



Ensemble forecasting shifts in climatically suitable areas for *Tropidacris cristata* (Orthoptera: Acridoidea: Romaleidae)

JOSE ALEXANDRE F. DINIZ-FILHO,¹ JOÃO CARLOS NABOUT,² LUIS MAURICIO BINI,¹ RAFAEL DIAS LOYOLA,¹ THIAGO FERNANDO RANGEL,³ DAVID NOGUES-BRAVO⁴ and MIGUEL B. ARAÚJO^{5,6} ¹Departamento de Ecologia, ICB, Universidade Federal de Goiás (UFG), Goiânia, GO, Brazil, ²Programa de Pós-Graduação em Ciências Ambientais, UFG, Goiânia, GO, Brazil, ³Department of Ecology and Evolution, University of Connecticut, Storrs, CT, USA, ⁴Department of Biology, Center for Macroecology, Evolution and Climate, University of Copenhagen, Copenhagen, Denmark, ⁵Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain and ⁶Cátedra de Biodiversidade Rui Nabeiro, CIBIO, Universidade de Évora, Largo dos Colegiais, Évora, Portugal

Abstract. 1. The effects of climate change on species' ranges have been usually inferred using niche-based models creating bioclimatic envelopes that are projected into geographical space. Here, we apply an ensemble forecasting approach for niche models in the Neotropical grasshopper *Tropidacris cristata* (Acridoidea: Romaleidae). A novel protocol was used to partition and map the variation in modelled ranges due to niche models, Atmosphere-Ocean Global Circulation Models (AOGCM), and emission scenarios.

2. We used 112 records of *T. cristata* and four climatic variables to model the species' niche using five niche models, four AOGCMs and two emission scenarios. Combinations of these effects (50 cross-validations for each of the 15 subsets of the environmental variables) were used to estimate and map the occurrence frequencies (EOF) across all analyses. A three-way ANOVA was used to partition and map the sources of variation.

3. The projections for 2080 show that the range edges of the species are likely to remain approximately constant, but shifts in maximum EOF are forecasted. Suitable climatic conditions tend to disappear from central areas of Amazon, although this depends on the AOGCM and the niche model. Most of the variability around the mapped consensus projections came from using distinct niche models and AOGCMs.

4. Although our analyses are restricted to a single species, they provide new conceptual and methodological insights in the application of ensemble forecasting and variance partition approaches to understand the origins of uncertainty in studies assessing species responses to climate change in the tropics.

Key words. Climate change, ensemble forecasting, niche models, Orthoptera, *Tropidacris*, variance partition.

Introduction

Knowing the geographic distribution of a species and the environmental suitability in each of its parts is a crucial step towards effective conservation strategies (Araújo & Williams, 2000; Thomas *et al.*, 2008). In general, this is achieved by niche models, which are built using different algorithms that establish

Correspondence: Jose Alexandre F. Diniz-Filho, Departamento de Ecologia, ICB, Universidade Federal de Goiás (UFG), Cx.P. 131, 74001-970 Goiânia, GO, Brazil.
E-mail: jafdinizfilho@gmail.com

correlations between a species' occurrence and the environmental variables measured at the same locations (reflecting thus the 'Grinnelian' component of ecological niche – *sensu* Soberón, 2007) (see Pearson & Dawson, 2003; Araújo & Guisan, 2006; Beaumont *et al.*, 2007; Colwell & Rangel, 2009 for conceptual discussions on niche models and their projections in geographical space). Once the models are established, they are projected into geographical space and thus can be used to estimate the species' potential geographic range based on a relatively small amount of information.

The use of niche models is usually considered as a useful first approach to overcome the so-called 'Wallacean' shortfall (i.e., the lack of knowledge about species distributions; see Bini *et al.*, 2006). Moreover, after building the niche model for a given species, it can also be projected into a different sets of environmental conditions, obtained for different time periods or biogeographic regions, thus being potentially useful to predict species' responses to climate change (e.g., Thomas *et al.*, 2004; Thuiller *et al.*, 2005a; Araújo *et al.*, 2006), reconstructing past distributions (e.g., Martinez-Meyer *et al.*, 2004; Kidd & Ritchie, 2006; Nogués-Bravo *et al.*, 2008; Nogués-Bravo, 2009) or predicting biological invasions (e.g., Peterson, 2003; Thuiller *et al.*, 2005b; Giovanelli *et al.*, 2008; Peterson *et al.*, 2008a).

