

CORRESPONDENCE



Data sets matter, but so do evolution and ecology

A response to Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, 17, 135–144.

In a recent paper, Peterson & Nakazawa (2008) (hereafter PN) contest key findings in our study (Fitzpatrick *et al.*, 2007) that suggest that *Solenopsis invicta* (hereafter the fire ant) underwent a niche shift upon its invasion of North America. Using niche-based models, we proposed that the fire ant established in environments similar to those found in its native range but subsequently spread into environments unlike those found within its native range – a pattern strikingly similar to that suggested by Broennimann *et al.* (2007) for spotted knapweed (*Centaurea maculosa*). PN counter that our findings are simply an artefact of the environmental variables we used to model the fire ant's distributions and suggest instead that selection of alternative variables can produce a more correct prediction of the fire ant's invasion. PN conclude that the biological explanations offered in Fitzpatrick *et al.* (2007) for the non-predictivity between the fire ant's native and invaded distributions, namely enemy release, genetic founder effects and hybridization, are not necessary. Here we respond to PN's criticisms.

We disagree with the contentions outlined in PN on the grounds that the authors (1) subjectively consider what represents a 'correct' prediction of the fire ant's niche, (2) do not discuss the potential for niches to be conserved along some environmental axes but not others and, most significantly, (3) do not adequately represent our original analyses in Fitzpatrick *et al.* (2007) by not testing the ability of the fire ant's invaded distribution to predict its native range using their alternative data sets. We demonstrate, using the procedures outlined in Fitzpatrick *et al.* (2007) and the set of environmental variables in PN that represents a subset of the variables used in Fitzpatrick *et al.* (2007), that the results from our original study stand.

Issue 1: subjective consideration of what constitutes a 'correct' prediction

PN state that, owing to small sample sizes, their 'test' of model predictions was qualitative. The failure of models to 'anticipate the full northward extent of the species' invasion was taken as an indication of poor generalization' (emphasis ours). We do not take issue with such a qualitative and subjective 'test' of model quality *per se*. But, if PN apply such a test to predictions of the invaded range, they must also apply the same 'test' to distributions predicted for the native range. PN seem satisfied with predictions of the fire ant's invaded distribution in North America as long as models anticipate at least a portion of the northern limit of the fire ant's invasion (not the full northern limit or the western limit) – no matter how low model agreement or how poorly models predict other portions of the fire ant's distributions (e.g. over-prediction of the fire ant's native range). In contrast, they dismiss models that fail their test of a 'correct' prediction, but that replicate the fire ant's native distribution in South America (upon which the models were based), including models that correctly predict the southern limit of the native range, which is roughly analogous to the north limit of the introduced range.

Our differences in interpretation originate, at least in part, from an essential difference between the goal of Fitzpatrick *et al.* (2007) and that of PN. Fitzpatrick *et al.* (2007) attempted to test for and offer hypotheses that might explain a niche shift, while PN attempt to replicate a well-documented invasion by selecting variables that generally predict the fire ant's invaded distribution, regardless of model performance elsewhere. Therefore, PN consider the fire ant's niche to be modelled 'correctly' when the prediction meets their criteria in the invaded range, even if models fail to predict the native range. We consider the fire ant's niche to be modelled 'correctly' when models predict all extents of both the native and invaded ranges, because if the niche of a species is conserved, then a single model should in principle predict both the native and the invaded range (Wiens & Graham, 2005). Such a gestalt evaluation, in tandem with

comparisons in bioclimatic space (rather than geographical space alone, e.g. using principal components analysis), is more likely to identify instances of niche shifts (or lack thereof) rather than a focus on particular characteristics of the predicted invaded distribution alone.

