Rarefaction of beta diversity

Carlo Ricotta[1,*] , Alicia T. R. Acosta[2], Giovanni Bacaro[3], Marta Carboni[2,4], Alessandro Chiarucci[5], Duccio Rocchini[6,7,8], Sandrine Pavoine[9]

[1]Department of Environmental Biology, University of Rome 'La Sapienza', Piazzale Aldo Moro 5, 00185 Rome, Italy.
[2]Department of Sciences, University 'Roma Tre', Viale Marconi 446, 00146 Roma, Italy. [3]Department of Life Sciences, University of Trieste, Via L. Giorgieri 10, 34127 Trieste, Italy. [4]Department of Biological Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON, M1C 1A4, Canada. [5]Department of Biological, Geological and Environmental Sciences, University of Bologna, Via Irnerio 42, 40126, Bologna, Italy. [6]Center Agriculture Food Environment, University of Trento, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy. [7]Center for Integrative Biology, University of Trento, Via Sommarive, 14, 38123 Povo (TN), Italy. [8]Fondazione Edmund Mach, Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy. [9]Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, 43 rue Cuvier, 75005 Paris, France.

* Correspondence author. E-mail address: carlo.ricotta@uniroma1.it

Abstract. Beta diversity has long been used to summarize the amount of variation in species composition among a set of N sampling units. However, while classical beta diversity provides an estimate of multiple-site dissimilarity among all sampling units, it is not informative on the changes

1

of multiple-site dissimilarity as a function of sampling effort. For gamma diversity, this pattern is usually represented as a species accumulation curve, which is the graph of the number of observed species when the number of plots varies from 1 to *N*. Here, we will show that species accumulation curves may also be used to summarize directional and non-directional beta diversity as a function of sampling effort. The behavior of the proposed measures of beta diversity is illustrated with one worked example on plant species in Mediterranean coastal vegetation. We believe this approach to the measurement of beta diversity provides a relevant contribution to summarize multiple-site dissimilarity as the result of a species turnover process, rather than as a static indicator. For directional species accumulation curves, the method, for which we provide a custom R function, further allows summarizing the spatial autocorrelation in species composition among plots along an a-priori defined spatial, temporal or environmental gradient.

Keywords: Community turnover; Directional and non-directional accumulation curves; Effective number of plots; Number of species; Spatial autocorrelation.

Introduction

Whittaker (1960, 1972) introduced the concept of beta diversity (the amount of variation in species composition among a set of sampling units) for linking mean local-scale diversity (or alpha diversity) to the diversity of the broader regional species pool (gamma diversity). Since then beta diversity has become a primary tool for connecting the spatial structure of species assemblages to ecological processes and there is a huge literature on its measurement and summarization (for a review, see e.g. Tuomisto 2010a, 2010b; Anderson et al. 2011). According to Whittaker's proposal, given a community composition matrix of *S* species in *N* plots (or sampling units), beta-diversity can be calculated as the ratio of the diversity of the pooled set of plots and mean species diversity within each plot, such that $\beta = \gamma/\alpha$. This metric of beta diversity summarizes multiple-site dissimilarity within the community composition matrix. However, it does not provide any

information on the changes of multiple-site dissimilarity when the number of plots varies from 1 to N. For gamma diversity, this variation is traditionally represented as a species accumulation curve, a graph of the number of observed species as a function of sampling effort or the area sampled (Gotelli and Colwell 2001). In this view, the species accumulation curve represents gamma diversity as the outcome of the sampling process rather than as a fixed summary statistic. A distinction is usually made between directional and non-directional accumulation curves: directional curves summarize how species accumulate along an a-priori defined spatial, temporal or environmental gradient. By contrast, non-directional curves measure turnover in community structure without reference to any specific gradient.

In this paper, we show that species accumulation curves may also be used for summarizing directional and non-directional beta diversity. While beta diversity measures have been developed for a long time, there is a knowledge gap on how multiple-site dissimilarity changes in relation to sampling effort. The present manuscript fills this gap and shows how species accumulation curves can be used to summarize beta diversity as a function of sampling effort. This can be very useful to researchers attempting to identify the necessary sampling effort to correctly characterize beta-diversity. We also show that comparing the directional beta diversity of a given set of plots with its non-directional counterpart provides a measure of the spatial autocorrelation of species composition among plots induced by the non-random species sorting along the ecological gradient.

The paper is organized as follows: first, we present an overview on directional and non-directional species accumulation curves and their associated beta diversity measures. Next, to show the behavior of the proposed measures, a worked example is used with data on plant species cover in 68 plots of Mediterranean coastal vegetation.