There are now several methods to model occurrences as a function of environmental variables and build the bioclimatic envelopes, which are the basis of niche models (see Segurado & Araújo, 2004; Elith *et al.*, 2006; Meynard & Quinn, 2007; Tsoar *et al.*, 2007; Phillips & Dudík, 2008; Elith & Graham, 2009 for recent reviews and comparative evaluations). These niche models range from very simple bioclimatic envelope models, in which only occurrences are used to estimate the amplitude of environmental variation which is suitable to the species (e.g., BIOCLIM, Busby, 1991), up to complex artificial-intelligence based algorithms (e.g., neural networks, Ripley, 1996). These methods have different mathematical and statistical properties, and possess different fitting abilities that may constrain the predictions (e.g., Segurado & Araújo, 2004; Elith *et al.*, 2006). Although it is possible to begin understanding the why different models provide different solutions (see Pearson *et al.*, 2006; Elith & Graham, 2009), there are many issues which have been intensively debated on (i) the theoretical basis underlying each of these methods, or model classes and subclasses (*sensu* Araújo & Guisan, 2006), (ii) the causal basis for the relationship between occurrences and environmental or evolutionary factors (Araújo & Pearson, 2005; Kearney *et al.*, 2008; Pearman *et al.*, 2008), (iii) the relationship between model fit and model transferability (Araújo & Rahbek, 2006; Randin *et al.*, 2006; Peterson *et al.*, 2007; Peterson & Nakazawa, 2008), and (iv) on how to correctly evaluate model fit (e.g., Liu *et al.*, 2005; Lobo *et al.*, 2008; Peterson *et al.*, 2008b).

To overcome some of these challenges, Araújo and New (2007) proposed that a more conservative approach to model species distributions should be based on a combination of models and their derived projections, built under different initial conditions (e.g., datasets, pseudo-absences), model parameterisation (e.g., distinct sets of environmental variables, or their polynomial expansions), model classes (e.g., different niche models) and boundary conditions [e.g., Atmosphere-Ocean

Global Circulation Models (AOGCMs)]. This approach has been termed 'ensemble forecasting' and has recently started to be used in species distribution modelling (Araújo *et al.*, 2005; Thuiller *et al.*, 2005a; Araújo *et al.*, 2006; Marmion *et al.*, 2009; O'Haney, 2009; Coetzee *et al.*, 2009; Roura-Pascal *et al.*, 2009; Diniz-Filho *et al.*, 2009a,b). Actually, some of the more complex niche models, such as Genetic Algorithm for Rule Set Production (GARP) (Stockwell & Noble, 1992), random forests (Breiman, 2001) and neural networks (Ripley, 1996), are often applied with the same reasoning, generating a large number of models, performing an evaluation and then combining the results of the best models (for review, see Araújo & New, 2007).

Here, we discuss the application of the ensemble forecasting approach to evaluate shifts in the suitable climatic conditions for a tropical insect, the giant grasshopper *Tropidacris cristata* (Carbonell, 1986). However, beyond a simple comparative evaluation of the different modelling techniques and the projected shifts in the distribution of climatically suitable areas, as usually performed, we applied a new protocol (Diniz-Filho *et al.*, 2009a) that allow us to partition the variance of ensembles of forecasts into multiple components and map the sources of uncertainty across niche models, AOGCMs, and emission scenarios, as well as their interactions.

Methods

The species

Tropidacris cristata belongs to the largest known grasshopper genus of the family Romaleidae (Orthoptera: Acridoidea). The species achieves almost 14 cm in length (frons to wing tips) and a wing span of ca. 24 cm. These grasshoppers are locally abundant and widely distributed, being found in different habitat types across their range. The genus was reviewed by Carbonell (1986) and three species (*T. cristata*, *Tropidacris collaris* and *Tropidacris descampsi*) were recognised. In South America, *T. collaris* is locally more abundant than *T. cristata*, but it does not expand towards Central America. *Tropidacris descampsi* is a recently described species based on a single specimen from Colombia. Also, *T. cristata* was subdivided by Carbonell (1986) into three subspecies, *Tropidacris cristata cristata*, occupying the central part of South America, *T. c. grandis*, occupying the southern part of South America (up to extreme south of Brazil and northern Argentina) and *T. c. dux*, occurring in Central America, from Costa Rica up to Yucatán and southern Mexico.

Data

We obtained a total of 112 records of occurrence of *T. cristata* (including the three subspecies) from Carbonell's (1986) comprehensive review of the genus, which were recorded in a grid with 1646 cells with 1° of latitude and longitude covering the Neotropics. A search for recent records of this species in different databases (e.g., Thomson ISI) and CRIA (<http://www.cria.org>).



Fig. 1. Locations of the giant grasshopper (*Tropidacris cristata*) occurrences used for niche modelling.

br) did not add new occurrences. These records are widely distributed along the Neotropics (Fig. 1) and suggest that species can be found mainly in wet forests of Central America up to Yucatán and southwestern part of Mexico, the Amazon and Atlantic Forest, and in parts of the Brazilian Cerrado and northern Argentina. The species is lacking in most of the dry areas of Central Brazil, in Cerrado and Caatinga biomes. We modelled the distribution of the species both for current and future (forecasted) climate conditions throughout the Neotropics.

The occurrences of *T. cristata* were modelled as a function of four climatic variables (mean annual rainfall and its variability, average temperature of the warmest and coldest months) derived from four different AOGCMs: CCSM3, CSIRO-MK3.0, UKMO-HadCM3, and ECHAM5/MPI-OM. For each AOGCM, we obtained the estimated current climate and the projected climate in two emission scenarios, A1 (more pessimistic in terms of CO₂ emissions) and B1 (more optimistic). Data were extracted from the World Climate Research Program's (WCRP) Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel dataset (Meehl *et al.*, 2007) (see Diniz-Filho *et al.*, 2009a for further details).