On these grounds, we take particular issue with PN's claim that four of the environmental data sets used in their paper could correctly predict the fire ant's potential to invade North America – even when considering their definition of a 'correct' prediction. These four data sets include data from: (1) the Intergovernmental Panel on Climate Change (IPCC), (2) the Center for Climate Research at the University of Delaware (CCR), (3) monthly surface reflectance values drawn from the normalized difference vegetation index (NDVI), and (4) a subset of the data layers from the WorldClim data set ('reduced WC2'; see Peterson & Nakazawa, 2008, for full descriptions of these data and citations). Of these four, only the 'reduced WC2' data set comes close to correctly predicting both the invaded and the native range using native range occurrence data (but see Issue 3 below). Both IPCC and CCR do a poor job of predicting the fire ant's invasive potential in North America. There are absences in the predicted distributions where fire ants are known to be present and regions with thin coverage (i.e. low model agreement). The IPCC data set predicts, also with low model agreement, that fire ants could invade areas north of the Arctic Circle. To consider these models as correct predictions of the fire ant's invasive potential is misleading. NDVI does anticipate the full northward extent of the fire ant's invasion. However, NDVI also over-predicts the native range (including its southern extent), suggesting that NDVI does not limit the fire ant's native distribution. This notion is strengthened by the fact that NDVI also predicts coastal Maine and regions of Canada north of Minnesota to be susceptible to invasion by fire ants. Because fire ant physiology has been intensively studied, we know that these northern regions are not suitable areas that fire ants have yet to colonize. Such over-prediction is to be expected when remotely sensed data are used as

surrogates for climate variables because distant regions may exhibit similar spectral signatures even if they have substantially different climates.

Issue 2: the potential for niches to be conserved along some environmental axes but not others

Given the amount of baggage that comes with the niche concept and its relationship to niche-based models, it is debatable whether differentiating between fundamental and realized niches is useful (Guisan & Thuiller, 2005; Araújo & Guisan, 2006; Soberón, 2007). However, in a general sense, distinguishing between fundamental and realized niches is a simple way to clarify the primary issue with projecting biological invasions using observed distributions of species in their native range. Further, distinguishing between fundamental and realized niches is useful when discussing niche conservatism because niche shifts can result from a change in the realized niche only (e.g. relaxation of biotic constraints on distribution with no change in climatic tolerances), or also from a change in both the realized and fundamental niche (Pearman *et al.*, 2008).

Niche-based models are applied and often discussed in the context of Hutchinson's niche concept. As defined by Hutchinson (1957), the fundamental niche represents the complete set of environmental conditions under which a species can persist, whereas the realized niche is the subset of those conditions within the fundamental niche that the species actually occupies. Because observed distributions of species reflect multiple determinants, including climatic tolerances, biotic interactions, and dispersal limitation, niche-based models developed using observed distributions will predict the geographic equivalent of the realized niche. When such a model is projected, the model identifies where the species is likely to invade as long as the combinations of biotic and abiotic constraints on the native distribution of the species remain unchanged and the species does not evolve. As has been widely theorized and empirically validated, changes to both realized and fundamental niches are possible during an invasion given the potential for release from biotic and other non-climatic constraints on distribution and adaptation (see Pearman *et al.*, 2008, for a recent review of these topics as well as a comprehensive list of examples of both niche shifts and niche conservatism drawn from many taxa).

A larger issue is the fact that there is no standard measure of what constitutes a niche shift. How much a species' niche has to

change for it no longer to be conserved is an open question. It is unlikely that any introduced species invades a new territory without experiencing some degree of niche shift, since it is highly unlikely that identical combinations of environmental conditions exist in both the native and introduced ranges – especially when considering more than a few environmental variables. Whether such niche shifts result from species realizing more of their fundamental niche or from founder effects or subsequent evolution that leads to change in both the realized and fundamental niche is irrelevant to our argument as niche-based models cannot distinguish these possibilities. Nonetheless, decades of evolutionary and ecological theory and a large body of empirical evidence documenting that invasive species can experience rapid evolution as well as release from biotic constraints on distribution suggest that niche shifts should be commonplace when species are introduced to new biogeographical settings.

