



On the detection of metacommunity structure

Steven J. Presley¹

Received: 18 June 2019 / Accepted: 12 February 2020 / Published online: 16 March 2020
© Akadémiai Kiadó Zrt. 2020

The metacommunity concept is an important approach in community ecology that explicitly recognizes that the composition of local communities and spatial variation in community composition, are influenced by both local (e.g., biotic interactions, environmental tolerances) and regional (e.g., dispersal, habitat fragmentation) factors (Leibold et al. 2004). Increased interest in broad-scale patterns of biodiversity and better understanding of the contributions of regional factors to local community assembly have changed our perception of the community concept from a localized group of interacting species to one in which the spatial distributions of species have become a greater focus for understanding patterns of co-occurrence and local biodiversity (Ricklefs 2008).

Although estimating the relative importance of ecological mechanisms is the primary goal of ecology, it is often necessary first to identify ecological patterns to facilitate the identification of likely mechanisms that may give rise to observed structures. Metacommunity structure is an emergent property that describes the relationships among species distributions along an environmental gradient (Leibold and Mikkelsen 2002). Many idealized metacommunity structures have been described, with each structure representing a distinct prediction about relationships between ecological mechanisms and species distributions. For example, Clementsian structures are characterized by coincident range boundaries that delimit sets of communities with similar species composition that are distinctive from other such sets; Gleasonian structures arise from species-specific responses to environmental variation; strong competition and priority effects may result in checkerboard patterns among ecologically similar species that have overlapping geographical ranges; or variation among species in dispersal ability,

habitat specialization, or abiotic tolerance may give rise to nested subsets.

Recently, a set of critiques have highlighted potential weaknesses in the ability of the elements of metacommunity structure (EMS) to effectively identify metacommunity structures and the importance of gradient selection or identification for this approach (Schmera et al. 2018). These critiques include (1) the viability of researcher-defined gradients to evaluate EMS, (2) nested subsets and range turnover not being opposite ends of a continuum, (3) lack of correspondence between negative coherence and checkerboards, and (4) statistical implications of using a series of tests to identify structures.

Reciprocal averaging and identifying response gradients

Much of the confusion about the implementation or interpretation of EMS is related to the importance of the gradient along which the elements are evaluated. Reciprocal averaging (also called correspondence analysis) is used to order sites and species in a presence/absence matrix by simultaneously optimizing the proximity of species with similar distributions and the proximity of communities with similar species compositions, allowing species occurrences to define the latent environmental gradient that is important to species distributions (Gauch 1982). This provides factor scores for both sites and species with respect to the same gradient, an advantage over ordination methods that can only order sites or species, but not both (e.g., nonmetric multidimensional scaling, NMDS; principal component analysis, PCA). Unlike NMDS or PCA, reciprocal averaging does not maximize the variation explained along each orthogonal axis, rather correspondence between matrix orientations is maximized (Gauch 1982). For site by species matrices, this is the degree to which sites and species could simultaneously be ordered perfectly (i.e., the degree of correspondence achieved).

A similar ordination problem manifests for researcher-defined gradients (Dallas et al. 2016): one can order sites

✉ Steven J. Presley
steven.presley@uconn.edu

¹ Department of Ecology and Evolutionary Biology, Center for Environmental Sciences and Engineering, Institute of the Environment, University of Connecticut, 3107 Horsebarn Hill Road, Storrs, CT 06269-4210, USA

along an environmental gradient (elevation, temperature, pH), but how does one objectively choose the order of species along that same gradient? Even if one can find a defensible method to order species along the same gradient, a problem arises in the attempt to evaluate statistical significance. A researcher can order empirical data along an environmental gradient, but matrices generated to create a null distribution for statistical evaluation have no environmental gradient along which to order sites or species. Schmera et al. (2018) correctly note the limitations of user-defined gradients for EMS, which likely is not statistically or methodologically viable. The mismatch between the gradient used to calculate empirical values and the gradient used to create distributions of randomly generated values leads to elevated type I or II error rates (Presley et al. 2019).