Niche models and partitioning sources of uncertainty

We modelled *T. cristata* distribution using the following 'presence-only' niche models (see Tsoar *et al.*, 2007): a simple surface range envelope model based on orthogonal limits, the BIOCLIM (Busby, 1991), Euclidean distances (EUCL), Mahalanobis distances (MAH; Farber & Kadmon, 2003), GARP (Stockwell & Noble, 1992), and Maximum entropy (MAXENT; Phillips *et al.*, 2006; Phillips & Dudík, 2008). These methods have been widely used and comparative analyses of their statistical performance can be found elsewhere (Segurado & Araújo, 2004; Elith *et al.*, 2006; Meynard & Quinn, 2007; Tsoar *et al.*, 2007; Phillips & Dudík, 2008; Elith & Graham, 2009), so only a brief description of their implementation will be provided here. BIOCLIM,

EUCL and MAH are strictly presence-only methods, because they use only environmental data, where the species occur to build the climatic envelopes (models), whereas both GARP and MAXENT generated pseudo-absence data to fit the environmental models. However, in all cases, pseudo-absences were used to allow model evaluation using receiving operating curve (ROC) and true skill statistics (TSS) (see below). GARP was implemented as a single-run and the best subset was defined by modelling different dataset partitions and variable combinations. MAXENT algorithm is the same available at <http://www.cs.princeton.edu/~schapire/maxent/> (see Phillips & Dudík, 2008) and was run using default options (including regularisation parameter equal to 1).

For each one of the niche models described above, a total of 750 models were generated. First, the full dataset was randomly partitioned into two subsets (training and projection), with 70% and 30% respectively, and this procedure was replicated 50 times. For each partition, we obtained a total of 15 models by performing a full factorial combination of the four environmental variables. Adding a quadratic term to temperature did not qualitatively improve the results. For each model, pseudo-absences were generated to allow building the ROC (see Fielding & Bell, 1997; Allouche *et al.*, 2006), and the ROC was used to transform quantitative predictions of models (e.g., distances from species' niche centroids as given by the Mahala distances) into a binary vector of 0/1, which indicates absence or presence of the species in each cell. The cut-off point in BIOCLIM, EUCL and MAH was established by using multiple delimitations of the bioclimatic envelope (99%, 95%, etc.) allowing us to generate the ROC.

Following Allouche *et al.* (2006), the evaluation of models was done by the TSS, which was used to weight the final estimated occurrence frequencies (EOF) vector before ensembling (see below) (after deleting models with TSS < 0.5). There are discussions about the validity of any of these statistics in terms of model fit and transferability (see Peterson *et al.*, 2007, 2008a,b; Lobo *et al.*, 2008) but here the focus is on generating a

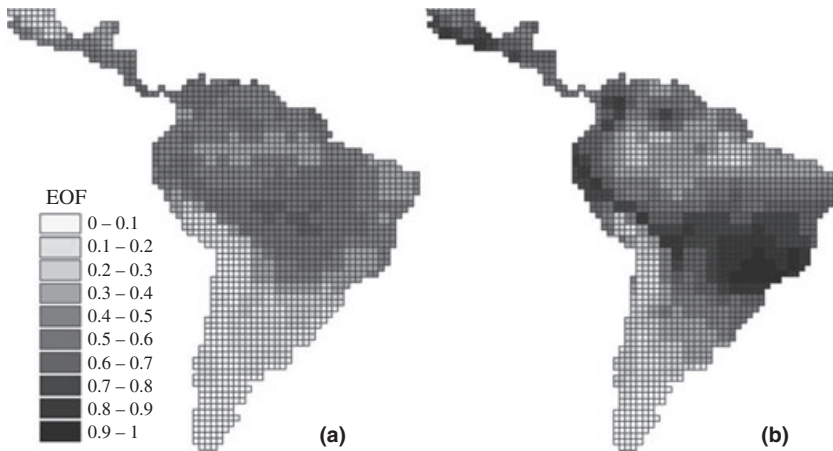


Fig. 2. Estimated frequencies of occurrence (EOF) for *Tropidacris cristata* in Neotropics, based on the consensus of 30 000 models based on different niche-modelling methods, Atmosphere-Ocean Global Circulation Models (AOGCMs) for current time (a) and future climate change scenarios (b).

consensus among alternative forecasts rather than using particular test statistics to select a ‘best’ model.

We analyse the frequency with which *T. cristata* is predicted in each cell across the whole set of models, called here EOF. Thus, our final analyses are actually based on a total of 30 000 models (750 models in each of the 40 combinations of niche model, AOGCM and emission scenario). Each of the 40 EOF vectors was then overlaid and an unweighted consensus of the frequencies was produced (see Marmion *et al.*, 2009). Weighting the EOF vectors by the average TSS of models did not qualitatively affect our results (especially because poor models, with low TSS statistics, were excluded before calculating EOFs). All these analyses were carried out in the integrated BIOENSEMBLES software (Diniz-Filho *et al.*, 2009a; Rangel *et al.*, 2009).

