Challenges in the application of geometric constraint models

Craig R. McClain[1]*, Ethan P. White[2,3] and Allen H. Hurlbert[4]

[1]*Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA,* [2]*Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA,* [3]*Department of Biology, Utah State University, Logan, UT 84322, USA and* [4]*National Center for Ecological Analysis and Synthesis 735 State Street, Suite 300, Santa Barbara, CA 93101, USA*

*Correspondence: Craig R. McClain, Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA. E-mail: cmcclain@mbari.org

ABSTRACT

Discerning the processes influencing geographical patterns of species richness remains one of the central goals of modern ecology. Traditional approaches to exploring these patterns have focused on environmental and ecological correlates of observed species richness. Recently, some have suggested these approaches suffer from the lack of an appropriate null model that accounts for species ranges being constrained to occur within a bounded domain. Proponents of these null geometric constraint models (GCMs), and the mid-domain effect these models produce, argue their utility in identifying meaningful gradients in species richness. This idea has generated substantial debate. Here we discuss what we believe are the three major challenges in the application of GCMs. First, we argue that there are actually two equally valid null models for the random placement of species ranges within a domain, one of which actually predicts a uniform distribution of species richness. Second, we highlight the numerous decisions that must be made to implement a GCM that lead to marked differences in the predictions of the null model. Finally, we discuss challenges in evaluating the importance of GCMs once they have been implemented.

Keywords

Debate, diversity, geometric constraint model, mid-domain effect, null model, species range, species richness.

INTRODUCTION

In short, the recent emphasis on the need to evaluate perceived patterns in community structure against null hypotheses is — as seen with hindsight — overdue. But legitimate enthusiasm for sound methodology must go hand in hand with the realization that null hypotheses in ecology, as elsewhere, depend on null models, and that all models make assumptions. If these assumptions are not appropriate, or create systematic biases, no amount of mathematical and statistical precision will produce biologically valid answers. (Harvey *et al.*, 1983)

The application of null models has a controversial history in ecology (Gotelli & Graves, 1996). One of the most recent of these debates centres on the mid-domain effect and geometric constraint models (GCMs; e.g. Cowell *et al.*, 2005; Hawkins *et al.*, 2005; Zapata *et al.*, 2005). These null models generate patterns of species richness in the absence of geographical gradients in ecological processes that affect richness, taking into account the naturally restricted area in which species can occur. In general, GCMs produce a central peak in species richness when species ranges are randomly placed within a bounded domain, and it has been argued that these stochastic models can in part explain such classic ecological patterns as the latitudinal gradient in species richness (Colwell & Hurtt, 1994; Willig & Lyons, 1998; Colwell & Lees, 2000).

In general, null models are designed to determine whether the observed pattern differs from expectations in the absence of some important process (Harvey *et al.*, 1983; Gotelli & Graves, 1996). In the case of GCMs, this is actually a large group of processes — all non-random processes predicted to generate gradients in species richness (Colwell & Hurtt, 1994; Colwell & Lees, 2000; Colwell *et al.*, 2004). Having decided which processes to exclude, the challenge is to design a procedure that will produce the pattern anticipated in the absence of those processes. However, a single best choice of procedure is not always obvious (Gotelli, 2000), making inferences from this type of approach more difficult. Here we suggest that at the present time GCMs suffer from three challenges in this regard: (1) uncertainty regarding the basic mode of random range placement and the consequent null pattern of richness that would be expected (i.e. a mid-domain peak or uniform richness) in the absence of the

identified processes; (2) uncertainty about how to generate the null pattern, even when the basic mode of range placement is agreed upon; and (3) uncertainty regarding how best to evaluate the importance of GCMs in explaining empirical richness patterns.

HOW SHOULD RANGES BE RANDOMLY PLACED IN THE ABSENCE OF ECOLOGICAL PROCESSES?

> The null hypothesis tested in any analysis of biogeographical data … is not that empirical patterns do not differ from random ones, but that they do not differ from patterns generated by a particular model of the world. (Colwell and Winkler, 1984)

Colwell and Hurtt (1994) originally proposed two general classes of null models for patterns of species richness generated by randomly distributed species ranges. Both of these null models involve placing a geographical range randomly onto a domain (envision throwing discs of varying size onto a mat on the floor). In the first model (their model 1), if a range overlaps a domain boundary the portion of that range that lies outside the boundary is cut off. As such, the centre of a terrestrial species' range could occur in the middle of an ocean, with only a small portion of the randomly placed range actually occurring on land. Colwell and Hurtt (1994) considered this model to represent a 'baseline against which to evaluate the effect of introducing additional assumptions in the models that follow' (p. 572). In the second type of model (their models 2 and 3, hereafter called geometric constraint models, or GCMs), if a range falls on a boundary it is picked up and thrown again until it falls entirely within the range (for more detailed descriptions of these models see Colwell & Hurtt, 1994; Sandel & McKone, 2006). These two types of null models are both logically valid null models for the random distribution of ranges in a bounded domain. However, they make very different assumptions about how the 'random' placement of ranges operates — model 1 suggests that there are no boundary constraint effects on species richness, whereas GCMs assume substantial boundary effects. As a result these different types of null model also predict very different diversity patterns — model 1 predicts that species richness will be uniformly distributed across the landscape, while GCMs predict a peak in richness at the centre of the domain (Sandel & McKone, 2006).