Methods

Non-directional species accumulation curves and beta diversity

Sample-based accumulation curves are constructed for a set of N equal-sized plots by plotting the number of detected species S when the number of sampled plots increases from 1 to N (Scheiner

3

83 2003; Scheiner et al. 2011). For non-directional curves, it is assumed that the position of the plots in

84 physical or environmental space is irrelevant. Therefore, accumulation curves are generally

85 constructed via permutation procedures by randomly resampling the N plots and then plotting the

86 average number of species found in 1, 2, ..., N plots (Gotelli and Colwell 2001; Colwell et al. 2004).

87 For standard resampling without spatial constraints, a mathematical expression for sample-based

88 accumulation curves was first derived by Shinozaki (1963) and then independently re-discovered by

89 several authors in the 1970s and then again in the 2000 (for a review, see Chiarucci et al. 2008). For

90 species incidence (presence and absence) data, the average number of species detected by sampling

91 all possible combinations of M equal-sized plots randomly and without replacement from a larger

92 set of N plots, is given by:

93

$$S_M = \sum_{i=1}^{S} \left[1 - \frac{\binom{N - N_i}{M}}{\binom{N}{M}} \right] = S - \sum_{i=1}^{S} \left[\frac{\binom{N - N_i}{M}}{\binom{N}{M}} \right]$$

94   (1)

95

96 where S_M is an estimator of the number of species in the M plots, S the total number of species in

97 the entire set of N plots $(N \geq M)$, and N_i the number of plots that contain species i. The binomial

98 ratio $\rho_{iM} = \binom{N - N_i}{M} \Big/ \binom{N}{M}$ is the probability that species i is not present in a randomly selected

99 sample of M plots. Therefore, $\rho_{iM} = 0$ for $M > N - N_i$.

100 Species accumulation curves have been used for comparing species richness in assemblages of

101 various sizes after rarefaction to a common level sampling effort or sample completeness (Gotelli

102 and Colwell 2001; Colwell et al. 2004; Chao and Jost 2012), and for extrapolating species richness

103 to a larger area sampled (Colwell et al. 2012; Chao et al. 2014, 2015). Eq. 1 is also used for

104 constructing individual-based accumulation curves relating the estimated number of species to the

105   number of sampled individuals. In this case, M and N represent the sampled number of individuals

106   and the overall population size, respectively (Hurlbert 1971; Simberloff 1972).

107   The estimated number of species S_M represents the gamma diversity in the M plots: $\gamma_M = S_M$.

108   By contrast, alpha diversity does not change with M being always equal to S_1. This is easily shown

109   considering that in Eq. 1 each plot appears an equal number of times for the calculation of S_M,

110   irrespective of the value of M, such that $\alpha_M = S_1$ (see Ricotta et al. 2012, Appendix S3). From S_M

111   and S_1 we can calculate an index of beta diversity for all values of M as (Chao et al. 2014):

112

$$\beta_M = \frac{S_M}{S_1} = \frac{\gamma_M}{\alpha_M}$$

113   (2)

114

115   According to Jost (2007), β_M can be interpreted as an 'effective number' of communities or

116   plots, the theoretical number of fully-distinct plots (no shared species) with mean species richness

117   equal to α_M needed to produce the given value of beta diversity (see also Wilson and Shmida

118   1984).

119   Like the original species accumulation curve, the curve of beta diversity built from the β_M

120   values is a concave and increasing function of M (proof in Appendix 1). Therefore, the graph of β_M

121   vs. M shows how the effective number of plots accumulates with sampling effort. Note however

122   that β_M summarizes multiple-site dissimilarity among plots in the range [1, M]. This range is

123   intuitively related to the very meaning of β_M as the effective number of maximally distinct plots.

124   For a community composition matrix of S species x N plots, if all plots are compositionally

125   identical, then $\beta_M = 1$ irrespective of the value of M. In that case, since $\alpha_M = \gamma_M$, one single

126   'effective' plot is enough to produce the observed value of β_M. By contrast, if the plots do not have

any species in common, $\beta_M = M$, meaning that M maximally distinct plots with mean species

richness α_M are needed to obtain the observed value of β_M. Therefore, since the range of β_M depends on M, the effective numbers of plots cannot be used for comparing beta diversity for different values of M. To remove this dependence, beta must first be rescaled onto the unit interval (Chao et al. 2012; Ricotta and Pavoine 2015). The simplest way to rescale a given quantity X

between zero and one is to use the linear transformation $(X - X_{\min})/(X_{\max} - X_{\min})$. This

transformation gives a normalized measure of non-directional beta diversity (Harrison et al. 1992;

Jost 2007):

$$\beta_\eta = \frac{\beta_M - 1}{M - 1}$$
(3)

which can be interpreted as the average proportion of non-overlapping species in the M plots. When

$M = 2$, it reduces to the classic Sørensen dissimilarity coefficient. The normalized measure β_η summarizes beta diversity in a fixed range of values ($0 \le \beta_\eta \le 1$) that is independent of M. That is, the minimum and maximum values that β_η can take are fixed constants that do not depend on the number of accumulated plots M. This lack of systematic constraints between β_η and M ensures that β_η can be used for comparing multiple-site dissimilarity at different values of M. For additional normalized dissimilarity measures of beta diversity, see Chao and Chiu (2016).