The primacy of coherence

EMS attempts to distinguish among many idealized structures using three characteristics of species distributions: coherence, range turnover, and range boundary clumping (Leibold and Mikkelsen 2002). Coherence determines if species distributions are molded by the same environmental gradient. However, responses to that gradient may differ among species (i.e., locations and extents of ranges differ among species), giving rise to a host of possible metacommunity structures. Coherence is characteristic of 12 metacommunity structures that are distinguished by analysis of range turnover and range boundary clumping (Presley et al. 2010). In contrast, checkerboards are thought to exhibit negative coherence (i.e., groups of species with ranges molded by multiple independent gradients), and random structures exhibit non-coherence (i.e., species ranges are not molded by shared environmental gradients).

In cases of non-coherence, the ordination does not effectively represent species distributions along a shared environmental gradient. Consequently, estimates of range turnover and range boundary clumping do not reflect the concepts they embody (turnover of species ranges along a gradient or clumping of range boundaries along a gradient) and are not evaluated for non-coherent metacommunities (Leibold and Mikkelsen 2002). The same metacommunity may exhibit different structure along different gradients, with gradients often representing structure at different spatial scales (Presley et al. 2009; López-González et al. 2012). Changing the gradient changes the structure. Schmera et al. (2018) state that they “see no strong theoretical support for the priority of coherence in metacommunity structuring”. This ignores the importance of the gradient of analysis or the need to determine if species distributions are molded by a common gradient. Further, Schmera et al. (2018) conclude that latent gradients identified via reciprocal averaging are not “real” environmental gradients,

despite providing no supporting evidence. This seems unlikely as many dozens of studies have evaluated EMS along many 100 s of latent environmental gradients for which supplementary analyses were able to identify underlying environmental gradients (see supplementary materials of Presley et al. 2019). In general, latent gradients have been associated with known important environmental gradients (e.g., disturbance, habitat type, elevation, latitude, water depth, salinity, temperature, precipitation, island size, habitat patch size, host identity), with results that are consistent with contemporary understanding of the ecology of the system.

Range turnover versus nestedness

Schmera et al. (2018) state that “the turnover test cannot always detect nested pattern, because turnover and nestedness are not necessarily the opposite endpoints of a continuum”. Schmera et al. (2018) correctly note that the EMS definition of turnover is not the same as other definitions (of which there are > 20; Tuomisto 2010). Uniquely, EMS evaluates characteristics of species ranges along a specific gradient, using a species range as a unit (turnover cannot happen within the range of a species), whereas other measures use species occurrences as units. There are two problems with the evaluation of the relationship between nested subsets and range turnover. First, analyses were conducted along different gradients: nested subsets along a richness-occurrence gradient and range turnover along (presumably) the primary axis of reciprocal averaging. Second, one would not evaluate range turnover without first considering coherence, as the metric does not represent the concept of range turnover along a shared gradient if it is not established that the ranges of species are molded by the same gradient (Leibold and Mikkelsen 2002). Previous critiques of EMS (Ulrich and Gotelli 2012) also failed to recognize that the same gradient must be used to effectively compare measures of range turnover to those of nested subsets, and that range turnover is a metric that requires its associated axis to exhibit coherence. The 10,000 randomly generated matrices produced by Schmera et al. (2018) were not evaluated for coherence prior to testing range turnover, nor did they report the number of random matrices consistent with positive coherence (though 3.03% exhibited negative coherence, an acceptable type I error rate).

