# Sa cal de alle e il del a al

## N d la J.G ell a dWe e Ul d de

N. J. Gotelli, Dept of Biolog, Uni. of Vermont, Burlington, VT 05405, USA. W. Ulrich (ulrich @umk.pl), Dept of Animal Ecolog, Nicolaus Copernicus Uni. in Toru\_nGagarina 9, PL-87-100 Toru\_nPoland.

is review identifies several important challenges in null model testing in ecology: 1) developing randomization algorithms that generate appropriate patterns for a specified null hypothesis; these randomization algorithms stake out a middle ground between formal Pearson–Neyman tests (which require a fully-specified null distribution) and specific process-based models (which require parameter values that cannot be easily and independently estimated); 2) developing metrics that specify a particular pattern in a matrix, but ideally exclude other, related patterns; 3) avoiding classification schemes based on idealized matrix patterns that may prove to be inconsistent or contradictory when tested with empirical matrices that do not have the idealized pattern; 4) testing the performance of proposed null models and metrics with artificial test matrices that contain specified levels of pattern and randomness; 5) moving beyond simple presence—absence matrices to incorporate species-level traits (such as abundance) and site-level traits (such as habitat suitability) into null model analysis; 6) creating null models that perform well with many sites, many species pairs, and varying degrees of spatial autocorrelation in species occurrence data. In spite of these challenges, the development and application of null models has continued to provide valuable insights in ecology, evolution, and biogeography for over 80 years.

A null model is a pattern generating model that is based on randomi ation of ecological data or random sampling from a kno n or imagined distribution.

e null model is designed ith respect to some ecological or e olutionar process of interest'. (Gotelli and Graves 1996)

From its origins in the analysis of species/genus ratios (Järvinen 1982), there is a long history of using null models to analyze patterns and test hypotheses in ecology, evolution and biogeography (Harvey et al. 1983). Although the general controversy in the 1970s over null models and competition has died down (Gotelli and Graves 1996), there are still many disputed aspects of testing and implementing null models. In this paper, we review some of the more recent challenges and controversial issues in the implementation and interpretation of null models in ecology. We focus primarily on the use of null models in biogeography, ecology, and macroecology.

### H ♥e e adc a II delaal

Classical Pearson–Neyman hypothesis testing (Graves 1978) addresses the dichotomy between a null hypothesis ( $H_0$ ) and its alternative ( $H_1$ ). If these hypotheses are mutually exclusive and collectively exhaustive, then the probability that  $H_0$  is true, given the data ( $P(H_0|\text{data})$ ), is  $P(H_0) = 1 - P(H_1)$ .

# Statistical null hypotheses

Classical Fisherian hypothesis testing based on equiprobability and/or on parametric distributions

### Null models

Pattern oriented

Increased ecological realism using an increasing number of constraints

Matrix constraints may account for: matrix fill, marginal totals, abundances ecological gradients, niche differences, phylogeny.

Randomization of observed data to obtain the null distribution

#### Mechanistic models

Process oriented

Theory of island biogeography

Multispecies Multispecies based

Metapopulation models

Neutral theory Detailed mechanistic models

Increasing people gized realism

Increasing ecological realism

Generation of patterns from first principles to obtain the null distribution

leads to excessive type II errors (Grant and Abbott 1980, Colwell and Winkler 1984). For this reason, Presley et al. (2010) recently advocated the use of the equiprobable–equiprobable model for testing for patterns of species distributions. However, the poor performance of this algorithm in the context of nestedness (Ulrich and Gotelli 2007a) and species co-occurrence (Gotelli 2000, Ladau 2008), suggests there is a real danger in overestimating the frequency of significant patterns by taking such a liberal approach.

Recently Kullback-Leibler information-based model choices have become popular as complementary approaches to classical hypothesis testing (Akaike 1973, Burnham and Anderson 2002). Information criteria assign probabilities to competing models with di erent numbers of free parameters and thus allow for a ranking of models from best to worst (Anderson 2008). In the context of null model analysis, we might ask whether information criteria are capable of quantifying the information content of di erently constrained null models. However, a simplistic use of information criteria is problematic because we cannot equate the number of null model constraints with the number of free parameters necessary for calculating information metrics. Moreover, null models cannot simply be ranked additively by the number of constraints they contain, but should instead be, nformge

metbiologl hyp0(e)0(aticm the eqirmignumpns. H,d thus qirm)-8.1(rmanTJ  $T^*$  .00291w [(addce of o)1by ohma(oblkstingith the ber ompetiar(e)0(dth theormation criteria isug ltld ghlo)1nsp

between the formal definition of a pattern and the working definition based on the index.

e history of the nestedness concept (Almeida-Neto et al. 2007, 2008, Ulrich et al. 2009) is a good example of how this mismatch of pattern and metric can cause confusion. Patterson and Atmar (1986) originally defined nestedness as 'that the species comprising a depauperate fauna should constitute a proper subset of those in richer faunas'. is definition of nestedness focuses on the species composition among sites, but does not consider the site composition among species. A proper nestedness metric should measure the

things, critics objected to the random placement of ranges within a bounded range because real ranges reflect species interactions with the environment (Hawkins and Diniz-Filho 2002).