In this vein, PN do not explore as a possible explanation for the ability of their models with fewer variables to better replicate the fire ant's invasion that niches may shift along some environmental axes while being conserved along others. There is little reason to think that a species' niche will shift along all environmental axes simultaneously. It is entirely plausible, and we would argue much more likely, for a species' niche to shift along one axis or a few axes such that they may tolerate, say, different moisture conditions, while conserving their tolerance of minimum temperature. Such a scenario may explain why the 'reduced WC2' data set predicts more of the fire ant's invaded distribution than variables used in our original analysis. The fire ant's niche may have shifted along an environmental axis represented by variables in the 'full WC2' data set, but which is not represented in the 'reduced WC2' data set. Further, given that dimensionality is reduced as environmental variables are removed from consideration, models will tend to produce a broader predicted niche (and distribution) because the number of possible constraints on the niche is correspondingly reduced as well. In any event, as we outline in Issue 3, our analysis using the 'reduced WC2' data set does not eliminate the necessity for biological explanations for the non-transferability of models between the fire ant's ranges as claimed by PN.

Issue 3: incomplete replication of our original analysis

Despite the availability of data describing the fire ant's invaded distribution, PN employed

only native distribution data in their analysis (and used slightly different native distribution data than the data used in our original analysis). We performed an analysis identical to that described in Fitzpatrick *et al.* (2007) using PN's 'reduced WC2' data set and the original Desktop GARP algorithm within the Open Modeller framework. We focus on the 'reduced WC2' data set because it represents a subset of the original variables used in Fitzpatrick *et al.* (2007). In addition, we used the 'ade4' package in R version 2.6.0 to test for niche conservatism by comparing the positions of native and invaded range distribution data in the climatic space resulting from a principal components analysis on the 'reduced WC2' data set. We weighted occurrences to ensure that both the invaded range (741 points) and the native range (74 points) had equal representation. The significance of the difference between the fire ant's native and invaded niches (i.e. the two clusters of points in PCA space) was assessed using a between-class analysis (see Broennimann *et al.*, 2007, for a relevant application) and by performing a Monte Carlo test (99 permutations) on the resulting between-class inertia percentage.

Our analysis using the 'reduced WC2' data set confirmed our original findings. When examined in 'reduced WC2' climatic space, the invaded niche of the fire ant is significantly different from its native niche (between-class inertia: 40.0%; $P < 0.01$), mainly along an axis associated with temperature (data not shown). This finding suggests that the fire ant has invaded colder temperatures than those characterizing its native distribution. This niche shift was revealed in geographical space when models developed using the 'reduced WC2' data set were projected (Fig. 1, right panel). Models developed using native range occurrences failed to predict the full northward extent of the fire ant's invasion (even when we considered model agreement as low as 25%; black shading in Fig. 1b, right panel), whereas models developed using invaded range occurrences also over-predicted the southern limit of the native range (Fig. 1d, right panel). These projections are nearly identical to those obtained in our original analysis (Fig. 1, left panel).

We continue to argue that these 'prediction errors' are biologically interesting and a more biologically rigorous model confirms our notion. Morrison *et al.* (2004) used a mechanistic, physiological model based on colony growth rates in North America (Korzukhin *et al.*, 2001) to predict the potential global extent of the fire ant's distribution. In accordance with our analysis, predictions from the colony-growth model also suggest that the