Range boundary clumping versus co-occurrence

Range boundary clumping measures the degree to which range boundaries are clumped along a gradient. Ulrich and Gotelli (2012) interpreted the fact that range boundary clumping was weakly associated with measures of species

co-occurrence as evidence that the metric performs poorly. This comparison has two major faults. First, co-occurrence metrics are independent of any gradient: individual occurrences are the focal units and sites may occur in any order. Therefore, the comparison used different gradients as range boundary clumping is associated with a specific gradient. Second, the metrics embody independent concepts. Imagine an elevational gradient along which species ranges are continuous. Species A occurs from 0 to 500 m and Species B from 50 to 450 m; these species exhibit positive co-occurrence but have no shared range boundaries. In contrast, Species C occurs from 500 to 1000 m, exhibiting negative co-occurrence with both Species A and B, but sharing one range boundary with Species B. Two species could have identical elevational distributions with two shared range boundaries, but co-occurrence could be positive or non-significant if each species had multiple embedded absences. Consequently, one would expect range boundary clumping to be only weakly associated with measures of co-occurrence.

Negative coherence and checkerboards

Checkerboards have been used as a metaphor for multiple patterns (Connor et al. 2013; Schmera et al. 2018; Presley et al. 2019). Checkerboards originally described geographically interspersed patterns of mutual exclusion by ecologically similar species (MacArthur et al. 1972). Subsequently, Leibold and Mikkelsen (2002) expanded this concept to entire metacommunities by adding the criterion that distributions of each mutually exclusive pair should be independent from other such pairs. Unfortunately, this definition is nearly identical to that of random metacommunity structure (i.e., non-coherence). Indeed, the only difference between random and checkerboard metacommunity structures is that in checkerboard structures each species has one mutually exclusive association, with all other associations being random. Consequently, the dominant mechanism for both of these proposed structures is randomness, and the developed methodology cannot distinguish between random and checkerboard structures (Presley et al. 2019). The idea that checkerboard structures *sensu* Leibold and Mikkelsen (2002) can be detected via analyses of coherence should be abandoned.

Series of tests and error rates

Schmera et al. (2018) calculate the likelihood of observing particular metacommunity structures by chance based on the required series of test results, and correctly note that these likelihoods make type II errors more likely for some structures than others. However, the single test used for nested subsets has unequal likelihood of observing each potential

structure by chance: 0.05 for nested and 0.95 for non-nested structures. Is it more problematic that Clementsian and Gleasonian structures have unequal rates of occurring by chance (0.00002 and 0.00059, respectively)? Conservative analysis shows that 9–13% of systems are nested (Ulrich and Gotelli 2007), given a 5% type I error rate, does this mean that about half of the systems identified as nested represent errors and that we should abandon this approach because it is only good at effectively identifying non-nested structures? I am playing the part of *Advocatus Diaboli* here to establish context for this potential concern. There is no reason to expect that all metacommunity structures should have an equal chance of occurring, because the constraints for some are much greater than for others. A review of the literature (Presley et al. 2019) shows that the frequency of reported empirical structures are not associated with the combined probabilities of type I errors: Clementsian structures are most common (28% of all structures) despite having the lowest probability of occurring by chance, quasi-structures occur only 60% as often as their counterparts despite being 19 times more likely to occur by chance. In all likelihood, responses to the environment are sufficiently strong to minimize issues related to unequal detection error rates among structures.

Concluding statement

The framework for EMS is not as straight forward as the analysis makes it appear, decisions made by users can have unexpected implications. Critical evaluations of EMS (Ulrich and Gotelli 2012; Schmera et al. 2018; Presley et al. 2019) have been useful in highlighting potential conceptual and methodological weaknesses and pitfalls that users need to understand to make informed decisions when conducting analyses.

Users should be mindful of the potential for over interpretation of EMS results and remember that these analyses are designed specifically to detect patterns that represent the relationships among species distributions along environmental gradients (i.e., to identify the best fit metacommunity structure). Additional analyses and additional data are required to identify processes that contribute to these patterns (e.g., Leibold et al. 2004; Cottenie 2005). Despite attempts to match metacommunity level processes (environmental filtering, dispersal, interspecific interactions) to metacommunity structures (Meynard et al. 2013), multiple processes can give rise to the same structure (Ovaskainen et al. 2019) or the same process can give rise to multiple structures (Presley et al. 2012). Moreover, it is likely that a combination of mechanisms (e.g., dispersal, habitat specialization, competition, predation, tolerance to abiotic conditions) contribute to the structure of each metacommunity, with the relative importance of these mechanisms differing

among metacommunities or through time. Identification of the importance of ecological processes associated with each metacommunity requires comprehensive analysis (Ovaskainen et al. 2019) and cannot be determined by the identification of pattern alone.

