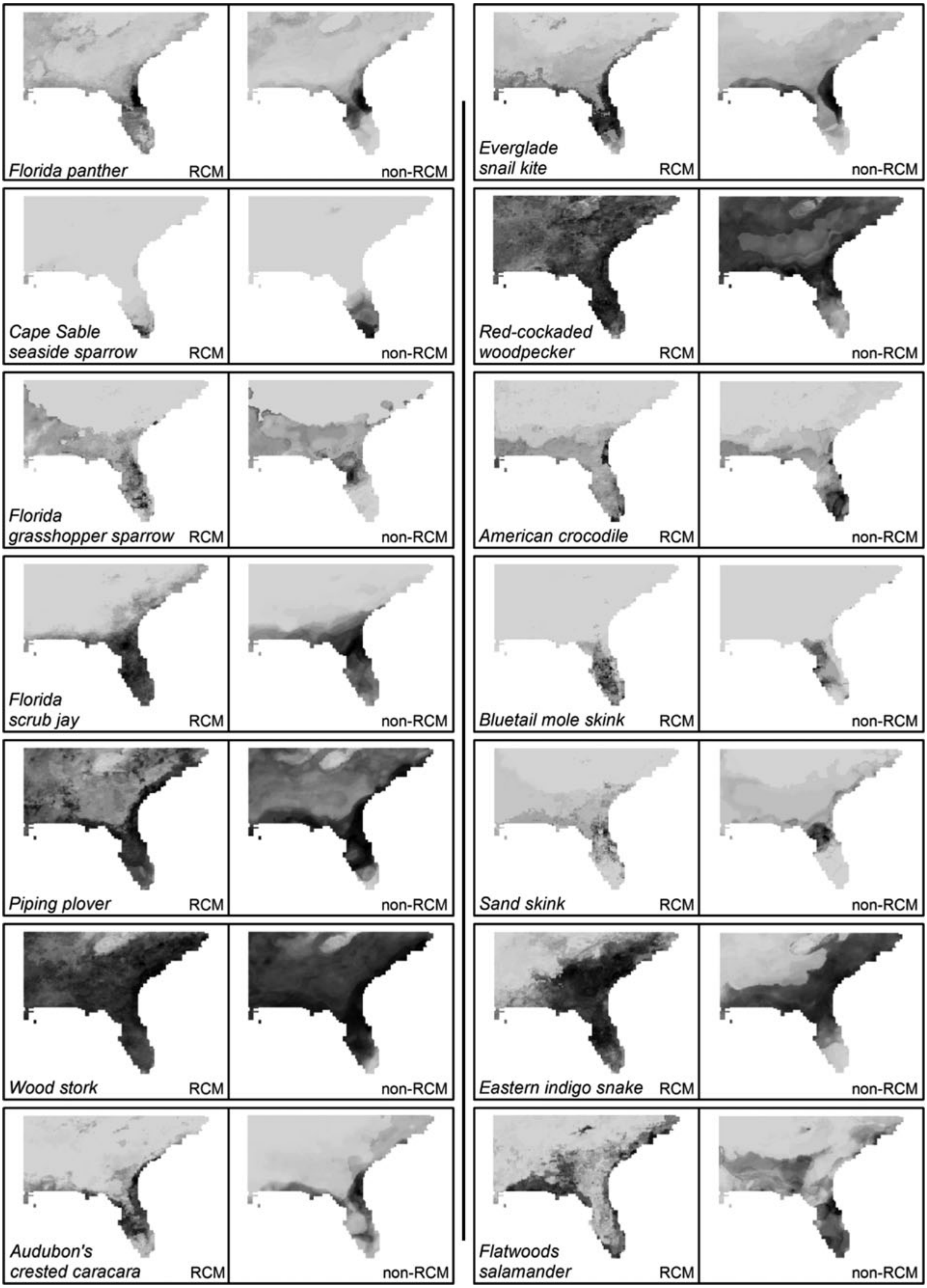
Table 4 Spatial correlation between GCMs (varying downscaling method) and downscaling method (varying GCMs), and spatial coincidence (calculated between downscaling methods; mean for the two GCMs) for all 14 study species climate envelope model predictions for 2041–2060