The existence of two logically valid null models for random range placement poses a problem for unambiguously testing the null expectation. Are there reasons to choose one model over another? Colwell and Hurtt (1994) rejected the first model based on deviations of observed data from secondary predictions of the model. Specifically, Colwell and Hurtt (1994) noted, 'If a hard-boundary version of [model 1] were realistic, we would expect a drop in mean latitudinal range and a stacking up of range limits at the northern extremes of these distributions…. Not only is there no such pattern, but mean latitudinal range actually peaks at or near the northern limit for most groups, as Stevens (1989) points out, and range limits do not tend to stack up (G. C. Stevens personal communication)'.

However, one cannot characterize a null model as being inappropriate because it does not reproduce empirically observed patterns. The purpose of a null model is to represent the observed pattern in the absence of some chosen process, and differences between the null model predictions and the data suggest that the relevant process is in fact operating. Thus, the appropriateness of a null model should be judged on its assumptions rather than on comparisons between the model output and the data. Were we to identify as valid only those null models that matched the data in all aspects, then by definition reality would always confirm the null expectation. Sandel and McKone's (2006) recent argument that these secondary patterns can be used to 'help in assessing the usefulness of the various models' suffers from the same logical fallacy. In this case, Colwell and Hurtt (1994) have argued that the fact that: (1) range edges do not 'stack up' in northern latitudes; and (2) average range area increases in northern latitudes, is evidence against the placement of ranges in accordance with model 1. However, it is clearly possible to generate these same empirical patterns by placing ranges according to model 1 while simultaneously imposing a gradient in species richness and/or average range size.

The confusion regarding the rejection of model 1 based on these secondary characteristics likely resulted because the predictions of both types of null models are contingent upon two assumptions: (1) the type of range placement; and (2) the absence of meaningful ecological gradients that affect richness. Therefore, deviations from the predictions of the model can occur either because the assumptions about the type of random range placement are wrong (as Colwell and Hurtt, 1994, assert) or because of the presence of a meaningful gradient in species richness. Since a gradient in species richness is typically what is being investigated in these types of studies, deviations from a null model that predicts an absence of gradients (or associated secondary characteristics) offer no evidence against using model 1 as a reasonable null.

Moreover, model 1 is not unique in failing to reproduce all aspects of patterns. GCMs also suffer from secondary inconsistencies with empirically observed data. These include a lack of symmetric unimodality in most observed data (Willig & Lyons, 1998; Jetz & Rahbek, 2001; Laurie & Silander, 2002; McClain & Etter, 2005), a mismatch with observed patterns of spatial turnover (Koleff & Gaston, 2001) and predicted zero values for species richness at the edges of domains (Laurie & Silander, 2002). As such, neither null model can be characterized as superior with respect to reproducing all secondary patterns.

In addition to the arguments regarding secondary characteristics, Colwell and Hurtt (1994) also suggest that GCMs may be more appropriate than model 1 based on 'theories of extinction and minimum viable population size'. Ranges cannot become arbitrarily small at domain boundaries because (*ceteris paribus*) there will be some minimum range size below which a population is unlikely to persist as a result of demographic stochasticity. While this may represent a reasonable argument against the strict implementation of model 1, it makes no distinction between GCMs and a version of model 1 that includes a minimum range size. This would result in a hybrid model, with a uniform

distribution of richness across most of the domain and a decrease in richness only near the edges. When the minimum viable range size is small compared with the domain, as would be expected for most continental domains, such a model would be mostly indistinguishable from model 1. In addition, Sandel and McKone (2006) have recently argued that model 1 (their 'truncation' model) is the more appropriate null based on the idea that species tend to occur under particular sets of environmental conditions. If the domain were somehow expanded it is possible that species at the domain's edge would encounter additional suitable habitat and have larger realized ranges. In other words, the observed range of a species that abuts the domain edge may only represent a small portion of its 'potential range' (Sandel & McKone, 2006).

Therefore at least two (more if one considers potential hybrid models) equally viable classes of null model exist for patterns of species richness in the absence of environmental gradients. To implement the first of Colwell and Hurt's (1994) null models no additional decisions need to be made. The expectation in the absence of geographical gradients is simply a uniform distribution of species richness, as has been assumed implicitly by studies that do not address the 'mid-domain effect'. However, to implement GCMs, a variety of decisions must be made, and as we summarize in the subsequent sections, these choices will affect the null expectation.

IF RANGE PLACEMENT IS BOUNDED, HOW SHOULD IT BE EXECUTED?