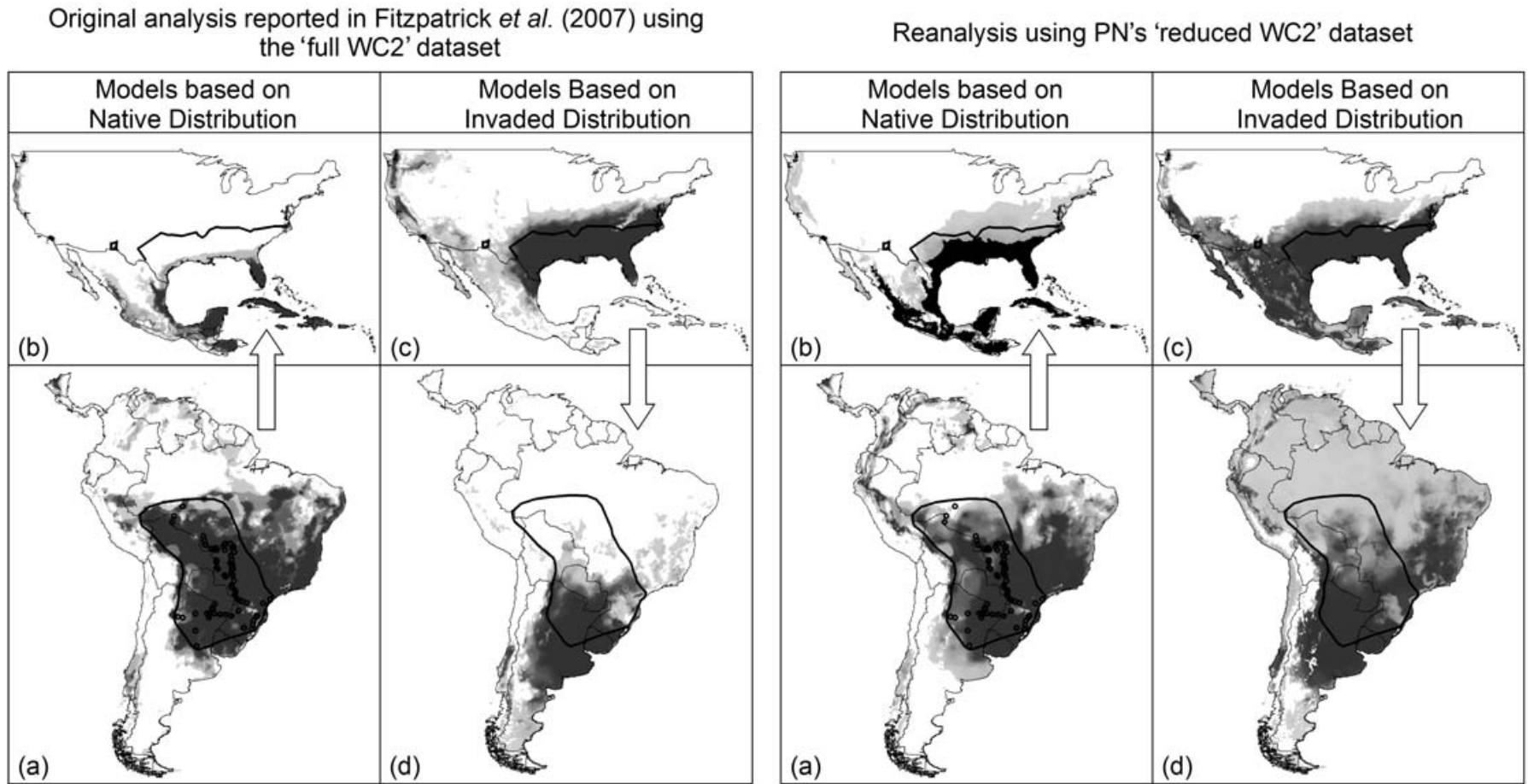


Figure 1 Potential distributions of *Solenopsis invicta* developed using niche-based models and two environmental data sets. The left panel is the original as published in Fitzpatrick *et al.* (2007) and contended by Peterson and Nakazawa (2008). The right panel replicates our original analysis using the reduced WorldClim data set (reduced WC2) of Peterson and Nakazawa (2008). In both panels, native range models represent (a) the potential native and (b) the potential invaded distributions of the fire ant based on 74 known occurrences in South America (a, open circles). Invaded range models represent (c) the potential invaded and (d) the potential native range of the fire ant based on the central points of 741 US counties (c, points not shown). Bold, solid lines indicate the approximate extent of the native (a, d) and invaded (b, c) range of the fire ant. Darker shading represents greater model agreement. Black shading in the right panel (b) represents areas where model agreement is at least 25%. As in the original analysis on the left, the 'reduced WC2' data set under-predicts the invaded range (b, right panel) and over-predicts the native range (d, right panel).

fire ant's native distribution could extend further south than its currently recognized boundary in South America (Morrison *et al.*, 2004). The most parsimonious explanation, supported by both niche-based and physiological models, is that the fire ant's niche was not conserved upon its invasion of North America. Whether this apparent niche shift represents a change in the fire ant's realized or fundamental niche remains unclear, because, to our knowledge, no such physiological model has been developed for fire ant populations in South America.