Common name	Spatial correlation (downscaling procedure)		Spatial correlation (GCMs)		Spatial coincidence (downscaling procedure)
	CCSM3	HadCM3	RCM	non-RCM	(mean of both GCMs)
Florida panther	0.512	0.797	0.468	0.660	0.364
Cape Sable seaside sparrow	0.489	0.764	0.556	0.821	*0.287*
Florida grasshopper sparrow	0.351	0.486	0.243	0.563	*0.218*
Florida scrub jay	0.844	0.865	0.786	0.923	0.698
Piping plover	0.456	0.665	0.521	0.794	0.394
Wood stork	0.594	0.702	0.650	0.759	0.862
Audubon's crested caracara	0.442	0.697	0.422	0.683	0.341
Everglade snail kite	0.562	0.700	0.665	0.568	0.147
Red-cockaded woodpecker	0.347	0.074	−0.051	0.384	0.644
American crocodile	0.439	0.705	0.746	0.662	*0.062*
Bluetail mole skink	0.664	0.630	0.595	0.623	*0.208*
Sand skink	0.541	0.565	0.443	0.632	0.086
Eastern indigo snake	0.647	0.693	0.685	0.713	0.584
Flatwoods salamander	0.524	0.502	0.538	0.559	0.212
Mean	0.530	0.632	0.519	0.667	0.365

For the spatial coincidence metric, *italics* indicate a species for which the 0.1 threshold was used

between climate projections can arise through the downscaling process of a single GCM. When implemented in CEMs, climate variables from these projections produced moderately to drastically different predictions, depending on the species and variables used. These results highlight the need for caution when interpreting CEM predictive maps for future climate regimes, particularly binary suitable/unsuitable climate maps (often termed "presence/absence"), a point reinforced by low spatial coincidence scores between CEM predictions in this study (mean = 0.365).

Fig. 2 Climate envelope model prediction maps for 14 terrestrial ▶ vertebrate species in the southeastern United States using the HadCM3 GCM and A2 emissions scenario. All maps represent predictions of climate suitability based on average conditions during the 20-year period, 2041–2060. *Darker shades* indicate higher climate suitability


Florida panther
RCM
non-RCM
Cape Sable seaside sparrow
RCM
non-RCM
Florida grasshopper sparrow
RCM
non-RCM
Florida scrub jay
RCM
non-RCM
Piping plover
RCM
non-RCM
Wood stork
RCM
non-RCM
Audubon's crested caracara
RCM
non-RCM
Everglade snail kite
RCM
non-RCM
Red-cockaded woodpecker
RCM
non-RCM
American crocodile
RCM
non-RCM
Bluetail mole skink
RCM
non-RCM
Sand skink
RCM
non-RCM
Eastern indigo snake
RCM
non-RCM
Flatwoods salamander
RCM
non-RCM

Precipitation projections showed much greater dissimilarity than temperature between RCM and non-RCM data, and we showed that this uncertainty can be transferred to CEM predictions. For the HadCM3 GCM, spatial predictions from CEMs constructed using different downscaling techniques were more discrepant when poorly resolved precipitation variables contributed greatly to CEM predictions. Precipitation is very sensitive to climate model parameterization and is highly spatially variable, two issues that can contribute to large discrepancies between different climate model (both GCM and RCM) projections (Maraun et al. 2010). While RCMs are initially driven by GCM projections, they have been shown to contribute added value in predicting precipitation (Maraun et al. 2010). However, RCMs generally show greater skill in the prediction of winter precipitation, and for moderate (as opposed to very heavy, particularly convective) precipitation (Maraun et al. 2010). Indeed, summer precipitation is difficult to model in the southeastern United States (and particularly peninsular Florida, where it is the wet season) due to the regional climatological factors, including the significant contribution of convective rainfall, sea breeze, thunderstorms, and squall lines (Stefanova et al. 2012). Regional climate models, which are equipped to deal with these regional-scale phenomena, may be expected to differentiate most greatly from the non-RCM-informed predictions for high-precipitation months, or dry-to-wet season transition months (as was the case here).