> Mathematical models and null models both emphasize a plurality of approaches to solving problems in community ecology. For example, a problem such as population growth in a resource-limited environment can be addressed with a variety of analytical tools, none of which provides a single 'right' answer to the question…. Similarly, many different null models can be constructed to generate community patterns in the absence of interspecific competition…. The diversity of null models has not always been appreciated. (Gotelli and Graves, 1996)

If a GCM is to be used as the null model for the random distribution of species ranges, then a number of decisions must be made during model implementation (Colwell et al., 2004). Each decision actually constitutes choosing a slightly different null model (Gotelli, 2000), and understanding which null model is appropriate to answer a given question is important. Unfortunately, in many cases there is no obviously correct choice regarding how the model should be implemented due to our limited understanding of the biology of geographical range formation, but these decisions must be made nonetheless. The aim of this section is not to provide a comprehensive review, nor in most cases to provide advice about which choices to make, but simply to illustrate the types of decisions that need to be made, the current debate regarding the most appropriate approaches to dealing with them and the potential effects of the different decisions on predicted richness patterns.

Dimensionality

For historical reasons of perception, data availability and the feasibility of analysis, the complex two-dimensional pattern of species richness on the globe has often been characterized as sets of one-dimensional gradients (Hawkins & Diniz-Filho, 2004). Mid-domain models were initially developed in this context as null models for one-dimensional latitudinal, elevational and bathymetric gradients (Colwell & Hurtt, 1994; Pineda & Caswell, 1998), and consequently the great majority (>90%) of studies examining geometric constraints have done so in only a single dimension (Colwell et al., 2004). Despite the role of latitude in inspiring GCMs, a mid-domain effect is expected for any arbitrary transect across a domain. In the absence of environmental gradients, a peak in richness should occur not just at the latitudinal mid-point of a domain, for example, but at the longitudinal mid-point, the north-west–south-east mid-point, etc. While this logic seems straightforward there is still debate in the literature over whether one dimension, two dimensions, or a combination of the two is most appropriate (Bokma et al., 2001; Diniz-Filho et al., 2002; Hawkins & Diniz-Filho, 2002; Zapata et al., 2003; Colwell et al., 2004). In addition, there are gradients that may be more naturally characterized in a single dimension such as depth, elevation, river courses, etc., and there may be reason to include three-dimensional versions of continental GCMs that incorporate variability in elevation (Colwell et al., 2004).

Range placement method

One of the first decisions to be made is how to place ranges within the bounded domain (Colwell & Hurtt, 1994; Sandel & McKone, 2006). Even in the simple one-dimensional case, a number of methods for randomly placing ranges within the domain exist (Connolly, 2005), and expanding into two dimensions complicates the matter further. The primary methods for randomly placing ranges in either one-dimensional or two-dimensional cases are: (1) range shuffling — ranges are randomly placed such that they lie entirely within the domain (the pencil box analogy of Colwell et al., 2004); (2) spreading dye — a random starting point is chosen within the domain and the range spreads until it has reached the appropriate size by randomly filling unoccupied cells adjacent to the range edge while remaining within the domain (Jetz & Rahbek, 2001); and (3) Poisson placement — where range expansion from a starting point is governed by a Poisson probability of halting expansion (Connolly, 2005). While all of these approaches are reasonable and intuitive they can produce noticeably different null expectations based on details of the underlying implementation (Connolly, 2005; Sandel & McKone, 2006).

Choosing a range size frequency distribution

The choice of an appropriate range size frequency distribution (RSFD) greatly affects the expected richness pattern of any geometric constraint model and therefore the goodness of fit to

empirical richness data (Colwell & Hurtt, 1994; Laurie & Silander, 2002; Colwell *et al.*, 2004). There are typically two choices: (1) range sizes and mid-points may be drawn from a uniform distribution of all possible combinations (i.e. all combinations occurring entirely within the domain); or (2) the RSFD may be constrained to match that of the observed data. The decision regarding which RSFD to use is a decision about how much the null model should be constrained by empirical data. Colwell *et al.* (2004, 2005) have recently argued strongly for using the observed RSFD, contending that the most appropriate null model should incorporate all aspects of biology (including the observed RSFD) while excluding only the factor of interest (i.e. non-random range placement). Others have argued that biological factors that may shape the observed RSFD may be 'smuggled in' to the null model, leading to an overestimation of the importance of geometric constraints alone (Zapata *et al.*, 2003; Connolly, 2005; Hawkins *et al.*, 2005). This debate over which RSFD is most appropriate for geometric constraints models is analogous to the debate over non-random community assembly over the past several decades. While some authors have argued that the appropriate null involves randomly filling the species-by-location matrix with the appropriate number of occupancies, others have suggested that either row sums or column sums or both must be maintained in the randomization (Colwell *et al.*, 2005). These different choices amount to the inclusion of different assumptions about processes occurring in the system (Harvey *et al.*, 1983) and can have marked impacts on the outcome and interpretation of the study (e.g. Gotelli, 2000).

Determining the domain limits

In order to conduct a GCM analysis it is necessary to define the relevant boundaries for the system being studied and the question being addressed. Colwell and Lees (2000) proposed four types of domain limits: (1) continental edges; (2) ecological regions of sub-continental scale (e.g. biomes); (3) distributional limits of the focal taxon; and (4) arbitrary boundaries (e.g. 20° N to 20° S latitude). Arguments both for (Colwell & Lees, 2000; Colwell *et al.*, 2004) and against (Koleff & Gaston, 2001; Zapata *et al.*, 2005; Kerr *et al.*, 2006) the different types of boundaries are present in the literature. However, generally speaking, the choice of domain boundaries is often not clear or intuitive, and this uncertainty is problematic given the dependency of the predicted richness peak on the exact boundary locations (Vetaas & Grytnes, 2002; McClain & Etter, 2005). This phenomenon is well illustrated by Willig and Lyons (1998) who examined the match between empirical patterns of richness in New World marsupials and GCM predictions using three different latitudinal domains: the full latitudinal extent of the New World (disagreement in peak location, poor fit), the latitudinal extent of the marsupial fauna (agreement in peak location, poor fit) and the smallest latitudinal extent to include 95% of the marsupial species (agreement in peak location, good fit). Depending on the choice of boundaries, very different conclusions may be reached about the importance of geometric constraints.