Normalization removes the (undesired) dependence of the range of β_η on M. However, it does not eliminate the dependence of beta on sampling effort, which is the main assumption of this paper. Therefore, for a given set of plots N, the variation in species composition β_η decreases as a function of the number of plots sampled (proof in Appendix 1). Due to this decreasing pattern,

149   calling the graph of β_η vs. M an 'accumulation curve' is not the most appropriate choice.

150   Therefore, like in Wilson and Shmida (1984), we use for these curves the more neutral term of

151   'community turnover' curves.

152 ***Directional species accumulation curves and beta diversity***

153   The examination of turnover in species composition along a gradient requires explicit

154   consideration of the order of plots in physical or environmental space (Vellend 2001). To this end,

155   given *N* plots ordered along a predefined gradient of interest, starting from one end of the gradient

156   we can construct a directional species accumulation curve by sequentially adding the plots one after

157   the other along that gradient and plotting the cumulative number of species found in 1, 2, ..., *N*

158   plots.

159   Because each directional accumulation curve is case-specific and depends on the distribution of

160   plots along a user-defined gradient, it cannot be derived theoretically and a mathematical expression

161   for this curve does not exist. Therefore, the values of alpha and gamma diversity along the gradient

162   need to be calculated separately for each curve based on the specific order of plots. Beta diversity

163   for all values of *M* can then be calculated as above:

164

165    $$\vec{\beta}_M = \frac{\vec{\gamma}_M}{\vec{\alpha}_M}$$ (4)

166   and

167    $$\vec{\beta}_\eta = \frac{\vec{\beta}_M - 1}{M - 1}$$ (5)

168

169   where the arrows on alpha, beta and gamma denote the directional nature of the turnover process.

170   Unlike a non-directional accumulation curve, in the directional case the value of $\vec{\alpha}_M$ (i.e. the mean

171   number of species in the *M* plots) is not constant. Rather, like $\vec{\gamma}_M$, it depends on the order of the

172   plots along the gradient. Accordingly, the order in which individual plots are added affects the

173   shape of the resulting accumulation curve, such that directional beta $\vec{\beta}_\eta$ is not necessarily a

174   monotonically decreasing function of *M*.

To complete the picture, 'constrained' species accumulation curves which account for the arrangement of plots in physical or environmental space were introduced by Scheiner (2003, his Type IIIA curves) and first explored by Chiarucci et al. (2009). Order-free accumulation curves tend to overestimate the rate of increase in species richness because they ignore the autocorrelation of species composition among the sampling units. In contrast, constrained accumulation curves describe the increase in species richness as a function of both the sampling effort and the autocorrelation in species composition among plots.

In building a constrained accumulation curve, adjacent sampling units are combined step by step using the spatial or environmental proximity among plots as a constraining factor. Given a set of N plots, for each plot a directional species accumulation curve is constructed by sequentially adding the first, second, ..., k-th nearest neighbor. This procedure is repeated for all N plots, generating N directional accumulation curves from which a mean constrained curve is calculated (Chiarucci et al. 2009; Bacaro et al. 2016). The resulting curve is thus an intermediate solution between a non-directional accumulation curve in which all possible combinations of 1, 2, ..., M plots are used for building the curve and a pure directional curve in which the N plots are ordered along a single spatial or environmental gradient.

Denoting by $\vec{\gamma}_{jM}$ and $\vec{\alpha}_{jM}$ the gamma and alpha diversity of the j-th directional community turnover curve ($j = 1, 2, ..., N$), the constrained beta diversity $\overline{\beta}_M$ for a sequence of M plots can be calculated in analogy to Eq. 2 as the ratio of mean gamma to mean alpha:

$$\overline{\beta}_M = \frac{\overline{\gamma}_M}{\overline{\alpha}_M} = \frac{\sum_{j=1}^{N} \vec{\gamma}_{jM} / N}{\sum_{j=1}^{N} \vec{\alpha}_{jM} / N}$$

(6)

such that

$$\overline{\beta}_\eta = \frac{\overline{\beta}_M - 1}{M - 1}$$

(7)

199   The reasons for calculating beta as $\bar{\beta}_M = \bar{\gamma}_M / \bar{\alpha}_M$ instead of $\bar{\beta}_M = 1/N \sum_{j=1}^{N} \vec{\gamma}_{jM} / \vec{\alpha}_{jM}$ were

200   discussed by Chao et al. (2014) in the context of the rarefaction of Hill numbers, and the reader is

201   addressed to this paper for additional details.