While this study provided a direct comparison of CEM predictions created with RCM versus non-RCM climate projections, there are certain limitations that need to be considered. The non-RCM (statistical) data used here were derived using a simple (change-factor) technique; undoubtedly, alternative statistically downscaled projections using more complex regional interactions may produce different results (e.g., Hellström et al. 2001). Our variable selection methodology, which favored the selection of precipitation variables relative to temperature, may have acted to heighten the dissimilarity between CEM predictions. Also, due to data limitations, we only compared CEMs created with projections from two GCMs within one emissions scenario, which limited our ability to provide a more robust assessment of the effect of downscaling on CEM predictions. Finally, the geographic domain of our predictions was limited by the CLAREnCE10 RCM domain (Fig. 1). While this area is probably adequate for modeling changes in suitable climate for most of the study species due to their currently limited ranges, several (including the red-cockaded woodpecker, wood stork, piping plover, and American crocodile) have ranges that extend well beyond the southeastern United States, and the CEMs were built using data from outside this range (i.e., some occurrences fell outside of the southeastern United States). While this does not produce inherently "incorrect" prediction maps, it does place a significant predefined constraint on their interpretation and usefulness for conservation planning.

In summary, CEM predictions created with dynamically versus non-dynamically downscaled data were only moderately correlated, and correlation varied by GCM. A similar moderate correlation was seen between GCMs, which also varied by downscaling approach. This suggests that, similar to the selection of GCM and emissions scenario (Real et al. 2010), choices regarding downscaling approach of the climate data for predictive ecological modeling are not insignificant. While choosing an RCM projection may seem preferable for species modeling, there are many considerations that should be made prior to their use. Due to the computationally intensive nature and expensive implementation of RCMs, fewer [GCM x emissions scenario] combinations are likely to be dynamically downscaled (Fowler et al. 2007), reducing the ability to bracket predictions made with RCMs with consensus or "ensemble" (multiple GCM x emission scenario) approaches. Likewise, RCMs introduce their own biases and can produce variable climate projections within one GCM (Fowler et al. 2007), but alternative RCMs are less likely to be available at fine spatial resolutions for most areas. In addition, RCMs by definition are sub-global and do not represent climate for very large geographic areas, reducing their utility for wide-ranging species (as was exemplified with several species in this study). Our results also suggest that precipitation variables, particularly from the wet season, have greater uncertainties which can sometimes contribute to more discrepant CEM predictions. When these variables are significant determinants in a region's hydrology, and in turn, the ability of a species to occupy the area, they cannot be ignored during variable selection. Despite the limitations described above, RCM-informed projections may be preferable for species dependent on specific hydrological regimes, due to the RCM's ability to provide a more accurate representation of the spatiotemporal structure of fine-scale phenomena affecting precipitation regimes (Stefanova et al. 2012). Clearly, trade-offs between uncertainty, ensemble approaches, spatial scale, and geographic domain need to be weighed prior to the selection of the climate projection product(s) that is most appropriate for each species. As the impact of climate change will likely produce species-specific effects (Real et al. 2010), selection of climate projections based on justifiable species-specific considerations will likely produce more informative models.

Acknowledgments We thank Lydia Stefanova at COAPS for assisting us with the dynamically downscaled (CLAREnCE10) projections and three anonymous reviewers for their valuable comments and suggestions. Funding for this work was provided by the U.S. Fish and Wildlife Service, the National Park Service (Everglades and Dry Tortugas National Parks) through the South Florida and Caribbean Cooperative Ecosystem Studies Unit, and the U.S. Geological Survey (Greater Everglades Priority Ecosystems Science). Views expressed here do not necessarily represent the views of the U.S. Fish and Wildlife Service. Use of trade, product, or firm names does not imply endorsement by the U.S. Government.

References

Acevedo P, Jiménez-Valverde A, Lobo JM, Real R (2012) Delimiting the geographical background in species distribution modelling. J Biogeogr 39:1383–1390

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol 43:1223–1232

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modeling. J Biogeogr 33:1677–1688

Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. Global Ecol Biogeogr 16:743–753

Araújo MB, Peterson AT (2012) Uses and misuses of bioclimatic envelope modeling. Ecology 93:1527–1539

Araújo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species-climate impact models under climate change. Glob Change Biol 11:1504–1513

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modeling. Ecol Model 157:101–118

Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol Model 222:1810–1819

Cahill AE, Aiello-Lammens ME, Fisher-Reid MC, Hua X, Karanewsky CJ, Ryu HY, Sbeglia GC, Spagnolo F, Waldron JB, Warsi O, Wiens JJ (2012) How does climate change cause extinction? Proc R Soc B, published online 17 Oct 2012. doi:10.1098/rspb.2012.1890

Chapman DS (2010) Weak climatic associations among British plant distributions. Glob Ecol Biogeogr 19:831–841

Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783–2792

Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32:66–77

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods Ecol Evol 1:330–340

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

Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int J Climatol 27:1547–1578

Franklin J (2009) Mapping species distributions: spatial inference and prediction. Cambridge University Press, New York

Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. Ecol Model 217:48–58

Graham CH, Loiselle BA, Velásquez-Tibatá J, Cuesta FC (2011) Species distribution modeling and the challenge of predicting future distributions. In: Herzog SK, Martínez R, Jørgensen PM, Tiessen H (ed) Climate change and biodiversity in the Tropical Andes, Inter-American Institute for Global Change Research (IAI) and Scientific Committee on Problems of the Environment. São José dos Campos, Brazil, pp 295–310. http://www.iai.int

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009

Hellström C, Chen D, Achberger C, Räisänen J (2001) Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. Clim Res 19:45–55

Hijmans RS, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution climate surfaces for global land areas. Int J Climatol 25:1965–1978

Hirzel AH, Hausser J, Chessel JD, Perrin N (2002) Ecological niche-factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83:2027–2036

Intergovernmental Panel on Climate Change (2007) Climate change 2007: the physical science basis. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Contribution of working group 1 to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 433–498

Kanamitsu M, Yoshimura K, Yhang Y-B, Hong S-Y (2010) Errors of interannual variability and trend in dynamical downscaling of reanalysis. J Geoph Res 115:D17115

Kremen C et al (2008) Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. Science 320:222–226

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3):18–22

Lomba A, Pellissier L, Randin C, Vicente J, Moreira F, Honrado J, Guisan A (2010) Overcoming the rare species modelling paradox: a novel hierarchical framework applied to an Iberian endemic plant. Biol Conserv 143:2647–2657

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

Maraun D, Wetterhall H, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. Rev Geophys 48:RG3003

Mitikka V, Heikkinen RK, Luoto M, Araújo MB, Saarinen K, Poyry J, Fronzek S (2008) Predicting range expansion of the map butterfly in northern Europe using bioclimatic models. Biodivers Conserv 17:623–641

New M, Lister D, Hulme M, Makin I (2002) A high-resolution data set of surface climate over global land areas. Clim Res 21:1–25

Phillips SJ, Dudík M, Elith J, Graham CM, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197

Povilitis A, Suckling K (2010) Addressing climate change threats to endangered species in U.S. recovery plans. Conserv Biol 24:372–376

Ramirez J, Jarvis A (2008) High resolution statistically downscaled future climate surfaces, International Center for Tropical Agriculture (CIAT); CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Cali, Colombia. http://www.ccafs-climate.org/statistical_downscaling_delta

Rapacciuolo G, Roy DB, Gillings S, Fox R, Walker K, Purvis A (2012) Climatic Associations of British species distributions show good transferability in time but low predictive accuracy for range change. PLoS ONE 7:1–7

Real R, Márquez AL, Olivero J, Estrada A (2010) Species distribution models in climate change scenarios are still not useful for informing policy planning: an uncertainty assessment using fuzzy logic. Ecography 33:304–314

Rowland EL, Davison JE, Graumlich LJ (2011) Approaches to evaluating climate change impacts on species: a guide to initiating the adaptation planning process. Environ Manage 47:322–337

R Development Core Team (2012) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna. www.R-project.org

Stefanova L, Misra V, Chan S, Griffin M, O'Brien JJ, Smith TJ III (2012) A proxy for high-resolution regional reanalysis for the Southeast United States: assessment of precipitation variability in dynamically downscaled reanalyses. Clim Dynam 38:2449–2466

Syphard AD, Franklin J (2009) Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. Ecography 32:907–918

Tabor K, Williams JW (2010) Globally downscaled climate projections for assessing the conservation impacts of climate change. Ecol Appl 20:554–565

VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? Ecol Model 220:589–594

Watling JI, Romañach SS, Bucklin DN, Speroterra C, Brandt LA, Pearlstine LG, Mazzotti FJ (2012) Do bioclimate variables improve performance of climate envelope models? Ecol Model 246:79–85

Wilby RL, Fowler HJ (2010) Regional climate downscaling. In: Fung F, Lopez A, New M (eds) Modelling the impact of climate change on water resources. Wiley, Chichester, pp 34–85

Wood AW, Maurer EP, Kumar A, Letternmaier DP (2002) Long-range experimental hydrologic forecasting for the eastern United States. J Geophys Res 107:4429