Including non-endemic species

The logic of GCMs is contingent upon the examination of species endemic to the domain (Colwell & Lees, 2000; Jetz & Rahbek, 2001; Whittaker *et al.*, 2001; Hawkins & Diniz-Filho, 2002; Colwell *et al.*, 2004; Pimm & Brown, 2004). However, the fact that non-endemic species exist and therefore must be dealt with in analyses creates confusion and ambiguity for choosing the appropriate null model. As the percentage of non-endemic species in the biota increases (e.g. biomes will have proportionally more non-endemic species than continents), this issue becomes increasingly important. Three approaches have been proposed for dealing with non-endemic species: (1) exclude them entirely from all analyses (Colwell & Lees, 2000; Jetz & Rahbek, 2001); (2) truncate the range at the domain edge and treat the portion occurring within the domain as an endemic (Diniz-Filho *et al.*, 2002; McCain, 2003); and (3) truncate the range at the domain edge, and force these non-endemic range fragments to remain attached to a domain edge during randomization (Colwell *et al.*, 2004). Colwell *et al.* (2004) have argued that this decision does not strongly affect GCM fits to the data (based solely on correlation coefficients), but it can certainly affect the general shape and magnitude of the predicted relationship (e.g. the third approach will clearly produce flatter richness peaks than the second approach), especially in cases where the number of non-endemic species is large.

Range porosity

One assumption of classic GCMs is that species are found at all localities within their geographical range. However, most species ranges are patchily occupied (Rapoport, 1982; Gaston, 2003; Hurlbert & White, 2005), and this patchiness has potential consequences for the implementation of GCMs (Hawkins *et al.* 2005). Pineda and Caswell (1998) deal with this by incorporating the actual patchiness of species distributions in the null predictions by randomizing the species-by-location matrix for point samples. However, this assumes a complete lack of meaningful range coherence, which is clearly not the case in natural systems (Rahbek & Graves, 2000; Colwell *et al.*, 2004). Alternatively, McClain and Etter (2005) varied the percentage of species randomly removed from point samples after ranges had been shuffled along a bathymetric gradient. While only at severe 'porosity' was the mid-domain peak eliminated (see also Zapata *et al.*, 2003; Colwell *et al.*, 2004), the magnitude and curvature of the predicted richness curve varied systematically with mean range porosity (McClain & Etter, 2005). Colwell *et al.* (2004) argue that range porosity is 'rarely a source of significant bias for most taxa' because ranges are generally not particularly porous and range contiguity would have to be near zero to completely remove the mid-domain richness peak. However, recent work on over 300 bird species has shown that, on average, species only occur on 50–60% of the surveys within their range (Hurlbert & White, 2005). While this level of range porosity does not eliminate the mid-domain peak, it does have substantial effects on GCM model predictions (McClain & Etter, 2005), and should thus be explicitly modelled when possible (e.g. Dunn *et al.*, 2006). In

addition, the fact that survey data reflect patterns at a relatively fine resolution with very little spatial autocorrelation remaining after accounting for environmental factors (Hurlbert & White, 2005) suggests that the rationale for applying GCMs to survey data is generally weaker.

In general, this section demonstrates that there are a large number of decisions to be made when implementing a GCM. In some cases there are strong but contradictory opinions about the most appropriate course of action. In other cases, there is simply a recognition that a decision must be made, but little suggestion for which approach, if any, is superior. Given the fairly substantial effects that some of these decisions have on the predictions of the null model (e.g. Willig & Lyons, 1998; Grytnes & Vetaas, 2002; Bellwood *et al.*, 2005; Connolly, 2005; McClain & Etter, 2005) this represents a significant challenge for the implementation and interpretation of GCMs. In other words, if it is deemed necessary to include a GCM in an analysis of species richness patterns, which GCM is it that should actually be used?

EVALUATING THE IMPORTANCE OF GEOMETRIC CONSTRAINTS

Part of the confusion over how best to evaluate GCMs stems from the historical context in which null models were developed. Community null models of the 1970s and 1980s were used to ask some version of the question, 'Do ecologically or taxonomically similar species co-occur less frequently than expected by chance?'. Such null models relied on the principle of falsification (Gotelli & Graves, 1996), categorizing empirical patterns as being either distinguishable or indistinguishable from random based on some metric of co-occurrence. The focus of GCMs, on the other hand, is on the spatial pattern of species richness, an inherently complex pattern which cannot be reduced to a single metric. As such, authors have evaluated GCMs in a number of different ways including: (1) overall model fit based on correlation (e.g. Willig & Lyons, 1998); (2) testing a plot of predicted versus observed richness values for a slope of 1 and intercept of 0 (e.g. Jetz & Rahbek, 2001; Romdal *et al.*, 2005); (3) comparison of empirical richness values to 95% confidence intervals around GCM predictions (e.g. McCain, 2003); (4) comparison of curve parameters such as curvature, height and peak location (e.g. McClain & Etter, 2005); (5) evaluating the degree to which the null model accounts for secondary predictions such as species turnover and the positions of geographical ranges (e.g. Koleff & Gaston, 2001); and (6) testing GCM predictions when potential environmental determinants of richness are controlled for (Jetz & Rahbek, 2002).

