

Ecology, 91(7), 2010, pp. 1981–1983
 © 2010 by the Ecological Society of America

On beta diversity decomposition: Trouble shared is not trouble halved

CARLO RICOTTA¹

Department of Plant Biology, University of Rome “La Sapienza,” Piazzale Aldo Moro 5, 00185 Rome, Italy

The concept of beta diversity was first introduced by Whittaker (1960, 1972) as the proportion by which the pooled species richness in a set of plots of some arbitrary size exceeds the average richness of species in individual plots. According to Whittaker’s multiplicative approach, for species presence and absence data, beta diversity is computed as $\beta = \gamma/\bar{\alpha}$, where $\bar{\alpha}$ is the average alpha diversity of single plots. In an alternative approach originally proposed by McArthur et al. (1966) and recently “rediscovered” by Lande (1996), beta diversity is computed additively as $\beta = \gamma - \bar{\alpha}$. In both cases, as beta diversity increases, individual plots differ more markedly from one another and sample a smaller proportion of the species occurring in the region (Koleff et al. 2003).

Though initially developed for dealing with species richness, both diversity decomposition methods can be usually extended to traditional diversity indices, like the Shannon entropy or the Simpson diversity, that are based on species relative abundances. Unfortunately, unlike the alpha and gamma components of diversity, beta diversity is not a genuine measure of “compositional diversity” (Ricotta 2007). Rather, as shown by Vellend (2001), it is conceptually closer to a measure of multivariate plot-to-plot dissimilarity. This ambiguity in the very meaning of beta diversity has ensured that its measurement remains “capricious” (sensu Sarkar and Margules 2002).

Accordingly, a number of alternative methods for measuring beta diversity have been proposed by several

authors. For instance, the classical reviews by Wilson and Shmida (1984) and Koleff et al. (2003) list 8 and 24 different measures of beta diversity, respectively. Izsak and Price (2001) suggested that the mean of the dissimilarities among plots may be used as a genuine measure of beta differentiation (see also Whittaker 1972, Legendre et al. 2005). Legendre et al. (2005) also showed that the variance of the species \times plots matrix is another meaningful measure of beta diversity. More recently, Anderson et al. (2006) proposed measuring beta diversity as the average dissimilarity from individual plots to their group centroid in multivariate space, while Ricotta and Burrascano (2009) used instead the mean asymmetric dissimilarity between the individual plots and the pooled set of plots.

All these measures have the merit of summarizing the variability in species composition among sampling units based on distinct objectives and motivations; from a statistical viewpoint, by reducing a multivariate data set of high dimension like plot-to-plot species heterogeneity into a single index, information is necessarily lost, and there is no ideal function capable of uniquely characterizing all aspects of beta diversity.

In this framework, Jost (2007) went a step further in developing the mathematical foundation for multiplicative partitioning of species diversity. Jost (2006, 2007) noted that if the ratio $\gamma/\bar{\alpha}$ is computed directly from traditional diversity indices, it necessarily approaches unity when diversity is high, apparently indicating complete similarity, even if the plots sampled are completely differentiated (no species in common). Jost also showed that for the Simpson index $1 - \sum_{i=1}^S p_i^2$ (where p_i is the relative abundance of species i and S is

Manuscript received 22 January 2009; revised 11 May 2009; accepted 12 May 2009. Corresponding Editor: A. M. Ellison. For reprints of this Forum, see footnote 1, p. 1962.

¹ E-mail: carlo.ricotta@uniroma1.it

total species richness), the beta produced by additive partitioning necessarily approaches zero when diversity is high, apparently indicating no differentiation, even if the plots do not share any species. This is because the existing definitions of multiplicative and additive beta diversity produce a beta with a hidden dependence on alpha.

Jost (2006) suggested that a solution consists in converting alpha and gamma diversities to their “equivalent number of species” or “numbers equivalent” D before taking the ratio between gamma diversity and average alpha diversity such that $D_\beta = D_\gamma / D_\alpha$.

As shown by Jost (2006), for all diversity indices that are functions of $\sum_{i=1}^S p_i^q$ ($0 \leq q \leq \infty$) their numbers equivalents are given by the formula

$$^qD = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)} \quad (1)$$

while D_β embodies the effective number of distinct communities or plots in the region, thus reconciling the notion of beta with compositional diversity. Jost (2007) further demonstrated that numbers equivalents allow the multiplicative decomposition of any diversity index D into two independent components, D_α and D_β that are free to vary independently and that completely determine D_γ .