202   ==The plot of $\bar{\beta}_\eta$ vs. *M* describes how 'constrained' beta diversity varies with increasing sampling==

203   ==effort.== For each value of *M*, the difference between $\bar{\beta}_\eta$ (or $\vec{\beta}_\eta$) and the non-directional beta reflects

204   the amount of autocorrelation in community composition among plots. This is because directional

205   beta summarizes turnover in community structure along an a-priori ecological gradient and is thus

206   directly influenced by the similarity in community composition among plots. By contrast, non-

207   directional turnover is calculated with permutation procedures which ignore any autocorrelation in

208   community structure. Therefore, a measure of autocorrelation for directional beta diversity can be

209   calculated as the normalized difference between directional and non-directional beta:

210

$$\bar{\beta}_\phi = \frac{\beta_\eta - \bar{\beta}_\eta}{\beta_\eta + \bar{\beta}_\eta}$$

211   (8)

212

213   with $-1 \le \bar{\beta}_\phi \le 1$. Since the compositional dissimilarity among plots will often increase more slowly

214   when sampling units are close to each other, usually $\beta_\eta > \bar{\beta}_\eta$ and $\bar{\beta}_\phi > 0$. ==Compositional==

215   ==autocorrelation occurs due to two processes: localized dispersal mechanisms and autocorrelation in==

216   ==environmental conditions.== We therefore expect greater habitat homogeneity to lead to lower

217   compositional heterogeneity among adjacent plots. If adjacent plots are compositionally very

218   similar to each other, the difference $\beta_\eta - \bar{\beta}_\eta$ will be high, such that the normalized difference $\bar{\beta}_\phi$

219   will be high too. By contrast, if autocorrelation in community composition among adjacent plots is

220   low, $\bar{\beta}_\eta$ will approach its random expectation β_η and $\bar{\beta}_\phi$ will be close to zero.

Worked example

We used vegetation plots sampled on recent (Holocene) coastal dunes occupying a narrow strip (< 150 m) of about 16 km length along the seashore on the Tyrrhenian coast of central Italy (Figure 1). The vegetation of the dune profile follows a compressed zonation of habitats along a clearly defined sea-inland environmental gradient that drives considerable turnover in species composition and functioning over short distances: from the pioneer communities of the upper beach to the woody communities (Mediterranean macchia and evergreen forests) of the inland fixed dunes (Acosta et al. 2003).

Vegetation sampling was undertaken in spring (April-May) of 2006. Randomly generated GPS coordinates were used to define the sampling location of 2 x 2m vegetation plots in the study area. In each plot, the cover of all vascular plant species was recorded using a 10% interval rank scale. A total of 68 plots were sampled and 61 species recorded (Appendix 2). For the 68 plots of coastal dune vegetation used in this study we produced:

i) A non-directional community turnover curve of β_η vs. M constructed according to Eq. (1-3).

ii) A spatially-constrained turnover curve of $\bar{\beta}_\eta$ vs. M (hereafter distance-based turnover curve), in which all plots are ordered according to their distance from the seashore. Carboni et al. (2011) showed that for coastal vegetation, distance from the seashore is related to a directional stress gradient of wind and soil parameters, ranging from the exposed conditions of the upper beach to the more sheltered conditions of the back dune. Therefore, plant communities located closer to the sea are generally subject to higher environmental stress and disturbance.

iii) A functionally-constrained turnover curve (hereafter trait-based turnover curve), in which all plots are ordered along an ecological gradient according to the functional characters of the most abundant species. For building this functional turnover curve, we defined an a-priori directional

11

gradient in functional trait space. To do so, we used a set of plant functional traits measured for a subset of 42 dominant species. This subset of species was chosen by selecting the most common and abundant species within each habitat along the sea-inland gradient, that collectively account for ~80% of the standing live biomass (see Carboni et al. 2013 for details). This threshold has been shown to ensure a satisfactory description of overall community properties (Pakeman and Quested 2007).