The relative importance of each of the sources of variation studied here (i.e., methods for niche modelling, AOGCMs and emission scenarios) on the EOF for *T. cristata* was quantified by using a three-way analysis of variance (ANOVA) without replication (Sokal & Rohlf, 1995), following the approach recently proposed by Diniz-Filho *et al.* (2009a). The relative sum of squares (SS) estimates the variance components, which is expressed as the percentage of explanation of each of the three main sources and their interaction. As these variance components were obtained for each cell in the grid, it is possible to map where each source is more important. A principal component analysis (PCA) of the EOF vectors (decomposing the pairwise correlation matrix between EOF maps) was also used to evaluate which levels within each source of uncertainty are more similar (see Thuiller, 2004; Araújo *et al.*, 2005, 2006).

Results

The consensus maps of the 20 EOF vectors obtained for the current time (i.e., four current climates for each AOGCM modelled using five niche-modelling methods) match the occurrence *T. cristata* throughout the Neotropics, from Northern Argentina up to western coast of Mexico, but relatively high frequencies (>0.5) appear mainly for Central South America (Fig. 2a). The fit statistics of the models, for each AOGCM vary, but relatively high TSS values around 0.7 were found, with values around

75% or higher of the presences being correctly identified by all models.

When using the emission scenarios A1 and B1 to project each model into future climatic conditions, the average EOF vector in 2080 did not change substantially the consensus map in terms of general distributional limits (Fig. 2b). However, there are clear changes of EOF patterns in 2080, characterised by a southeast shift of the higher frequencies, from Central American towards the southeast, and an increase in EOF in Central America and northwestern Amazon. Gains in EOF, in other words, in the suitability of climatic conditions for the species, appear in the southeast coast of Brazil, southern Andes and western coast of

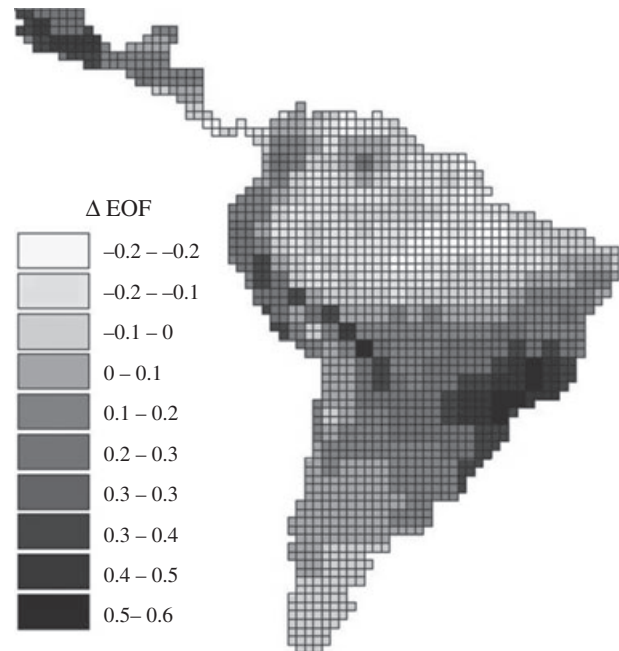


Fig. 3. Difference (Δ) between estimated occurrence frequencies (EOF) vectors between 2080 and 2000, with darker regions showing gain (i.e., higher EOF in 2080 than in 2000) and brighter ones indicating loss in EOF (i.e., lower EOF in 2080 than in 2000).

Table 1. Median proportions of the total sum of squares from the three-way ANOVA performed for each grid cell covering the Neotropics, evaluating the relative contributions of method for niche models, Atmospheric-Ocean Global Circulation Models (AOGCM) and emission scenarios to the variability in forecasting EOF of *Tropidacris cristata*. Minimum and maximum values in the maps are also given (see also Fig. 4).

Source	SS (%) median	Min–max
Niche model	29.1	2.0–93.3
AOGCM	28.5	0.4–82.6
Scenario	2.3	0.0–44.7
Niche model × AOGCM	17.7	2.7–56.0
Niche model × scenario	1.0	0.0–13.0
AOGCM × scenario	4.2	0.0–34.7
Third-order interaction	2.9	0.0–14.7

Mexico, whereas reductions in EOF appear in the Central Amazon (Fig. 3).