Based on simulated data, Veech and Crist (2010) contest this result and argue that: “When evaluating the statistical dependence of alpha and beta diversity, it is important to remember that a third variable, gamma diversity, is involved [...]. This a priori knowledge of the value of gamma suggests that alpha and beta are not *conditionally independent*. Beta is completely determined from gamma and alpha. Procedurally, gamma and alpha are calculated first and then beta is calculated as either gamma – alpha or gamma/alpha. Alpha and beta would be conditionally independent [...] if the value of alpha did not determine the value of beta (or vice versa) given a known gamma. Each of the three variables, alpha, beta, and gamma are *pairwise independent*. This means that for each of the pairs [(alpha, beta), (alpha, gamma), and (beta, gamma)] neither variable would determine the other without knowing the value of the third variable not in the pair.”

Though Veech and Crist correctly note that, given the data, all the metrics are determined, exactly because they are calculated from the data, I cannot fully agree with their approach. The problem here is in recognizing what these metrics really are conditional upon. For instance, the “design” of a study (as opposed to its results) only constrains the total number of plots we sample, not the total number of species sampled or their occurrences. Consequently, *before* we get the data, alpha, beta, and gamma could potentially assume any values (within the constraints of their definitions; e.g., alpha must be less than or equal to gamma, and so forth). In the present debate, what seems to be of interest is the values that

beta can take, conditional on the value of alpha in the data. Given this objective, the only information we should use is the number of plots sampled (N), which is set before we have the data in hand, and the quantity we are conditioning on (alpha).

In this view, for species presences and absences, multiplicative beta can be expressed as

$$\beta = N / \bar{N}_i \quad (2)$$

where \bar{N}_i is the mean number of species presences in the N plots. This latter way of expressing beta also immediately tells us that in multiplicative diversity partitioning maximum beta is necessarily constrained by N such that $\beta \leq N$. Accordingly, a null model that first chooses the number of plots, then alpha at random, then beta at random within the possible values defined by Eq. 2, and then determines what gamma should be, would not have the correlations obtained by Veech and Crist (because they constrain beta based on alpha *and* gamma).

Also, gamma is in no way independent on alpha, as for a given N , gamma is constrained within the values $\bar{\alpha} \leq \gamma \leq N \times \bar{\alpha}$. This dependence of gamma on alpha is a basic component of the doubling property of Jost (2006, 2007) and of the replication principle of Ricotta (2008). For instance, both conditions require that, under some circumstances, there is a linear dependence of gamma on alpha. According to the doubling property, given two equally large and completely distinct species assemblages, each with diversity D , if these assemblages are combined, the diversity of the combined assemblages should be $2D$.

This semi-additive property is at the core of the independence between alpha and beta demonstrated by Jost (2007). Most diversity indices violate this property, but their numbers equivalents do not. Therefore, apart from species richness that represents its own numbers equivalent, we can confidently conclude that the multiplicative partitioning of numbers equivalents is the best possible choice for getting independent alpha and beta components; the next step will now consist in extending this partitioning scheme to diversity measures that incorporate information about the degree of ecological dissimilarity between species, such as, e.g., the Rao (1982) quadratic entropy. For a short review on such measures see Schmera et al. (2009) and references therein.

Yet, this is not the end of the history; as noted by one anonymous referee, independence of beta on alpha is not a good reason for “letting the tail of statistical convenience wag the dog of ecological inquiry.” In particular, different beta metrics are measuring different quantities (i.e. average number of species not observed for additive beta, or “effective number of communities” for multiplicative beta). Therefore, the key question we should ask of a beta metric is: does it measure the thing we are biologically interested in? If the metric has statistical properties that make patterns in beta easy to

analyze and interpret, so much the better. But if not, this is not necessarily a good reason to abandon it in favor of something statistically well behaved that is not actually the quantity we are most interested in. Rather, as covariances between statistics calculated from the same data ought to be something that can be handled by generating the appropriate statistical expectations (either via analytical probability theory or possibly by bootstrapping or Monte Carlo methods), we simply need to do the hard work of coming up with valid tests for patterns in that beta metric.