Three quantitative life-history traits were chosen based on their relevance to the functional ecology of plants in coastal dune environments (Carboni et al. 2016): plant height (cm), leaf dry mass content (LDMC; $mg·g^{-1}$) and seed mass (mg). These traits provide a good representation of the species global spectrum of form and function (Diáz et al. 2016) and are also correlated with the directional stress gradient of wind and soil parameters which shape the zonation of the coastal dune vegetation (Carboni et al. 2011, 2013). Because of environmental sorting of species according to habitat preferences along the sea-inland stress gradient, we expect that specific species combinations with particular suites of traits (e.g. short annual species with succulent leaves on the upper beach) should gradually be replaced along the gradient by other dominant species with different traits (e.g. tall perennial species on the main dune ridge). Carboni et al. (2013) showed that there was a significant relationship between sea-inland stress and community functional composition.

First, we calculated the Bray-Curtis pairwise dissimilarity among species based on the selected functional traits. Before calculations, traits were square root-transformed when necessary and scaled to unit length. We next calculated the pairwise functional dissimilarity among plots according to Pavoine and Ricotta (2014, Eq. 3) The resulting plot-to-plot functional distance matrix was then used for constructing the constrained turnover curve of $\overline{\beta}_\eta$ vs. M. All curves were constructed using a custom R function available as an electronic appendix to this paper (Appendix 3).

Results

The directional and non-directional species accumulation curves of γ_M and $\overline{\gamma}_M$ vs. M, together

with the corresponding community turnover curves of β_η and $\overline{\beta}_\eta$ vs. M. are shown in Figure 2 and

3, respectively. Figure 4 shows the amount of autocorrelation of directional beta diversity $\overline{\beta}_\phi$ as a

function of sampling effort. As expected, the shape of the curves differs depending on which

methods are used to construct the curves, thus providing a useful tool to characterize biodiversity

patterns and to explore the relevance of selected functional traits or environmental variables in

shaping community structure. Due to the very high autocorrelation of plots in functional space, the

trait-based accumulation curve increases less steeply than the corresponding non-directional curve.

That is, functionally similar plots are on average compositionally more similar to each other than

expected by chance alone. As shown in Figure 3, the lower rate of species accumulation of the trait-

based curve compared to the random non-directional curve results in a lower rate of turnover in

community structure and thus in lower beta diversity.

On the other hand, the plot of the distance-based accumulation curve is very similar to that of the

non-directional curve. That is, at least for our case study, distance from the seashore is a less

effective indicator of the species sorting than trait-based differences. This is probably due to the

effects of natural and anthropogenic disturbance, which tend to modify the natural zonation of

vegetation along the dune profile giving rise to a complex mosaic of fragmented and intermingled

communities. As shown in Figure 4, for the distance-based curve, the autocorrelation of directional

beta diversity is much lower than that of the trait-based curve and becomes negligible for $M > 45$,

meaning that after this threshold the compositional turnover of the constrained distance-based curve

does not substantially differ from random expectation. This example clearly shows how different

degrees of autocorrelation in geographical or functional space can affect the estimates of gamma

and beta diversity of community turnover curves.

Discussion

In this paper, we showed that community turnover curves provide a tool for summarizing various types of directional and non-directional beta diversity as a function of sampling effort. Azovsky (2011) separated two conceptually different curves: species-area and species-sampling effort relationships. The present manuscript deals mostly with species-sampling effort relationship, although directional analysis of real gradients with ordered plots goes a step further simple species-sampling effort relationships *sensu* Azovsky (2011). The same approach can be extended for constructing incidence- or abundance-based community turnover curves either mathematically or with resampling methods for any taxonomic, functional or phylogenetic diversity metric (Walker et al. 2008; Ricotta et al. 2012; Chao et al. 2014, 2015). The resulting measures of beta diversity can then be compared among habitats sampled with different effort after rarefaction to the sample size of the smaller datasets. For examples, see e.g. Bacaro et al. (2016). In addition, if directional or non-directional beta diversity is calculated for a nested series of plots of different size, the resulting turnover curves can be used to analyze how beta diversity scales with local sample area, thus extending our results to species-area relationships (see Palmer and White 1994; Mokany et al. 2013).

Finally, turnover curves can also be used for calculating the amount of autocorrelation in community composition for directional beta diversity compared to its non-directional counterpart. Species are almost never randomly distributed, either due to heterogeneity of environmental factors or to non-random dispersal mechanisms. Therefore, autocorrelation is a crucial aspect of directional turnover curves and appropriate methods should be used in diversity studies if there is evidence of significant clustering of species in physical or environmental space (Legendre 1993; Fortin and Dale 2009; Bacaro et al. 2016).