However, the MDE served as a very e ective null model because it excluded geographical gradients in historical e ects or contemporary climate and demonstrated that species richness gradients can arise entirely from simple geometric constraints (Colwell et al. 2004). ese constraints are a realistic alternative to the implicit null hypothesis in many correlative studies where species have no dispersal constraints and can occur in any grid cell within a domain that has appropriate climatic conditions (Gotelli et al. 2009). More recent analyses have used the range cohesion e ect embodied in MDE in stochastic models that also include environmental e ects (Rahbek et al. 2007).

For very large matrices, and for matrices sampled at large spatial scales, the homogeneity assumption cannot be justified and traditional null models should be applied with caution. Recently Navarro-Alberto and Manly (2009) showed that any di erence either in occurrence probabilities of species across sites (non-uniform column degree distributions) or species (non-uniform row degree distributions) causes some degree of spatial autocorrelation. Null models that do not correct for autocorrelation may therefore too often point to non-randomness. To our knowledge, the e ect of autocorrelation on matrix structure has not been studied systematically, although Ulrich (2004) demonstrated that a neutral model with limited spatial dispersal can generate binary presence-absence matrices that are statistically segregated. Autocorrelation in species occurrences should cause a tendency towards matrix compartments with regions of higher and lower fill. For large matrices, even very small degrees of autocorrelation will be identified as being significant (Burnham and Anderson 2002).

A second type of autocorrelation is the repetition og]TJ T\* .1196

tionroRTsvn(l/y catric 2016 T\$n 822aly.12108w)16 (eTD) T004ese .0.0422166 3h 8220019 .9yH0] T(b) T2



- Anderson, D. R. 2008. Model based inference in the life sciences. Springer.
- Atmar, W. and Patterson, B. D. 1993. e measure of order and disorder in the distribution of species in fragmented habitat.
   Oecologia 96: 373–382.
- Atmar, W. and Patterson, B. D. 1995. e nestedness temperature calculator: a Visual Basic program, including 294 presence—absence matrices. AICS Res., Univ. Park, NM and Field Museum, Chicago <a href="http://aics-research.com/nestedness/tempcalc.html">http://aics-research.com/nestedness/tempcalc.html</a>>.
- Bacallado, J. J. 1976. Notas sobre la distribucion y evolucion de la avifauna Canaria. In: Kunkel, G. (ed.), Biogeography and ecology in the Canary Islands. Junk, pp. 413–431.
- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecol. Biogeogr. 19: 134–143.
- Brualdi, R. A. and Sanderson, J. G. 1999. Nested species subsets, gaps, and discrepancy. Oecologia 119: 256–264.
- Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd ed. – Springer.
- Chao, A. et al. 2010. Phylogenetic diversity measures based on Hill numbers. – Phil. Trans. R. Soc. B 365: 3599–3609.
- Clark, P. J. and Evans, F. C. 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. – Ecology 35: 445–453.
- Colwell, R. K. and Winkler, D. W. 1984. A null model for null models in biogeography. – In: Strong, D. et al. (eds), Ecological communities: conceptual issues and the evidence. Princeton Univ. Press, pp. 344–359.
- Colwell, R. K. and Lees, D. C. 2000. e mid-domain e ect: geometric constraints on the geography of species richness. Trends Ecol. Evol. 15: 70–76.
- Colwell, R. K. et al. 2004. e mid-domain e ect and species richness patterns: what have we learned so far? Am. Nat. 163: E1–E23.
- Colwell, R. K. et al. 2005. e mid-domain e ect: there's a baby in the bathwater. Am. Nat. 166: E149–E154.
- Connor, E. H. and Simberlo , D. 1979. e assembly of species communities: chance or competition? Ecology 60: 1132–1140.
- Connor, E. F. et al. 2000. Individuals–area relationships: the relationship between animal population density and area. – Ecology 81: 734–748.
- Diamond, J. M. 1975. Assembly of species communities. In: and Anderson, E\* -.25Tw367[(163676TJ621445-453.)]T[(.-er6238-1.162TD12888)433D2888