There is little reason to believe that predicted distributions based on species distribution models will ever match observed distributions perfectly. Certainly some prediction errors will prove to be uninteresting and related to data quality or statistical inaccuracies. Therefore, it is important to point out such potential sources of uncertainty in both our original analysis and that presented here. For example, the environmental conditions that fire ants experience on the ground are likely to differ vastly in some regions from those characterized by temporally and spatially generalized climate data – especially in regions such as the desert south-west of the United States where fire ants persist mainly where irrigation is prevalent. This fact alone could account for some modelling discrepancies and highlights the caution required when using niche-based models to test hypotheses regarding species–climate relationships (Araújo *et al.*, 2005). Further, there is now a consensus among researchers that projections can vary widely with the statistical technique used to model geographical distributions and therefore a range of modelling techniques and ensemble forecasting (Araújo & New, 2007) should ideally be used to reduce and quantify such model-based uncertainty. In both the analysis here and our original analysis we used only one algorithm, GARP. An investigation of the ability of other statistical approaches to predict the invasion of the fire ant (and other invasive species) is warranted. In fact the well-studied fire ant could serve as an excellent test of the ability of different techniques to project invasions. Finally, in keeping with our interest in replicating our original analysis, we did not validate our findings using all of PN's data sets, namely IPCC, CCR or NDVI.

Nonetheless, we view certain model errors as biologically interesting and necessitating

biological explanations – since it is biological processes that species distribution models notoriously ignore. Niches can change owing to drift, enemy release, selection, hybridization and simply as a consequence of genetic founder effects during invasion. Some or all of these factors could result in niche shifts that are potentially detectable at the broad spatial scales at which niche-based models are commonly applied. Understanding the prevalence of and mechanisms behind such shifts is of theoretical and applied interest and may facilitate improvements in our ability to anticipate both biological invasions and the potential impacts of climate change on biodiversity. We agree that the role of environmental data sets in these issues merits careful investigation. However, by implying that model errors are simply the result of variable selection and do not warrant biological explanations, PN may have inadvertently exposed niche modelling studies to yet another criticism.

ACKNOWLEDGEMENTS

M.C.F. acknowledges support from the University of Tennessee in the form of a Yates Dissertation Fellowship and through the Department of Ecology and Evolutionary Biology. We thank an anonymous referee, Gregory Crutsinger, William Hargrove, J. P. Lessard, David Nogués-Bravo and Daniel Simberloff for improving an early draft of this paper.

MATTHEW C. FITZPATRICK¹*, ROBERT R. DUNN² AND NATHAN J. SANDERS¹

¹*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996-1610, USA*

E-mail: mfitzpatrick@utk.edu

²*Department of Zoology, North Carolina State University, Raleigh, NC 27695, USA*

*Correspondence: Matthew C. Fitzpatrick, University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996-1610, USA.

E-mail: mfitzpatrick@utk.edu

REFERENCES

- Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modeling. *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., Thuiller, W. & Erhard, M. (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.
- Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.
- Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24–33.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.
- Korzukhin, M.D., Porter, S.D., Thompson, L.C. & Wiley, S. (2001) Modelling temperature-dependent range limits for the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the United States. *Environmental Entomology*, **30**, 645–655.
- Morrison, L.W., Porter, S.D., Daniels, E. & Korzukhin, M.D. (2004) Potential global range expansion of the invasive fire ant, *Solenopsis invicta*. *Biological Invasions*, **6**, 183–191.
- Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time: species distribution modeling meets phylogenetics. *Trends in Ecology and Evolution*, **23**, 149–158.
- Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135–144.
- Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115–1123.
- Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 519–539.

Editor: José Alexandre F. Diniz-Filho

doi: 10.1111/j.1466-8238.2008.00389.x