Previous methods to assess multivariate spatial community structure include distance-decay plots (Nekola and White 1999; Qian and Ricklefs 2007), multivariate variograms (Wagner 2003) and Mantel correlograms (Borcard and Legendre 2012). Our proposal adds a new tool to the ecologist

toolbox as it allows to summarize autocorrelation in species composition along spatial, temporal or environmental gradients from community turnover curves.

Directional species accumulation curves have been typically used to summarize the decrease in species accumulation due to spatial autocorrelation along an environmental gradient compared to non-directional curves (e.g. Chiarucci et al. 2009; Bacaro et al. 2016). The same 'directional' approach can also be used to maximize the rate of species accumulation with a given effort so as to improve the efficiency of species inventories (Palmer et al. 2002; Rocchini et al. 2005). Because the sampling order of plots determines the slope of the accumulation curve, an optimal strategy to maximize species accumulation consists of using external data to choose the order of plots so that the resulting accumulation curve is the steepest possible (Rocchini et al. 2005). That curve is a special kind of directional turnover curve, which accumulates species at a faster rate compared to the corresponding random non-directional curves. As a consequence of this increased accumulation rate, the normalized difference between the non-directional and directional beta diversity measures takes negative values showing a general tendency towards a negative correlation in the compositional similarity of subsequent plots. This illustrates how our proposed metric of autocorrelation based on turnover curves can also be used as a synthetic criterion for defining optimization strategies in vegetation sampling, for example to show at what distance species composition is no longer autocorrelated. Thus, one can use this critical distance between sampling units to avoid spending time and resources on redundant information.

The idea of calculating an index of beta diversity from species accumulation curves has already been proposed in the past. Ricotta et al. (2002) suggested calculating beta diversity as the slope of the linearized species accumulation curve in semilogarithmic space according to Gleason (1922); Crist and Veech (2006) proposed an additive model of diversity decomposition for which beta diversity is obtained as the difference in species richness between the last and first points of the species accumulation curve: $\beta = \gamma - \alpha = S_N - S_1$. The main difference between these previous works and ours is that we do not treat beta diversity as a fixed quantity: the graph of beta vs. *M*

15

shows how community turnover varies when M increases from 1 to N. Therefore, using accumulation curves, multiple-site dissimilarity or beta diversity is represented as the result of a species turnover process, rather than as a static indicator. For $M = N$, β_M reduces to the traditional Whittaker's (static) beta diversity for a set of N sampling units $\beta = \gamma_N / \alpha_N$.

We also showed that using the classical multiplicative formula $\beta_M = \gamma_M / \alpha_M$, a measure of beta is obtained which is easily interpretable in terms of 'effective number of plots'. However, due to the constraint $1 \leq \beta_M \leq M$ we cannot directly use this measure for comparing beta diversity for different numbers of plots. To remove the dependence on M, we can use the normalized measure $\beta_\eta = (\beta_M - 1)/(M - 1)$. This scaled version of beta diversity can be used for comparing community turnover for different values of M.

In principle, because of the fundamental constraint $M\alpha_M \geq \gamma_M \geq \alpha_M$ (Chao et al. 2012), a similar scaling can also be used for defining a normalized measure of gamma diversity $\gamma_\eta = (\gamma_M - \alpha_M)/(M\alpha_M - \alpha_M)$. However, since $\gamma_M = \alpha_M \times \beta_M$, a little algebra shows that $\gamma_\eta = \beta_\eta = (\beta_M - 1)/(M - 1)$, meaning that there is no fundamental difference between normalized beta diversity and normalized gamma diversity. For a given number of plots, the rescaled version of gamma diversity is *per se* a measure of community turnover, such that normalized beta diversity can be also expressed in terms of the relationship between gamma diversity and its extreme values: if $\gamma_M = \alpha_M$ then $\beta_\eta = 0$; At the other extreme, if $\gamma_M = M\alpha_M$ then $\beta_\eta = 1$. Therefore, the scaled versions of beta diversity and gamma diversity lead to the same normalized measure of species turnover.

To conclude, by its very nature, community ecology usually deals with high-dimensional multivariate data and calls for summarizing methods and indices. However, by condensing the structure of a multidimensional community composition matrix into a single indicator, information is inevitably lost, such that a variety of indices is needed to answer a variety of ecological questions

16

371 (Pavoine 2016). Therefore, accurate knowledge of the basic properties of diversity measures

372 becomes a necessary condition for comparing different habitats in ecologically meaningful ways.

373 Given this, relying on metrics standardized for sampling effort can provide better means for

374 identifying the most relevant biotic and abiotic drivers of community composition in different

375 habitats and taking appropriate management decisions.

376

380

381 **Supporting Information**

382 **Appendix 1.** Proofs.