Some of these methods of evaluation are hypothesis tests in the tradition of earlier null models that ask whether the entire pattern (method 2) or some aspect of it (method 4) differs from the null expectation. Despite this tradition of falsification in community null models, Colwell *et al.* (2004) argue against using such an 'all-or-nothing' approach. This is understandable given that the complexity of the pattern being predicted allows for 'partial matching' of null model predictions. Along these lines Colwell and Lees (2000, p. 75) suggest that 'Departure from the expected richness peak, under an appropriate null model, but

not the peak itself, requires biological or historical explanation…'. Connolly *et al.* (2003) discuss and interpret patterns of fish and coral diversity in just this way, focusing on deviations from the null predictions. However, this approach assumes that GCM predictions have logical primacy (see Roughgarden, 1983, for a discussion of logical primacy and null models) over other candidate predictor variables in explaining richness patterns. This can be problematic as can be seen by considering the case where richness increases linearly across a domain as a direct result of some gradient. Since GCM predictions are by definition hump-shaped, are we to seek out ecological variables that explain the J-shaped pattern of the residuals or the simpler linear relationship (Fig. 1)? Furthermore, one could envision a scenario in which richness varies across the domain in a manner consistent with the predictions of a GCM, and yet the ranges are distributed in a distinctly non-random, perfectly nested manner (Fig. 2). While this nested distribution of ranges begs for a biological explanation (e.g. nested physiological tolerances, unimodal variation in resource availability with nested competitive abilities, etc.), a strong correlation between GCM predictions and the empirical pattern would suggest that there was nothing further to explain.

More recently, Colwell *et al.* (2004, p. E14) stated that 'there is no justification for treating [mid-domain effect] predictions as primary just because they are generated by a null model'. Instead,
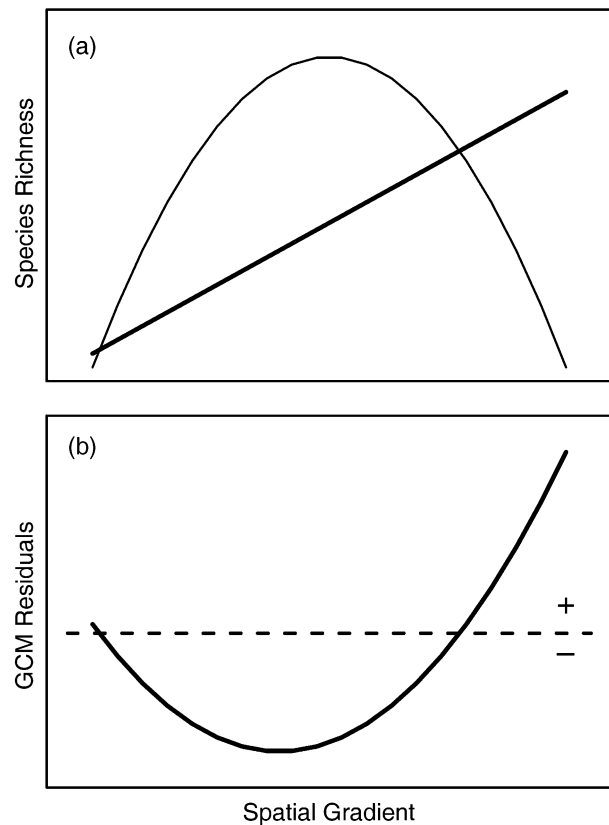

Figure 1 An example where species richness increases linearly across the domain (bold line, a). GCM predictions are hump-shaped (solid line, a). Residuals from GCM expectations are J-shaped (b).
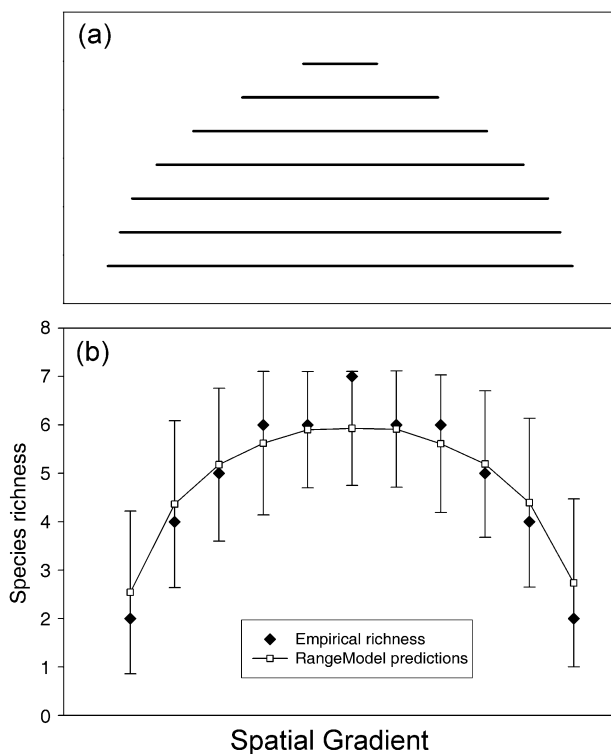
Figure 2 An example of richness varying across a one-dimensional domain in a manner consistent with the predictions of a geometric constraint model, but with ranges distributed in a distinctly non-random, perfectly nested manner. (a) Each line segment represents a species' range along a hypothetical spatial gradient. Ranges are spaced along the *y*-axis purely for illustrative purposes. (b) Empirical pattern of species richness and the pattern predicted by the random shuffling of the ranges in (a) using Colwell's (2006) RangeModel software. Error bars represent 95% confidence intervals based on 1000 simulations.