The three-way ANOVA applied to the EOF vectors in each cell shows that the median of the variation in future projections are due to differences in niche models is around 29%, but going as high as 93% for some regions in southeastern coast of Brazil and Patagonia (where the species is currently not found, so the higher component appears just because a few models predict the species there with low EOF, whereas others give values close to zero) (Table 1). The maps showing this SS (Fig. 4a) indicates that the largest differences among methods appear in the Brazilian southeast coast. The second most important source of variation is AOGCM (median of 28%, with values going up to 83%) and the interaction between AOGCM and niche model, which explains about 17% across cells. The map of the SS attributable to AOGCM (Fig. 4b) shows a higher uncertainty in the northeast coast of South America and in the Brazilian northeastern region, in the dry Caatinga biome (Fig. 4b). Finally, the map of the interaction between method and AOGCM shows a patch of

high uncertainty in the central part of Amazon, with a peak in its eastern portion (Fig. 4c).

As niche models and AOGCMs account for most of the variability in the ensemble of model projections, we reduced the dimensionality of the EOF vectors using a PCA based on A1 scenario only, to make the visualisation of the similarity among the maps easier. The first principal component accounted for 63% of the correlation among maps whereas the second one reported 9%. The loadings of the PCA on these two axes (Fig. 5) show that all niche models tend to show a similar pattern, but MAXENT based on the ECHAM AOGCM tend to give a slightly different result, as well as GARP and EUCL based on CCSM3. Some of the niche models (MAXENT and Euclidean distances) are more variable than others when based on different AOGCMs, which explains the relatively high interaction term in the ANOVA. Thus, predictions based on average EOF vectors for *T. cristata* are similar, even though they may differ for more regional analyses and interact in complex ways with climatic models used.

Discussion

Geographic range and shifts in climatically suitable areas for Tropidacris cristata

Changes in geographical ranges or suitable climatic areas for species under climate change detected by species distribution and niche modelling have been predicted for several groups of organisms, including several species of insects. They are actually one of the most important contributions of macroecology to the evaluation of climatic change impacts (Kerr *et al.*, 2007). However, these analyses usually focus on changes in range limits and overall distributional patterns and are usually performed for temperate northern hemisphere, where changes are more pronounced due to steepest climatic gradients (but see Bush & Hooghiemstra, 2005). Here, our analyses with a broadly distrib-



Fig. 4. Maps of variance component (relative sum of squares) for the effect of niche-modelling methods (a), Atmosphere-Ocean Global Circulation Models (AOGCMs) (b) and the interaction between both sources of uncertainty (c).

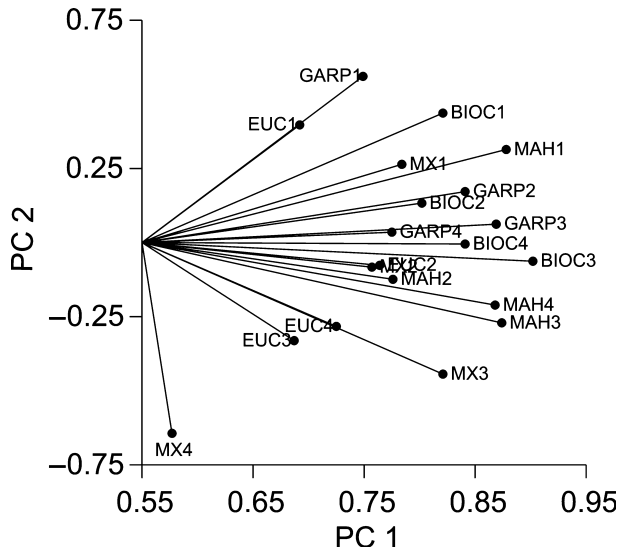


Fig. 5. Principal component loadings on the first two axes (PC1 and PC2) from the correlation among the maps for A1 scenario built using all niche-modelling methods and Atmosphere-Ocean Global Circulation Models (AOGCMs). The first and the second principal components accounted for 63% and 9.3% of the variation among maps, respectively. The niche-modelling methods are BIOCLIM (BIOC), Euclidean distances (EUC), Mahalanobis distances (MAH), Genetic Algorithm for Rule Set Production (GARP) and Maxent (MX), whereas the AOGCMs are numbered sequentially as CCSM3 (1), CSIRO (2), HADCM3 (3) and ECHAM (4). Only the emission scenario A1 was considered here because differences between A1 and A2 scenarios account for <3% of the variation among maps.

uted Neotropical species of grasshopper allow a clearer interpretation of the effects of climate change in the tropics, not only in terms of limits, but mainly in terms of shifts in suitable climatic areas for the species.

The current geographic range of *T. cristata* is well modelled based on the four climatic variables, according to the TSS evaluation statistics, and as suggested by the occurrence data that limits it from northern Argentina up to southern Mexico (but niche models expand slightly northward, up to California). The maps produced reflect the broad range of occurrences and the general description of species distribution by Carbonell (1986). In the models for the current time, the maximum EOF is found in central Brazil, in the ecotone between Cerrado and Amazon. Carbonell (1986) pointed out that *T. cristata* is found in moist forests, and the absence of the species in the Caatinga ecoregion is probably certain, and that species must occur only in 'islands' of dense vegetation inside the driest part of the Brazilian Cerrado. EOF captured only partially this pattern, probably because of the coarse resolution of our spatial units and the lack of more detailed data on vegetation cover (which is important to characterise small patches of humid vegetation in the relative dry region of Cerrado in central Brazil). This scale issue is important and further work is necessary to throw light on how to combine broad-scale macroecological studies on suitability patterns with more refined analyses at local and regional scales.