383 **Appendix 2.** Composition matrix and functional trait values of the dominant species for the coastal

384 dune communities of central Italy.

385 **Appendix 3.** R function for the calculation of directional and non-directional accumulation curves,

386 beta diversity and autocorrelation as a function of sampling effort M.

References

Acosta, A.T.R., Stanisci, A., Ercole, S., Blasi, C. (2003) Sandy coastal landscape of the Lazio region (central Italy) Phytocoenologia 33: 715–726.

Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B., Stegen, J.C., Swenson, N.G. (2011) Navigating the multiple meanings of β–diversity: a roadmap for the practicing ecologist. Ecol. Lett. 14: 19–28.

Azovsky, A.I. (2011) Species-area and species-sampling effort relationships: disentangling the effects. Ecography 34: 18–30.

Bacaro, G., Altobelli, A., Cameletti, M., Ciccarelli, D., Martellos, S., Palmer, M.W., Ricotta, C., Rocchini, D., Scheiner, S.M., Tordoni, E., Chiarucci, A. (2016) Incorporating spatial autocorrelation in rarefaction methods: Implications for ecologists and conservation biologists. Ecol. Indic. 69: 233–238.

Borcard, D., Legendre, P. (2012) Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. Ecology 93: 1473–1481.

Carboni, M., Acosta, A.T.R., Ricotta, C. (2013) Are differences in functional diversity among plant communities on Mediterranean coastal dunes driven by their phylogenetic history? J. Veg. Sci. 24: 932–941.

Carboni, M., Santoro, R., Acosta, A.T.R. (2011) Dealing with scarce data to understand how environmental gradients and propagule pressure shape fine-scale alien distribution patterns on coastal dunes. J. Veg. Sci. 22: 751–765.

Carboni, M., Zeleny, D., Acosta, A.T.R. (2016) Measuring ecological specialization along a natural stress gradient using a set of complementary niche breadth indices. J. Veg. Sci. 27: 892–903.

Chao, A., Chiu, C.-H. (2016) Bridging the variance and diversity decomposition approaches to beta diversity via similarity and differentiationmeasures. Methods Ecol. Evol. 7: 919–928.

18

412 Chao, A., Chiu, C.-H., Hsieh, T.C. (2012. Proposing a resolution to debates on diversity
413   partitioning. Ecology 93: 2037–2051.

414 Chao, A., Chiu, C.-H., Hsieh, T.C., Davis, T., Nipperess, D.A., Faith, D.P. (2015) Rarefaction and
415   extrapolation of phylogenetic diversity. Methods Ecol. Evol. 6: 380–388.

416 Chao, A., Gotelli, N.J., Hsieh, T.C., Sander, E., Ma, K.H., Colwell, R.K., Ellison, A.M. (2014)
417   Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in
418   species diversity studies. Ecol. Monogr. 84: 45–67.

419 Chao, A., Jost, L. (2012) Coverage–based rarefaction and extrapolation: standardizing samples by
420   completeness rather than size. Ecology 93: 2533–2547.

421 Chiarucci, A., Bacaro, G., Rocchini, D., Fattorini, L. (2008) Discovering and rediscovering the
422   sample–based rarefaction formula in the ecological literature. Community Ecol. 9: 121–123.

423 Chiarucci, A., Bacaro, G., Rocchini, D., Ricotta, C., Palmer, M., Scheiner, S. (2009) Spatially
424   constrained rarefaction: incorporating the autocorrelated structure ofbiological communities into
425   sample–based rarefaction. Community Ecol. 10: 209–214.

426 Colwell, R.K., Chao, A., Gotelli, N.J., Lin, S.-Y., Mao, C.X., Chazdon, R.L., Longino, J.T. (2012)
427   Models and estimators linking individual–based and sample–based rarefaction, extrapolation and
428   comparison of assemblages. J. Plant. Ecol. 5: 3–21.

429 Colwell, R.K., Mao, C.X., Chang, J. (2004) Interpolating, extrapolating, and comparing incidence–
430   based species accumulation curves. Ecology 85: 2717–27.

431 Crist, T.O., Veech, J.A. (2006) Additive partitioning of rarefaction curves and species–area
432   relationships: unifying α–, β– and γ–diversity with sample size and habitat area. Ecol. Lett. 9:
433   923–932.

434 Diáz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M.,
435   Wirth, C., Prentice, I.C., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles,
436   A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Wright, S.J., Sheremet'ev, S.N., Jactel,
437   H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S.,

Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D., Gorné, L.D. (2016) The global spectrum of plant form and function. Nature 529: 167–171.

Fortin, M.J., Dale, M.R.T. (2009) Spatial autocorrelation in ecological studies: a legacy of solutions and myths. Geogr. Anal. 41: 392–397.