Acknowledgements SJP was supported by the National Science Foundation (DEB-1239764 and DEB-1546686) and by the Center for Environmental Sciences and Engineering at the University of Connecticut. C. Higgins provided comments on an earlier version of this manuscript.

References

- Connor, E. F., Collins, M. D., & Simberloff, D. (2013). The checkered history of checkerboard distributions. *Ecology*, *94*, 2403–2414.
- Cottenie, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, *8*, 1175–1182.
- Dallas, T. A., Kramer, A. M., Zokan, M., & Drake, J. M. (2016). Ordination obscures the influence of environment on plankton metacommunity structure. *Limnology and Oceanography Letters*, *1*, 54–61.
- Gauch, H. G. (1982). *Multivariate analysis in community ecology*. Cambridge: Cambridge University Press.
- Leibold, M. A., Holyoak, M., Mouquet, M., Amarasekare, P., Chase, J. M., Hoopes, M. F., et al. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, *7*, 601–613.
- Leibold, M. A., & Mikkelsen, G. M. (2002). Coherence, species turnover, and boundary clumping: Elements of meta-community structure. *Oikos*, *97*, 237–250.
- López-González, C., Presley, S. J., Lozano, A., Stevens, R. D., & Higgins, C. L. (2012). Metacommunity analysis of Mexican bats: Environmentally mediated structure in an area of high geographic and environmental complexity. *Journal of Biogeography*, *39*, 177–192.
- MacArthur, R. H., Diamond, J. M., & Karr, J. R. (1972). Density compensation in island faunas. *Ecology*, *53*, 330–342.
- Meynard, C. N., Lavergne, S., Boulangeat, I., Garraud, L., Van Es, J., Mouquet, N., et al. (2013). Disentangling the drivers of metacommunity structure across spatial scales. *Journal of Biogeography*, *40*, 1560–1571.
- Ovaskainen, O., Rybicki, J., & Abrego, N. (2019). What can observational data reveal about metacommunity processes? *Ecography*, *42*, 1877–1886.
- Presley, S. J., Cisneros, L. M., Patterson, B. D., & Willig, M. R. (2012). Vertebrate metacommunity structure along an extensive elevational gradient in the tropics: A comparison of bats, rodents and birds. *Global Ecology and Biogeography*, *21*, 968–976.
- Presley, S. J., Higgins, C. L., López-González, C., & Stevens, R. D. (2009). Elements of metacommunity structure of Paraguayan bats: Multiple gradients require analysis of multiple axes of variation. *Oecologia*, *160*, 781–793.
- Presley, S. J., Higgins, C. L., & Willig, M. R. (2010). A comprehensive framework for the evaluation of metacommunity structure. *Oikos*, *119*, 908–917.
- Presley, S. J., Mello, J. H. F., & Willig, M. R. (2019). Checkerboard metacommunity structure: An incoherent concept. *Oecologia*, *190*, 323–331.
- Ricklefs, R. E. (2008). Disintegration of the ecological community. *American Naturalist*, *172*, 741–750.
- Schmera, D., Podani, J., Botta-Dukát, Z., & Erős, T. (2018). On the reliability of the elements of metacommunity structure framework for separating idealized metacommunity patterns. *Ecological Indicators*, *85*, 853–860.
- Tuomisto, H. (2010). A diversity of beta diversities: Straightening up a concept gone awry. Part I. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*, 2–22.
- Ulrich, W., & Gotelli, N. J. (2007). Null model analysis of species nestedness patterns. *Ecology*, *88*, 1824–1831.
- Ulrich, W., & Gotelli, N. J. (2012). Pattern detection in null model analysis. *Oikos*, *122*, 2–18.