they argue that GCM predictions should be treated identically to other ecological variables and included in a multivariate framework. Thus far only a few studies examining GCMs have simultaneously considered other physical, biological or historical correlates of richness (e.g. Jetz & Rahbek, 2002; Bellwood *et al.*, 2005; Mora & Robertson, 2005; Watkins *et al.*, 2006). This approach alleviates the types of problems illustrated in Fig. 1 and thus is clearly preferable as pointed out by Colwell *et al.* (2004). However, the use of regression models evaluates the fit or partial fit of GCM predictions in a relative rather than absolute sense (Zapata *et al.*, 2003; Colwell *et al.*, 2004). Predicted richness and empirical richness can be highly correlated and yet substantially different in magnitude, and thus such analyses overestimate the importance of GCMs. Environmental correlates do not have this problem because coefficients for these variables intentionally fit to the data in order to determine the model (i.e. environmental models have free parameters whereas GCMs by definition do not). As a result of this difference, GCM predictions need to be treated differently from environmental variables and a method for integrating these two distinct types of predictors has not yet been proposed.

One area of analysis of GCMs that has been almost entirely ignored is consideration of power. Power is the probability of detecting a non-random (or non-null) pattern when one is actually present. Knowing the power of null models is important, because a null model with very low power will almost always suggest that non-random processes are not operating, even when they are operating quite strongly (e.g. Toft & Shea, 1983; Losos *et al.*, 1989; Kelt *et al.*, 1995; Gotelli, 2000). In a multivariate framework, power determines how easily the effects of different predictor variables can be distinguished from one another. For example, one-dimensional evaluations of richness patterns have inherently less power than two-dimensional analyses because of a simplification of pattern and concomitant loss of information. Collapsing variables and predictions onto a single dimension is likely to increase multicollinearity, reducing the ability to distinguish among variables. For example, were we to partition the variance in latitudinal species richness explained by net primary productivity and a one-dimensional GCM we would likely find that much of this variance cannot be uniquely ascribed to either variable, greatly confounding the interpretation of these results. Detailed analyses of power are necessary to understand the strength with which the absence of process can be inferred from GCM results. Along these lines, it is worth pointing out that if one-dimensional simplifications are used for studying GCMs then the strength of the mid-domain effect should be approximately equal for all one-dimensional axes that cross the centre of the domain (*ceteris paribus*). Thus, support for a GCM could be based on the strength of the weakest one-dimensional relationship, or perhaps on some measure of the central tendency of the distribution of strengths across all possible axes (e.g. the average correlation coefficient). If a GCM explains 80% on only one axis but little on most of the others, then the logical inference is that there is only weak evidence for a mid-domain effect, and that environmental variables that correlate with the primary axis are candidate explanatory variables for species richness. This observation raises serious questions about support for GCMs drawn from one-dimensional analyses that have focused solely on a single axis for which a strong gradient is known to exist.

CONCLUSIONS

In this paper we have highlighted three major challenges in the current application of null models for species richness. The first, and perhaps most critical, of these is that the primary published justification for choosing GCMs over the simpler type 1 null model is flawed. A null model cannot be rejected because it does not match empirical patterns. The null model must be chosen because it is appropriate for the question being asked, and deviations from that model indicate that the excluded process (or processes) is in fact operating. Unfortunately, there is no way to know which of the two general types of null model presented by Colwell and Hurtt (1994), or potential hybrids, is in fact justified by the question being asked: 'what patterns of species richness would arise under minimal biological assumptions, in a world without geographical gradients in adaptation, speciation rate, extinction rate, habitat diversity, productivity, competition,

predation, or dispersal?' (p. 570). As such, until we better understand the processes responsible for geographical range formation and how they would operate in the presence of boundaries to range expansion, the choice of null model is a matter of preference rather than objective reasoning.