Finally, in spite of the many advantages offered by the multiplicative diversity decomposition of numbers equivalents, we should ask what is lost in transforming raw diversity measures to their numbers equivalents. Many authors have proposed a set of basic criteria that an index of diversity should meet to reasonably behave in ecological research (e.g., Patil and Taillie 1982, Routledge 1983, Wilson and Shmida 1984, Lande, 1996, Jost 2007). However, the usual outcome is that no single index can satisfy even the most basic of these criteria. This is because as diversity theory mirrors the intrinsically complex and nonlinear essence of ecological processes, it is also a fundamentally complex and nonlinear discipline. In this view, a desirable property of an ecologically meaningful (beta) diversity index is the so-called sum property. In simple terms, the diversity index needs to be decomposable into species-level patterns such that, given a diversity measure H that conforms to the sum property, the measure is decomposable into species-level patterns and the sum of single species diversities gives the pooled diversity of the species collection. That is, $H = \sum_{i=1}^S H_i$, where H_i is the contribution of species i to H .

In this way, the sum of single species diversities gives the pooled diversity of the species assemblage (see Ricotta et al. 2004). In a similar context, Patil and Taillie (1982) termed this property “dichotomy” because the diversity of species i would be unchanged if the other species were grouped into a single complementary category.

When dealing with beta diversity, a usual question to ask is which species contribute more to plot-to-plot heterogeneity? From Eq. 2, it is easily shown that for species presence and absence data, the contribution of species i to beta is proportional to the inverse of its number of presences in the N plots. Unfortunately, this simple result cannot be generalized to numbers equivalents. Due to the non-linearity of the transformation of raw diversities to numbers equivalents (see Eq. 1), these latter ones cannot be decomposed into single-species contributions. In this case, to capture the importance of single species or species groups in shaping the compositional heterogeneity of a given set of plots, different measures of beta diversity need to be used.

To conclude, though the Jost definition of beta behaves better than previous measures as concerns its independence on within-plot diversity, a perfect measure of beta diversity does not exist and none of the measures

proposed to date is entirely satisfactory. Therefore, as we do not leave in a perfect world, we are forced to use the type of beta diversity measure that is less inadequate to solve a specific problem.

ACKNOWLEDGMENTS

I thank one anonymous referee for very constructive comments on a previous draft of this paper.

LITERATURE CITED

- Anderson, M. J., K. E. Ellingsen, and B. H. McCauley. 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9:683–693.
- Izask, C., and R. G. Price. 2001. Measuring β -diversity using a taxonomic similarity index, and its relation to spatial scale. *Marine Ecology Progress Series* 215:69–77.
- Jost, L. 2006. Entropy and diversity. *Oikos* 113:363–375.
- Jost, L. 2007. Partitioning diversity into independent alpha and beta components. *Ecology* 88:2427–2439.
- Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology* 72:367–382.
- Lande, R. 1996. Statistics and partitioning of species diversity and similarity among multiple communities. *Oikos* 76:5–13.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75: 435–450.
- McArthur, R., H. Recher, and M. Cody. 1966. On the relation between habitat selection and species diversity. *American Naturalist* 100:319–332.
- Patil, G. P., and C. Taillie. 1982. Diversity as a concept and its measurement. *Journal of the American Statistical Association* 77:548–567.
- Rao, C. R. 1982. Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology* 21:24–43.
- Ricotta, C. 2007. A semantic taxonomy for diversity measures. *Acta Biotheoretica* 55:23–33.
- Ricotta, C. 2008. Computing additive beta-diversity from presence and absence scores: a critique and alternative parameters. *Theoretical Population Biology* 73:244–249.
- Ricotta, C., and S. Burrascano. 2009. Testing for differences in beta diversity with asymmetric dissimilarities. *Ecological Indicators* 9:719–724.
- Ricotta, C., A. Chiarucci, and G. C. Avena. 2004. Quantifying the effects of nutrient addition on community diversity of serpentine vegetation using parametric entropy of type α . *Acta Oecologica* 25:61–65.
- Routledge, R. D. 1983. Evenness indices: are any admissible? *Oikos* 40:149–151.
- Sarkar, S., and C. Margules. 2002. Operationalizing biodiversity for conservation planning. *Journal of Biosciences* 27(S2): 299–308.
- Schmera, D., T. Erös, and J. Podani. 2009. A measure for assessing functional diversity in ecological communities. *Aquatic Ecology* 43:157–167.
- Veech, J. A., and T. O. Crist. 2010. Diversity partitioning without statistical independence of alpha and beta. *Ecology* 91:1964–1969.
- Vellend, M. 2001. Do commonly used indices of β -diversity measure species turnover? *Journal of Vegetation Science* 12: 545–552.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- Whittaker, R. H. 1972. Evolution and measurement of species diversity. *Taxon* 21:213–251.
- Wilson, M. V., and A. Shmida. 1984. Measuring beta diversity with presence-absence data. *Journal of Ecology* 72:1055–1064.