Even so, it is interesting to note that the other widespread species of the genus, *T. collaris*, does occur in these driest habitats.

The comparison between current and projected ranges (2080) indicated minor changes in the distributional limits of the species, which is expected for large-ranged species whose niche is defined by very broad bioclimatic envelopes. Cutting the EOF under a majority consensus rule (of 50%), the range will be reduced from 12.7 (± 0.3) millions of km² to 9.8 (± 4.1) millions of km² in 2080 (averaging the EOFs for current and future climates), a difference that represents an average reduction of ca. 22%. Notice that there is a higher variance in the number of cells with EOF higher than 0.5 in 2080, so that changes actually go from 70% of losses to 60% gains, for different niche-modelling methods and AOGCMs. However, EOF *per se* changes substantially between 2000 and 2080. The environmental suitability decreases across the entire range and the location of the maximum EOF is shifted from central Brazil towards the Brazilian southeastern coast (Fig. 2).

The clearest impact occurs in the northern Amazon region, where a strong reduction in EOF is expected under climate change, although it is important to consider that EOF in current climate was low anyway. Because of the broad range size of *T. cristata*, its bioclimatic envelope is also wide and it encompasses the future climatic conditions in the Amazon. However, future conditions there are far from the species' optimum in the current time, which can explain the reduced EOF for 2080 and the associated low delta in respect to current one. Indeed, the delta EOF is mainly explained by shifts in two out of the four environmental variables used for the analyses. A multiple ordinary least squares (OLS) regression revealed that delta EOF is well explained by the four environmental variables between current climate and climate modelled for 2080 ($R^2 = 0.483$), with higher standardised coefficients for maximum temperature and average precipitation. The estimated values of this multiple regression model captures well the higher (negative) delta EOF in the Amazon, which will indeed become hotter and drier according to the models used here.

Thus, it is interesting to note that although AOGCMs predict larger changes in climate (mainly temperature) in temperate regions of the northern hemisphere, perhaps the severest impact on biodiversity will occur in the tropics (see also Colwell *et al.*, 2008). As pointed out by Bush and Hooghiemstra (2005), '...because moist tropical systems hold such huge diversity, and because the vast majority of species are thought to have restricted niches..., the potential exists for small climate perturbations to have a profound effect'. Although, *T. cristata* has a large geographic distribution, our analyses support this assertion, but it is important to stress that this does not appear in a simple evaluation of range limits, but instead in the reduction of the average environmental suitability across species range.

Ensemble forecast and sources of uncertainty

Both current and future EOFs discussed here are based on an ensemble forecasting approach that combines several niche models, AOGCMs and emission scenarios to provide a conser-

vative picture of the future range of *T. cristata* (see Marmion *et al.*, 2009). However, our analyses, beyond providing a simple consensus map, also allow quantifying and mapping the sources of uncertainty associated to such approach.

The most important source of uncertainty in establishing climatically suitable areas is the niche model used, but this is followed by relatively large effects of AOGCM and the interaction between niche model and AOGCM. The PCA revealed that the EOF vectors are similar in general and do not reveal clear contrasts between simple 'bioclimatic envelope models' (such as BIOCLIM or Euclidian and Mahalanobis) and the more complex computer-intensive and optimisation models (i.e., MAX-ENT and GARP). Although further analyses using other niche models may reveal more consistent patterns comparing these different 'model subclasses', it is important to note that when generating EOF predictions based on a large number of different subsamples of dataset and variable combination, TSS was used to eliminate models with bad fit. As a consequence, under the reasoning of Araújo and New (2007), combining predictions from different models (and other sources of uncertainty, such as AOGCMs) can provide, in principle, a more conservative estimate of species' distribution or response to climate change, even when mixing different niche models. At the same time, evaluating how these models vary in geographical space may be helpful to understand the origins of uncertainty and the properties of different models.

The shift observed in EOF maps between 2000 and 2080 shows two clear patterns: an increase in EOF in the southeastern coast of Brazil and a decrease in EOF in the Amazon. The first one is associated with niche model uncertainty. On the other hand, the reduction in EOF in the central Amazon seems to be more consensual across niche models (low variance components – see Fig. 3) and is more clearly related to a reduction in precipitation and increase in temperature in the region. Even so, the relatively high eigenvalue of the first principal component, which explains more than 60% of the variance among niche models, reveals that all models tend to agree about these patterns, although the detailed 'shape' of increasing EOF along the southeastern coast of South America varies a bit between them. In general, models based on the ECHAM AOGCM produce more restricted increases in EOF in this region, for all niche-modelling methods.