The second challenge is the fact that even if GCMs are considered to be the most appropriate general type of null model, the exact null expectation is contingent on a number of choices that must be made during model implementation. These choices can lead to substantially different null predictions for species richness patterns (Bellwood *et al.*, 2005; Connolly, 2005; McClain & Etter, 2005; Sandel & McKone, 2006). Therefore, there is not just a single null prediction, but a suite of them, and there is rarely a concrete justification for choosing some realizations over others (Colwell *et al.*, 2004). One is left to make these choices arbitrarily, or one could be tempted to make choices that maximize the fit to the data. Neither of these approaches represents an objective test of a specific null hypothesis. At the minimum, researchers should detail the specific decisions made in implementing a GCM.

Finally, even if one particular method of implementing GCMs could be viewed as superior, it is not clear how best to evaluate these types of randomization-based null models and their contributions to observed patterns. There is significant disagreement in the literature over whether GCMs: (1) should be treated like more traditional null hypotheses and either accepted or rejected (Pineda & Caswell, 1998; Connolly *et al.*, 2003), or whether their predictions should be included as predictive variables in multivariate analyses (Colwell *et al.*, 2004; 2005); (2) should be given logical primacy (Colwell & Lees, 2000; Bokma *et al.*, 2001) or treated as equivalent to other variables (Colwell *et al.*, 2004; 2005); or (3) should be evaluated using secondary predictions (Koleff & Gaston, 2001; Laurie & Silander, 2002; Connolly *et al.*, 2003) or only based on species richness patterns *per se* (Colwell *et al.*, 2004). In addition, there are problems related to power and multicollinearity between null model predictions and relevant environmental variables. At the moment there is no straightforward way to deal with these issues.

In general, we have sought to highlight the fact that random range placement models can be carried out in a variety of different ways, depending upon assumptions regarding the basic nature of the null, the inclusion of non-endemic species, the definition of the domain, etc. Not all such models will predict a peak in richness at the centre of the domain — some will predict uniform richness — and among those that do predict a peak there will be differences in the specific form of the predicted richness pattern. This suggests that the 'mid-domain effect' *per se* is neither as clean nor as universal a null hypothesis as has sometimes been suggested, and indicates that substantial additional work is necessary before real inference can be drawn from GCM analyses.

REFERENCES

Bellwood, D.R., Hughes, T.P., Connolly, S.R. & Tanner, J. (2005) Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters*, **8**, 643–651.

Bokma, F., Bokma, J. & Mönkkönen, M. (2001) Random processes and geographic species richness patterns: why so few species in the north? *Ecography*, **24**, 43–49.

Colwell R.K. (2006) *RangeModel: a Monte Carlo simulation tool for assessing geometric constraints on species richness. Version 5* (available at http://viceroy.eeb.uconn.edu/rangemodel).

Colwell R.K. & Hurtt, G.C. (1994) Nonbiological gradients in species richness and a spurious Rapoport effect. *The American Naturalist*, **144**, 570–595.

Colwell, R.K. & Lees, D.C. (2000) The mid-domain effect: geometric constraints on the geography of species richness. *Trends in Ecology & Evolution*, **15**, 70–76.

Colwell, R.K. & Winkler, D.W. (1984) A null model for null models in biogeography. *Ecological communities: conceptual issues and the evidence* (ed. by D.R. Strong, D. Simberloff, L.G. Abele and A.B. Thistle), pp. 344–359. Princeton University Press, Princeton, NJ.

Colwell, R.K., Rahbek, C. & Gotelli, N.J. (2004) The mid-domain effect and species richness patterns: what have we learned so far? *The American Naturalist*, **163**, E1–E23.

Colwell, R.K., Rahbek, C. & Gotelli, N.J. (2005) The mid-domain effect: There's a baby in the bathwater. *The American Naturalist*, **166**, E149–E154.

Connolly, S.R. (2005) Process-based models of species distributions and the mid-domain effect. *The American Naturalist*, **166**, 1–11.

Connolly, S.R., Bellwood, D.R. & Hughes, T.P. (2003) Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology*, **84**, 2178–2190.

Diniz-Filho, J.A.F., de Sant'Ana, C.E.R., de Souza, M.C. & Rangel, T.F.L.V.B. (2002) Null models and spatial patterns of species richness in South American birds of prey. *Ecology Letters*, **5**, 47–55.

Dunn, R.R., Colwell, R.K. & Nilsson, C. (2006) The river domain: why are there more species halfway up the river? *Ecography*, **29**, 251–259.

Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.

Gotelli, N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology*, **81**, 2606–2621.

Gotelli, N.J. & Graves, G.R. (1996) *Null models in ecology*. Smithsonian Institution Press, Washington, DC.

Grytnes, J.A. & Vetaas, O.R. (2002) Species richness and altitude: a comparison between null models and interpolated plant species richness along the Himalayan altitudinal gradient, Nepal. *The American Naturalist*, **159**, 294–304.

Harvey, P.H., Colwell, R.K., Silvertown, J.W. & May, R.M. (1983) Null models in ecology. *Annual Review of Ecology and Systematics*, **14**, 189–211.

Hawkins, B.A. & Diniz-Filho, J.A.F. (2002) The mid-domain effect cannot explain the diversity gradient of Nearctic birds. *Global Ecology and Biogeography*, **11**, 419–426.

Hawkins, B.A. & Diniz-Filho, J.A.F. (2004) 'Latitude' and geographic patterns in species richness. *Ecography*, **27**, 268–272.