Gleason, H.A. (1922) On the relation between species and area. Ecology 3: 158–162.

Gotelli, N.J., Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4: 379–391.

Harrison, S., Ross, S.J., Lawton, J.H. (1992) Beta diversity on geographic gradients in Britain. J. Animal Ecol. 61: 151–158.

Hurlbert, S.H. (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology 52: 577–586.

Jost, L. (2007) Partitioning diversity into independent alpha and beta components. Ecology 88: 2427–2439.

Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74: 1659–1673.

Mokany, K., Jones, M.M., Harwood, T.D. (2013) Scaling pairwise β-diversity and α-diversity with area. J. Biogeogr. 40: 2299–2309.

Nekola, J.C., White, P.S. (1999) The distance decay of similarity in biogeography and ecology. J. Biogeogr. 26: 867–878.

Pakeman, R.J., Quested, H.M. (2007) Sampling plant functional traits: what proportion of the species need to be measured? Appl. Veg. Sci. 10: 91–96.

Palmer, M.W., Earls, P., Hoagland, B.W., White, P.S., Wohlgemuth, T. (2002) Quantitative tools for perfecting species lists. Environmetrics 13: 121–137.

Palmer, M.W., White, P.S. (1994) Scale dependence and the species-area relationship. Am. Nat. 144: 717–740.

Pavoine, S., Ricotta, C. (2014) Functional and phylogenetic similarity among communities, Methods Ecol. Evol. 5: 666–675.

Pavoine, S. (2016) A guide through a family of phylogenetic dissimilarity measures among sites. Oikos 125: 1719–1732.

Qian, H., Ricklefs, R.E. (2007) A latitudinal gradient in large–scale beta diversity for vascular plants in North America. Ecol. Lett. 10: 737–744.

Ricotta, C., Pavoine, S. (2015) A multiple–site dissimilarity measure for species presence/absence data and its relationship with nestedness and turnover. Ecol. Indic. 54: 203–206.

Ricotta, C., Pavoine, S., Bacaro, G., Acosta, A.T.R. (2012) Functional rarefaction for species abundance data, Methods Ecol. Evol. 3: 519–525.

Ricotta, C., Carranza, M.L., Avena, G.C. (2002) Computing β–diversity from species-area curves. Basic Appl. Ecol. 3: 15–18.

Rocchini, D., Andreini Butini, S., Chiarucci, A. (2005) Maximizing plant species inventory efficiency by means of remotely sensed spectral distances. Global Ecol. Biogeogr. 14: 431–437.

Scheiner, S.M. (2003) Six types of species–area curves. Global Ecol. Biogeogr. 12: 441–447.

Scheiner, S.M., Chiarucci, A., Fox, G.A., Helmus, M.R., McGlinn, D.J., Willig, M.R. (2011) The underpinnings of the relationship of species richness with space and time. Ecol. Monogr. 81: 195–213.

Shinozaki, K. (1963) Note on the species area curve. Proceedings of the 10th Annual Meeting of Ecological Society of Japan, 5 (in Japanese).

Simberloff, D. (1972) Properties of the rarefaction diversity measurement. Am. Nat. 106: 414–418.

Tuomisto, H. (2010a) A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. Ecography 33: 2–22.

Tuomisto, H. (2010b) A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecography 33: 23–45.

Vellend, M. (2001) Do commonly used indices of β-diversity measure species turnover? J. Veg. Sci. 12: 545–552.

489 Wagner, H. H. (2003) Spatial covariance in plant communities: integrating ordination, variogram

490   modeling, and variance testing. Ecology 84:1045–1057.

491 Walker, S.C., Poos, M.S., Jackson, D.A. (2008) Functional rarefaction: estimating functional

492   diversity from field data. Oikos 117: 286–296.

493 Whittaker, R.H. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. Ecol Monogr

494   30: 279–338.

495 Whittaker, R.H. (1972) Evolution and measurement of species diversity. Taxon 21: 213–251.

496 **Figure captions**

497

498 **Figure 1.** Location of the study area.

499

500 **Figure 2.** Directional and non-directional species accumulation curves $\overline{\gamma}_M$ and γ_M vs. M. for the

501 coastal dune communities of central Italy.

502

503 **Figure 3.** Directional and non-directional community turnover curves $\overline{\beta}_\eta$ and β_η vs. M for the

504 coastal dune communities of central Italy.

505

506 **Figure 4.** Plot of the amount of autocorrelation of directional beta diversity $\overline{\beta}_\phi$ as a function of

507 sampling effort M for the coastal dune communities of central Italy.