On the other hand, the highest uncertainty from AOGCMs in the northeast coast of South America and in the northeastern of Brazil (in the dry Caatinga ecoregion), as well as the highest interaction between niche-modelling method and AOGCM in the central Amazon, are probably associated with uncertainties in forecasting precipitation patterns (see Berthelot *et al.*, 2005). This is a well known pattern among AOGCMs, and, on average, it is indeed expected that climate change will affect Amazon by reducing precipitation in the region (see Salazar *et al.*, 2007).

Concluding remarks

Although our analyses are restricted to a single species, it provides new conceptual and methodological insights for under-

standing the role of climate change in the tropics. Conceptually, the most important novelty presented here is the change of focus from occurrence only (distributional limits) to environmental suitability as given by EOF. Indeed, although range limits of *T. cristata* remains approximately constant, there are large changes in EOF that indicate that the Amazon will become more unsuitable for this species. Methodologically, the variance partition and mapping approach shows that the areas of highest suitability in 2080 are associated with the highest uncertainty attributable to the use of different niche models, whereas areas of decreasing EOF are associated with uncertainties related to the use of different AOGCMs. We expect that the application of ensemble forecasting and variance partition approach encourages more studies about the origins of uncertainty in species distribution modelling.

Acknowledgements

This paper is dedicated to the memory of the late Dr A. Mesa, for his lifelong dedication to insect genetics, systematic and evolution. This work was developed under the project 'Bioimpacto' supported by BBVA Foundation. Work by JAFDF, LMB, RDL and JCN in macroecology and biodiversity has been continuously supported by CNPq and CAPES. DNB thanks the Danish National Research Foundation for support to the Center for Macroecology, Evolution and Climate, whereas TFR has been supported by CAPES-Fullbright, NSF and University of Connecticut fellowships. MBA is also funded by the Spanish Ministry of Science and Innovation (CGL2008-01198-E/BOS).

References

- 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.
- Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.
- Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.
- Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distributions with climate. *Ecography*, **28**, 693–695.
- Araújo, M.B. & Rahbek, C. (2006) How does climate change affect biodiversity? *Science*, **313**, 1396–1397.
- Araújo, M.B., Thuiller, W. & Pearson, R.G. (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, **33**, 1712–1728.
- Araújo, M.B., Whittaker, R.J., Ladle, R. & Erhard, M. (2005) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, **14**, 529–538.
- Araújo, M.B. & Williams, P.H. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, **96**, 331–345.
- Beaumont, L.J., Pitman, A.J., Poulsen, M. & Hughes, L. (2007) Where will species go? Incorporating new advances in climate modeling into projections of species distributions. *Global Change Biology*, **13**, 1368–1385.

- Berthelot, M., Friedlingstein, P., Ciais, P., Dufresne, J.-L. & Monfray, P. (2005) How uncertainties in future climate change predictions translate into future terrestrial carbon flux. *Global Change Biology*, **11**, 959–970.
- Bini, L.M., Diniz-Filho, J.A.F., Rangel, T.F.L.V.B., Bastos, R.P. & Pinto, M.P. (2006) Challenging Wallacean and Linnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. *Diversity and Distributions*, **12**, 475–482.
- Breiman, L. (2001) Random forest. *Machine Learning*, **45**, 5–32.
- Busby, J.R. (1991) BIOCLIM – a bioclimate analysis and prediction system. *Nature Conservation: Cost Effective Biological Surveys and Data Analysis* (ed. by C.R. Margules and M.P. Austin), pp. 64–68. CSIRO, Australia.
- Bush, M.A. & Hooghiemstra, H. (2005) Tropical biotic responses to climate change. *Climate Change and Biodiversity* (ed. by T.E. Lovejoy and L. Hannah), pp. 125–156. Yale University Press, New Haven and London, US/UK.
- Carbonell, C.S. (1986) Revision of the Neotropical Genus *Tropidacris* (Orthoptera: Acridoidea, Romaleidae, Romaleinae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, **138**, 366–402.
- Coetzee, B.W.T., Robertson, M.P., Erasmus, B.F.N., van Rensburg, B.J. & Thuiller, W. (2009) Ensemble models predict Important Bird Areas in southern Africa will become less effective for conserving endemic birds under climate change. *Global Ecology and Biogeography*, **18**, 701–710.
- Colwell, R.K., Brehm, G., Cardelús, C.L., Gilman, A.C. & Longino, J.T. (2008) Global warming, elevational range shifts and lowland biotic attrition in the wet tropics. *Science*, **322**, 258–261.
- Colwell, R.K. & Rangel, T.F. (2009) Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19651–19658.
- Diniz-Filho, J.A.F., Bini, L.M., Rangel, T.F.L.V.B., Loyola, R.D., Hof, C., Nogués-Bravo, D. & Araújo, M.B. (2009a) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, **32**, 897–906.
- Diniz-Filho, J.A.F., Oliveira, G., Bini, L.M., Loyola, R.D., Nabout, J.C. & Rangel, T.F.L.V.B. (2009b) Conservation biogeography and climate change in the Brazilian Cerrado. *Natureza Conservação*, **7**, 100–112.
- Elith, J. & Graham, C.H. (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., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, **160**, 115–130.
- 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.
- Giovanelli, J.G.R., Haddad, C.F.B. & Alexandrino, J. (2008) Predicting the potential distribution of the alien invasive American bullfrog (*Lithobates catesbeianus*) in Brazil. *Biological Invasions*, **10**, 585–590.
- Kearney, M., Phillips, B.L., Tracy, C.R., Christian, K.A., Betts, G. & Porter, W.P. (2008) Modeling species distributions without using species distributions: the cane toad in Australia under current and future climates. *Ecography*, **31**, 423–434.
- Kerr, J.T., Kharouba, H.M. & Currie, D.J. (2007) The macroecological contribution to global change solutions. *Science*, **316**, 1581–1584.
- Kidd, D.M. & Ritchie, M.G. (2006) Phylogeographic information systems: putting the geography into phylogeography. *Journal of Biogeography*, **33**, 1851–1865.
- Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.
- Lobo, J.M., Jimenez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modeling. *Diversity and Distributions*, **15**, 59–69.
- Martinez-Meyer, E., Peterson, A.T. & Hargrove, W.W. (2004) Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography*, **13**, 305–314.
- Meehl, G.A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J.F.B., Stouffer, R.J. & Taylor, K.E. (2007) The WCRP CMIP3 multi-model dataset: a new era in climate change research. *Bulletin of the American Meteorological Society*, **88**, 1383–1394.
- Meynard, C.N. & Quinn, J.F. (2007) Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, **34**, 1455–1469.
- Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, **18**, 521–531.
- Nogués-Bravo, D., Rodriguez, J., Hortal, J., Batra, P. & Araújo, M.B. (2008) Climate change, humans, and the extinction of the woolly mammoth. *PLoS Biology*, **6**, e79. DOI: 10.1371/journal.pbio.0060079.
- O'Haney, J.R. (2009) Neural ensembles: a neural network based ensemble forecasting program for habitat and bioclimatic suitability analysis. *Ecography*, **32**, 89–93.
- Pearman, P.B., Guisan, A. & Broennimann, O. (2008) Niche dynamics in space and time. *Trends in Ecology and Evolution*, **23**, 149–158.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelopes useful? *Global Ecology and Biogeography*, **12**, 361–171.
- Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L.J., Segurado, P., Dawson, T.E. & Lees, D.C. (2006) Model-based uncertainty in species' range prediction. *Journal of Biogeography*, **33**, 1704–1711.
- Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology*, **78**, 419–433.
- Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: and example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135–144.