Hawkins, B.A., Diniz-Filho, J.A.F. & Weis, A.E. (2005) The mid-domain effect and diversity gradients: is there anything to learn? *The American Naturalist*, **166**, E140–E143.

Hurlbert, A.H. & White, E.P. (2005) Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. *Ecology Letters*, **8**, 319–327.

Jetz, W. & Rahbek, C. (2001) Geometric constraints explain much of the species richness pattern in African birds. *Proceedings of the National Academy of Sciences USA*, **98**, 5661–5666.

Jetz, W. & Rahbek, C. (2002) Geographic range size and determinants of avian species richness. *Science*, **297**, 1548–1551.

Kelt, D.A., Taper, M.L. & Meserve, P.L. (1995) Assessing the impact of competition on community assembly — a case-study using small mammals. *Ecology*, **76**, 1283–1296.

Kerr, J.T., Perring, M. & Currie, D.J. (2006) The missing Madagascan mid-domain effect. *Ecology Letters*, **9**, 149–159.

Koleff, P. & Gaston, K.J. (2001) Latitudinal gradients in diversity: real patterns and random models. *Ecography*, **24**, 341–351.

Laurie, H. & Silander, J.A. (2002) Geometric constraints and spatial pattern of species richness: critique of range-based null models. *Diversity and Distributions*, **8**, 351–364.

Losos, J.B., Naeem, S. & Colwell, R.K. (1989) Hutchinsonian ratios and statistical power. *Evolution*, **43**, 1820–1826.

McCain, C.M. (2003) North American desert rodents: a test of the mid-domain effect in species richness. *Journal of Mammalogy*, **84**, 967–980.

McClain, C.R. & Etter, R.J. (2005) Mid-domain models as predictors of species diversity patterns: bathymetric diversity gradients in the deep sea. *Oikos*, **109**, 555–566.

Mora, C. & Robertson, D.R. (2005) Causes of latitudinal gradients in species richness: a test with fishes of the Tropical Eastern Pacific. *Ecology*, **86**, 1771–1782.

Pimm, S.L. & Brown, J.H. (2004) Domains of diversity. *Science*, **304**, 831–833.

Pineda, J. & Caswell, H. (1998) Bathymetric species-diversity patterns and boundary constraints on vertical range distributions. *Deep-Sea Research Part II — Topical Studies in Oceanography*, **45**, 83–101.

Rahbek, C. & Graves, G.R. (2000) Detection of macro-ecological patterns in South American hummingbirds is affected by spatial scale. *Proceedings of the Royal Society of London Series B*, **267**, 2259–2265.

Rapoport, E.H. (1982) *Areography: geographical strategies of species*. Pergamon Press, Oxford.

Romdal, T.S., Colwell, R.K. & Rahbek, C. (2005) The influence of band sum area, domain extent, and range sizes on the latitudinal mid-domain effect. *Ecology*, **86**, 235–244.

Roughgarden, J. (1983) Competition and theory in community ecology. *The American Naturalist*, **122**, 583–601.

Sandel, B.S. & McKone, M.J. (2006) Reconsidering null models of diversity: do geometric constraints on species ranges necessarily cause a mid-domain effect? *Diversity and Distributions*, **12**, 467–474.

Stevens, G.C. (1989) The latitudinal gradient in geographical range — how so many species coexist in the tropics. *The American Naturalist*, **133**, 240–256.

Toft, C.A. & Shea, P.J. (1983) Detecting community-wide patterns — estimating power strengthens statistical inference. *The American Naturalist*, **122**, 618–625.

Vetaas, O.R. & Grytnes, J.A. (2002) Distribution of vascular plant species richness and endemic richness along the Himalayan elevation gradient in Nepal. *Global Ecology and Biogeography*, **11**, 291–301.

Watkins, J.E., Jr, Cardelus, C., Colwell R.K. & Moran R.C. (2006) Species richness and distribution of ferns along an elevational gradient in Costa Rica. *Journal of Botany*, **93**, 73–83.

Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–370.

Willig, M.R. & Lyons, S.K. (1998) An analytical model of latitudinal gradients of species richness with an empirical test for marsupials and bats in the New World. *Oikos*, **81**, 93–98.

Zapata, F.A., Gaston, K.J. & Chown, S.L. (2003) Mid-domain models of species richness gradients: assumptions, methods and evidence. *Journal of Animal Ecology*, **72**, 677–690.

Zapata, F.A., Gaston, K.J. & Chown, S.L. (2005) The mid-domain effect revisited. *The American Naturalist*, **166**, E144–E148.

BIOSKETCHES

Craig McClain is a post-doctoral fellow at the Monterey Bay Aquarium Research Institute. His research interests are characterizing and understanding the complex biodiversity and body size patterns of deep-sea organisms.

Ethan White is currently an NSF Postdoctoral Fellow in Bioinformatics. His research interests are primarily macroecological, combining theory with empirical analysis to address questions about emergent spatiotemporal patterns of species diversity, abundance and body size as well as the processes that drive them.

Allen Hurlbert is broadly interested in macroecology, with specific research foci including both local and global-scale patterns of species richness, turnover in species composition in space and time, the structure of species geographical ranges and avian ecology.

Editor: Pablo Marquet