- Peterson, A.T., Papes, M. & Eaton, M. (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, **30**, 550–560.
- Peterson, A.T., Papes, M. & Soberón, J. (2008b) Rethinking receiver operating characteristic analysis: applications in ecological niche modeling. *Ecological Modelling*, **213**, 63–72.
- Peterson, A.T., Stewart, A., Mohamed, K.I. & Araújo, M.B. (2008a) Shifting global invasive potential of European plants with climate change. *PLoS ONE*, **3**, e2441, DOI: 10.1371/journal.pone.0002441.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.
- Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, **33**, 1689–1703.
- Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Araújo, M.B. (2009) *BIOENSEMBLES 1.0. Software for Computer Intensive Ensemble Forecasting of Species Distributions Under Climate Change*. Goiás, Madrid, Évora.
- Ripley, B.D. (1996) *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, UK.
- Roura-Pascal, N., Brotons, L., Peterson, A.T. & Thuiller, W. (2009) Consensual predictions of potential distributional areas for invasive species: a case study of Argentine ants in the Iberian Peninsula. *Biological Invasions*, **11**, 1017–1031.
- Salazar, L.F., Nobre, C.A. & Oyama, M.D. (2007) Climate change consequences on the distribution in tropical South America. *Geophysical Research Letters*, **34**, L09708, DOI: 10.1029/2007GL029695.
- Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.
- Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.
- Sokal, R.R. & Rohlf, F.J. (1995) *Biometry*, 3rd edn. W.H. Freeman, San Francisco, California.
- Stockwell, D.R.B. & Noble, I.R. (1992) Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematics and Computers in Simulation*, **33**, 385–390.
- Thomas, C.D., Bulman, C.R. & Wilson, R.J. (2008) Where within a geographic range do species survive best? A matter of scale. *Insect Conservation and Ecology*, **1**, 2–8.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y., Erasmus, B.F.N., De Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., Van Jaarsveld, A.S., Midgley, G.F., Miles, L.J., Ortega-Huerta, M.A., Townsend Peterson, A., Phillips, O. & Williams, S.E. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller, W. (2004) Patterns & uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.
- Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M. & Prentice, I.C. (2005a) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8245–8250.
- Thuiller, W., Richardson, D.M. & Pysek, P. (2005b) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.
- Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. (2007) A comparative evaluation of presence only methods for modelling species distribution. *Diversity and Distributions*, **13**, 397–405.

Accepted 19 February 2010

First published online 1 April 2010

Editor: Raphael K. Didham

Associate editor: Robert Ewers